International Series in
Operations Research & Management Science

Chung-Yee Lee
Qiang Meng *Editors*

Handbook of Ocean Container Transport Logistics

Making Global Supply Chains Effective

Springer

International Series in Operations Research & Management Science

Volume 220

For further volumes
http://www.springer.com/series/6161

Chung-Yee Lee • Qiang Meng
Editors

Handbook of Ocean Container Transport Logistics

Making Global Supply Chains Effective

Springer

Editors
Chung-Yee Lee
Dept. of Industrial Engineering & Logistics Mgmt
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon
Hong Kong SAR

Qiang Meng
Dept. of Civil
& Environmental Engineering
National University of Singapore
Singapore
Singapore

ISSN 0884-8289 ISSN 2214-7934 (electronic)
ISBN 978-3-319-11890-1 ISBN 978-3-319-11891-8 (eBook)
DOI 10.1007/978-3-319-11891-8
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014953931

© Springer International Publishing Switzerland 2015
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

As international trade continues to grow rapidly and supply chains become more globalized, many operations have been outsourced and moved offshore. About 90 % of the international trade volume was facilitated by ocean transportation. If we assume that half of China's exports by value in 2013 were moved by ocean transport and on average it takes one month for the goods to reach the consignee, then about 85B\$[1] worth of goods would have been caught up during the transportation. Due to environmental and bunker cost considerations, international air cargo transport has been reduced from 3.1 % in 2010 to 1.7 % in 2013[2], with a shift towards more economic transportation modes, especially ocean transport. Hence, ocean transport logistics has played a very significant role in global supply chains.

Although ocean transport logistics has been well studied in maritime economics and operations research/management science, many important issues have yet to receive the attention they deserve. In this book, we reveal the interaction among parties along the chain, including shippers, terminal operators and line carriers. We examine the impact of ocean transport logistics on global supply chains and address many important topics to shed new light on the subject.

This book is organized into three parts. The first part talks about the innovative development of terminal operators and the competition they face. The second part delves into the many tactical and operational aspects of managing shipping liners, including empty container repositioning, disruption management, slow steaming, bunker purchasing, ship route schedule, and transport network design, and evaluates their corresponding challenges and opportunities. The third part studies the impact of ocean logistics transport on global supply chains. The 18 chapters of the book all highlight the immediate effect of ocean transport logistics on global supply chains.

---

[1] http://www.funggroup.com/eng/knowledge/research/ChinaTradeQuarterly1Q13.pdf.

[2] International Air Transport Association.

Part I: Container Terminal Operation: Innovations, Trends, Competition and Business Models

According to a survey by Notteboom[3], about 65 % of delay in ocean shipping is due to port congestion. Clearly, terminal operation efficiency is crucial to improving ocean shipping reliability. Due to bunker and emission cost reduction considerations, carriers tend to use huge vessels (capacity up to 18,000 TEU) and also adopt slow streaming. On the other hand, the strong demand for on time and/or fast delivery from shippers forces the carrier to cut the turn-around time in the terminal to allow time for slow steaming. All these have put huge pressure on terminal operators to improve their efficiency and have also made terminal competition fiercer than ever before. Hence, in Part I, we have five chapters that examine the innovations, trends, competition and business models of container terminal operations.

In the **first chapter**, Jiang, Chew and Lee study innovative container terminal designs. They first examine the issue of how to measure the port connectivity by proposing a new connectivity framework from a network perspective that can be used to generate a quantifiable measure of port connectivity. They also discuss the management of storage yards in transshipment ports. They further discuss innovative terminal designs that can serve as potential solutions for transshipment ports. Instead of using AGVs or ALVs to transport the containers in automated container terminals (ACTs), they introduce two innovative ACT systems: the “frame bridge system” designed by Shanghai Zhenhua Heavy Industries Co. Ltd., and the “GRID system” designed by BEC Industries. These revolutionary ideas aim to achieve a quantum leap in handling efficiency and productivity to support future shipping in an economically and environmentally sustainable manner.

In Chap. 2, Kim and Lee study the current trends and future challenges of container terminal operations. They review various planning and control activities in container terminals and define decision-making problems for operation planning and control. They also discuss new trends in the technological development for each decision-making process. They then describe the functions of terminal operation systems (TOS), the software used to implement the decision-making processes. Commercial TOS, including Navis SPARCS N4, CATOS, Mainsail Vanguard, TOPS, and OPUS, plus two famous noncommercial TOS, PortNet for PSA and nGen for HIT, are introduced and compared. Finally, they highlight recent trends of TOS responding to changes in the technological and market environment.

In Chap. 3, Notteboom and de Langen report an up-to-date and detailed analysis of the dynamics of the European container port system—the second most important container port system in the world. Their discussion and conclusion section summarizes very nicely their findings and also provides clear insights into the current drivers of the container port competition in Europe. They identify a number of key

---

[3] Notteboom, T.E. 2006. The time factor in liner shipping services. Maritime Economics & Logistics, 8 (1), 19–39.

success factors, including capacity, proximity to a hinterland with strong cargo generating and receiving capacities, access to sea and hinterland, strong sea and land connectivity, port and terminal efficiency, right pricing and a supply chain approach.

In Chap. 4, Lee and Lam examine whether major Asian ports have evolved into fifth generation ports (5GP) or if they remain fourth generation ports (4GP) to this day. They use the revised concept of 5GP to evaluate inter-port competition among four Asian ports—Shanghai, Singapore, Hong Kong and Busan—in a comprehensive way to reflect the cross-sectional, longitudinal and horizontal aspects of the port evolution. A novel approach with an empirical test combining a description method and a quantitative method is employed to study port competition and competitiveness.

In Chap. 5, Lu and Chang study the selection of a business model for container terminal operations. Recently, the Taiwan International Ports Corporation (TIPC) was set up to replace the former port authority of Kaohsiung, Keelung, Taichung and Haulin in Taiwan. They use TIPC as a case study to empirically identify the crucial criteria for choosing a business model for container terminal operations. An analytical hierarchy process approach was adopted to assess the relative importance of these criteria. Their results show that benefit and operational capability are the two most important criteria.

Part II: Shipping Liners: Tactical and Operational Management

According to a survey by Merge Global[4], the biggest portion (around 50 %) of the revenue in the whole ocean transport logistics service provider chain goes to carriers. There are many important issues in the shipping liner industry that are worth studying. In Part II, we report some studies on tactical and operational management issues, including empty container repositioning, disruption management, bunker purchasing, ship route schedule, slow steaming, and transport network design.

In Chap. 6, Song and Dong provide a comprehensive and critical survey on empty container repositioning for container shipping liners. After analyzing the main reasons for empty container repositioning operations, they provide a literature review with an emphasis on modeling empty container reposition problems from the network perspective. They then discuss possible solutions to the empty container repositioning problems from the logistics channel perspective followed by solutions from the methodical modeling technique perspective. Finally, they present two specific models aiming to tackle the empty container repositioning problems in stochastic dynamic environments considering both laden and empty container management.

In Chap. 7, Tsang and Mak further formulate the empty container repositioning problem for liner shipping as a multistage stochastic programming problem. Their model specifically handles the stochastic nature of demand and long transportation lead time. As they are able to reformulate the computationally intractable stochastic

---

[4] Merge Global. 2008. Insomnia: Why challenges facing the world container industry make for more nightmares than they should. American Shipper, July, 68–85.

program into a tractable cone program, a commercial solver can be used to find a solution. They also demonstrate that the robust model outperforms other simple policies.

In the shipping liner business, the route schedule is usually planned, fixed and announced either three or six months in advance, and then at the operational level, the vessel will stick to the schedule as closely as possible. But vessels will often encounter unexpected disruptions, such as port congestion, severe weather conditions, or even port closure. **In Chap. 8**, Qi studies the problem of how to dynamically revise the operation plan at the execution stage when a disruption occurs. Problem modelling and formulation are provided and a few key results of the solution scheme and managerial insights are derived. This is a rather new research area in ocean transport logistics, though it has been well studied in the airline industry. The chapter concludes by suggesting a few interesting topics for future research.

In Chap. 9, Plum, Pisinger and Jensen investigate an important optimal bunker purchasing problem in container shipping lines because the bunker cost constitutes a major component of the daily operating cost of a liner container ship. They first explain the bunkering issue in the liner container shipping industry. A base model for a single-containership bunker pursing is built taking into account the practical operational constraints. They further present a mixed-integer programming model with a novel solution approach for the bunker purchasing with contracts and discuss possible extensions of the model. Numerical experiments are carried out, and further research directions are highlighted.

In Chap. 10, Wang, Alharbi and Davy address the tactical-level interactions between container port operators and container shipping lines. They examine, in particular, a practical route schedule design for tactical liner ships that involves the interaction between container shipping lines and port operators on the availability of port time windows at each port of call. With some mild assumptions, they formulate the problem as a nonlinear non-convex optimization model and design an efficient dynamic-programming-based solution algorithm. A case study based on a trans-pacific ship route is conducted to assess the efficiency of the designed solution algorithm. Four specific future research directions are discussed.

In Chap. 11, Psaratis and Kontovas comprehensively examine the slow steaming strategies adopted by shipping lines. They present taxonomy of sailing speed models and analyze the main trade-offs. A decision model combining sailing speed and route choice is developed. Some examples are presented to introduce the main issues related to slow steaming. They point out that solutions for optimal environmental performance are not necessarily the same as those for optimal economic performance. A private operator of shipping lines would most certainly choose optimal economic performance as a criterion if policy-makers intend to influence the operator's decision to achieve a social optimum.

In Chap. 12, Wang and Liu contribute a comprehensive overview of existing studies on global container transport network design. After introducing the fundamentals and unique features of a container liner shipping network, a framework for container transport network design is proposed. They discuss the four special network design problems examined in the literature—ship route design with or without container transshipment operations, feeder shipping network design, hub-and-spoke shipping network design and general shipping network design. Five model formulations for the general shipping network design are presented. They end the chapter with suggestions for future research.

Part III: Shippers and Global Supply Chain Management

In Part III, we study the impact of ocean transport logistics on global supply chains. We shed light on this issue by investigating several related topics. These topics are purchasing transportation services from the shipper's viewpoint, ocean transport and the facilitation of trade, modelling global container transport demand, cooperation and competition in logistics operation, hinterland transportation as well as green corridors in the supply chain.

Shippers are the major service users in ocean transport logistics. Clearly, minimizing transportation costs is very important for global shippers who need to move their cargo containers all over the world. **In Chap. 13**, Xu and Lai introduce a general optimization model for the transportation service procurement problem (TSPP). After reviewing various existing solution methods for different variants and their extensions, they formulate a new general optimization model and discuss extensions to the existing results. Further research topics are also discussed, such as stochastic setting of the problem, trade-offs between transit time and freight cost, contract coordination and mechanism design.

In Chap. 14, Veenstra investigates the ocean transport part of international trade transaction. In particular, he highlights a number of processes related to ocean transport that generate uncertainty and costs in logistics chains. He concludes that certain ocean transport processes incur time loss and uncertainty and affect the efficiency of logistics and supply chains. He predicts that "if such frictions exist, there will be a tendency to move from a market relationship to a more hierarchical relationship between parties involved in the transaction." Examples from the Port of Rotterdam and its European hinterland are also provided.

In Chap. 15, Tavasszy, Ivanova and Halim examine the methods and techniques used in the analysis of the global container transport demand. Although the modeling of the global container transport demand can follow the generic architecture available for freight transport modeling, they find that few studies in the literature focus on global container transport modeling. They first model the global container demand between regions as the outcome of the process of production, consumption and trade. Based on the region-to-region demand, they proceed to model container demand for transport services by mode and route, including the container demand for maritime and inland port services.

In Chap. 16, Lee and Song examine the environmental challenges recently faced by maritime logistics operators, and investigate ways in which these operators can effectively manage competition and co-operation with their rivals to better respond to those challenges and thus achieve their strategic goals. They establish a theoretical framework to show the positive relationship between co-operative networks, knowledge acquisition and the value of maritime logistics. A comprehensive survey of existing literature reveals that a high level of co-operation in a co-operative network facilitates knowledge acquisition, and competition promotes the positive impact of co-operation on knowledge acquisition. The acquired knowledge helps to increase the value of maritime logistics. They conclude that this outcome will certainly provide maritime operators with strategic insights into the identification of determinants and/or sources of competitive advantage and greater organizational performance from inter-organizational coordination and knowledge-based perspectives.

Hinterland transportation is increasingly important for ocean transport logistics, especially in the European container port system. In **Chap. 17**, Bouchery, Fazi and Fransoo analyze the most important features of such container transportation systems for the hinterland supply chain. In addition to reviewing the current state of the art and identifying avenues for future research at the network design level, they also characterize those important factors influencing the trade-off between intermodal transportation and truck-only deliveries. A case study of coordination at an intermodal barge terminal in the Netherlands is also provided. A better information system has been identified as a crucial component of efficient hinterland intermodal transportation. This is an interesting area worth further investigation by the operations management community.

Finally, **in Chap. 18**, Panagakos, Psaraftis and Holte present the concept of green corridors and analyse their possible impact on the supply chain. A green corridor was introduced by the European Commission in 2007 aiming at reducing the environmental and climate impact of freight logistics. This chapter mainly focuses on surface fright transport, including maritime transport. It is well known that consolidation of large volumes of freight transport over long distances can reduce transport cost and emission and hence rail and waterborne transport have certain advantages if arranged effectively. They report the analysis performed under the SuperGreen project. They also provide an example that compares the deep sea service linking China to Europe (Shanghai – Le-Havre – Hamburg range) and the trans-Siberian rail link between Beijing and Duisburg/EU. They conclude that in terms of costs and CO_2 emissions (on a per tonne-km basis), deep sea shipping has significant advantages over rail transport although the latter is faster.

Acknowledgements

We would like to thank many colleagues for contributing chapters to this book. We are also grateful to the Research Grants Council of Hong Kong SAR, China, for their funding support (T32-620/11) Cheong Ying Chan Professorship, and NOL Fellowship Programme of Singapore.

Contents

Part I Container Terminal Operation: Innovations, Trends, Competition and Business Models

Part II Shipping Liners: Tactical and Operational Management

Part III Shippers and Global Supply Chain Management

Contributors

Abdurahim Alharbi School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW, Australia

Ronald Aprilyanto Halim TNO and Delft University of Technology, Delft, The Netherlands

Yann Bouchery School of Industrial Engineering, Technische Universiteit Eindhoven, Eindhoven, The Netherlands

Pei-Hsuan Chang Department of Logistics and Maritime Studies, C.Y. Tung International Centre for Maritime Studies, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

Ek Peng Chew Department of Industrial and Systems Engineering, National University of Singapore, Singapore, Singapore

Pam Davy School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW, Australia

Peter W. de Langen Port and Logistics Advisory (PLA), Eindhoven University of Technology, Eindhoven, The Netherlands

Jing-Xin Dong Business School, Newcastle University, Newcastle upon Tyne, UK

Stefano Fazi School of Industrial Engineering, Technische Universiteit Eindhoven, Eindhoven, The Netherlands

Jan C. Fransoo School of Industrial Engineering, Technische Universiteit Eindhoven, Eindhoven, The Netherlands

Even Ambros Holte Marintek, Trondheim, Norway

Olga Ivanova TNO and Delft University of Technology, Delft, The Netherlands

Peter N. Jensen Maersk Oil Trading, København K., Denmark

Xinjia Jiang College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing, China

Kap Hwan Kim Department of Industrial Engineering, Pusan National University, Gumjeong-gu, Busan, Republic of Korea

Christos A. Kontovas Department of Transport, Technical University of Denmark, Kgs. Lyngby, Denmark

Xiaofan Lai Department of Logistics and Maritime Studies, The Hong Kong Polytechnic University, Kowloon, Hong Kong

Jasmine Siu Lee Lam Division of Infrastructure Systems and Maritime Studies, School of Civil and Environmental Engineering, Nanyang Technological University, Singapore, Singapore

Eon-Seong Lee Department of Maritime Logistics and Management, Australian Maritime College, Univerisity of Tasmania, Launceston, Australia

Hoon Lee Institute of Logistics Systems, Total Soft Bank Ltd, Haeundae-gu, Busan, Republic of Korea

Loo Hay Lee Department of Industrial and Systems Engineering, National University of Singapore, Singapore, Singapore

Paul Tae-Woo Lee Department of Business Administration, Soochow University, Taipei, Taiwan

Zhiyuan Liu Institute of Transport Studies, Department of Civil Engineering, Monash University, Clayton, VIC, Australia

Chin-Shan Lu Department of Logistics and Maritime Studies, C.Y. Tung International Centre for Maritime Studies, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

Ho-Yin Mak Department of Industrial Engineering and Logistics Management, The Hong Kong University of Science and Technology, Hong Kong, China

Theo E. Notteboom Transportation Management College - Dalian Maritime University, China and ITMMA - University of Antwerp, Belgium and Antwerp Maritime Academy, Belgium

George Panagakos Department of Transport, Technical University of Denmark, Kgs. Lyngby, Denmark

David Pisinger DTU Management Engineering, Technical University of Denmark, Lyngby, Denmark

Christian E. M. Plum DTU Management Engineering, Technical University of Denmark, Lyngby, Denmark

Maersk Line, København K., Denmark

Harilaos N. Psaraftis Department of Transport, Technical University of Denmark, Kgs. Lyngby, Denmark

Xiangtong Qi Department of Industrial Engineering and Logistics Management, The Hong Kong University of Science and Technology, Clearwater Bay, Kowloon, Hong Kong

Dong-Ping Song School of Management, University of Liverpool, Liverpool, UK

Dong-Wook Song Transport Research Institute, Edinburgh Napier University, Edinburgh, UK

Lóránt A. Tavasszy TNO and Delft University of Technology, Delft, The Netherlands

Ho-Tak Tsang Department of Industrial Engineering and Logistics Management, The Hong Kong University of Science and Technology, Hong Kong, China

Albert Veenstra School of Industrial Engineering and Innovation Sciences, Department of Operations, Planning, Accounting and Control, Eindhoven University of Technology, Eindhoven, The Netherlands

Shuaian Wang School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW, Australia

Strome College of Business, Old Dominion University, Norfolk, Virginia, USA

Zhou Xu Department of Logistics and Maritime Studies, The Hong Kong Polytechnic University, Kowloon, Hong Kong

Part I
Container Terminal Operation: Innovations, Trends, Competition and Business Models

Chapter 1
Innovative Container Terminals to Improve Global Container Transport Chains

Xinjia Jiang, Ek Peng Chew and Loo Hay Lee

Abstract Due to globalization, the container traffic has been growing rapidly and it was observed that the transshipment activities of many major ports are increasing at a faster speed. One of the main reasons is due to the increasing vessel size and the preference of ocean liners in adopting the “hub-and-spoke” system. Connectivity plays an important role for these major ports to remain competitive in capturing transshipment markets. In this chapter, we propose novel performance indicators which can be used to measure the impact of connectivity from the network perspective. These indicators can be used for port benchmarking and also provide useful insights for port operators, such as who are their competitors, and how fragile they are in the competition. We also introduce yard management strategies which aim at improving land productivity while retaining the operational efficiency in transshipment ports. As there is a strong drive for sustainability to address issues such as land scarcity, labor shortage and using of green energy, there is a need for new and innovative concept for designing the port of the future. We have proposed an innovative double storey automated container port—Sustainable Integrated Next Generation Advanced Port (SINGAPort) in this regard.

1.1 Container Terminals in the Global Container Transport Chains

Since the introduction of standardized steel containers, this concept has achieved a worldwide acceptance with the development of supporting transport facilities and handling equipment. As the international trade grows, the volume of world container transportation has boomed during the last two decades. According to Drewry Shipping Consultants, the annual container throughput has increased around 6.7 times

---

L. H. Lee (✉) · E. P. Chew
Department of Industrial and Systems Engineering, National University
of Singapore, 10 Kent Ridge Crescent, Singapore 119260, Singapore
e-mail: e-mail: iseleelh@nus.edu.sg

X. Jiang
College of Economics and Management, Nanjing University
of Aeronautics and Astronautics, 29 Yudao Street, Nanjing 210016, China

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_1

from 88.150 million TEUs (20-foot equivalent unit) in 1990 to 588.905 million TEUs in 2011. As the crucial connection points for various container transportation modes, container terminals have played an important role to improve global container transport chains.

Traditionally, a container terminal mainly focuses on "gateway activities", which handle containers between inland transportation modes and container vessels. For gateway activities, standardized containers packed with valuable export goods and resources are assembled from inland manufacturers via trains, trucks and small barges. When a container vessel arrives at the port, it will not only pick up these export cargos, but also discharge a large batch of import containers packed with goods from other parts of the world. With the gateway activities, container terminals become the life line of international trade and economic prosperity of coastal cities. Many new terminals are built to attract or facilitate the international trade as well as stimulating economic growth of coastal cities. Typical examples can be found in China, such as Shanghai, Ningbo, Dalian, etc. These container terminals not only increase portals of global container transport chains, but also improve the container handling efficiency with cutting-edge technologies.

Compared with the gateway activities, the growth of annual transshipment activity is even faster, at 11.7 times, from 15.504 million TEUs in 1990 to 181.596 million TEUs in 2011. The portion of transshipment among the total container traffic has increased from 17.6 % in 1990 to 30.8 % in 2011. The transshipment activities focus on vessel-to-vessel container handling service, instead of connecting the inland transportation modes and container vessels. One of the main reasons, which lead to this growing trend of transshipment activities, is the increasing popularity of "hub-and-spoke" pattern for container transport chains. This "hub-and-spoke" pattern allows the container traffic in the major routes to be combined and transported via mega vessels, which will not call at all the final destinations of containers, but only at some hub terminals. The container transportation between hubs and spoke terminals will then be carried out by feeder vessels. This hub-and-spoke pattern not only helps to reduce transportation cost, but also provides more flexibility in ship routing. With the increasing vessel size and shipping alliance, it is expected to see more hub-and-spoke patterns in the global container transport chains.

Due to the increasing importance of transshipment activities, the future improvement of global container transport chains may depend on the major container terminals which have large portion of transshipment activities, such as Singapore, Hong Kong, Tanjung Pelepas, Salalah, Port Said, Gioia Tauro, Malta Freeport, and Algeciras, as well as the new rising deep-water ports. To improve the performance of container terminals, various KPIs have been used to measure the efficiency of port operations at the quayside, the landside and the storage yard. However, the existing KPIs are mainly focusing on the operation efficiency of each individual port, while the efficiency of transshipment activities depends on the connectivity of the port network. Then, how to measure the port connectivity? On the other hand, the operation efficiency of a transshipment port have great impact on the transport chains, but what are the management issues challenging the transshipment ports? What are the innovative management methodologies and terminal designs that can be applied to improve the performance of a transshipment port? To answer these questions, we need to first look at the future trends and challenges related to transshipment activities, which will be discussed in details in the following section.

1.1.1 Trends of Container Transportation

1. Increasing vessel size

To achieve the economy of scale, the size of container ships has been increasing steadily over the past decades, as shown in Fig. 1.1. In 1970s, the largest container ship had a capacity of 2500 TEUs. That increased to 5000 TEUs in 1990s and 6800 TEUs in early 2000. In 2006, Emma Maersk was launched to be the first container ship in the E-class of eight owned by the A. P. Moller-Maersk Group. At that time, Emma Maersk was known as the largest container ship ever built, which can carry around 14,770 TEUs. In June 2013, the Triple E ships were launched by Maersk, which can carry around 18,000 TEUs.

On the other hand, the average vessel size has also been increasing steadily. According to World Shipping Council, the existing and future number of container vessels in each size range can be shown in Table 1.1. By 1 July 2013, the average size of existing container vessels is 3401 TEUs. However, the average size of vessel on order is 7368 TEUs, which is more than two times the current average size. Around 50 % of container vessels on order are above 7500 TEU, while the percentage of such vessels among the existing vessels is only 11 %.

High-volume container ships can provide significant cost savings to liner companies because of the economy of scale in major trade lanes. A 12,000 TEU ship can be operated with the same 13 or 14 crewmembers required by a ship with half the capacity. Per-unit costs of capital investment and fuel consumption are substantially less than those for two vessels of half the size. Thus, the large container ships dominate the order books. By the time the on-order vessels are put into use, around 40 % of the container transport capacity will be performed by only 14 % of all container vessels. It is expected to see more and more containers transported via larger vessels, while the container traffic in major routes might be combined to fulfill the mega vessels. The increasing vessel size also urges the development in port facilities, since it requires ports to be equipped with modern handling equipment, especially specialized gantry cranes. The trend towards larger vessels is likely to encourage ports to invest further in port handling equipment.

2. Increasing transshipment activities

The efficiency of port operations serves as an indicator of a country's economic development, and at a global level more than 85 % of international trade is transported through seaports (Liu 2008). Since the application of standardized transportation facilities and handling equipment, container transportation has achieved a worldwide acceptance. This together with the development of maritime transportation leads to a

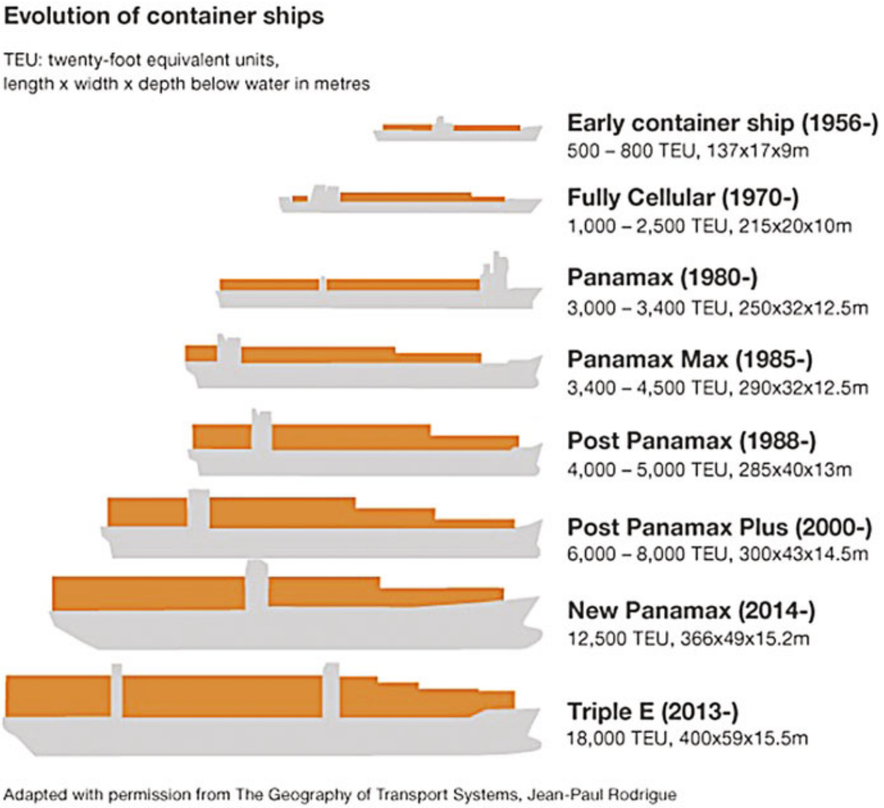


Fig. 1.1 Evolution of container ships. (Source: The geography of transport systems, Jean-Paul Rodrigue)

boom of container transportation market. As shown in Fig. 1.2, the total throughput of container terminals has been expanding significantly during the last few decades.

However, the economic crisis of 2007–2010, also known as the Great Recession, has greatly impacted the global economy. In 2009, the world GDP declined by 1 %, for the first time in the post-World War II era. The recession has also hit global trade, with trade volumes declining by as much as 25 % in 2009 from 2008's level, which in turn badly hit the container shipping industry. However, the setback has been quickly recovered. The annual throughput level has surpassed the level in 2008, while the throughput level in 2011 is even higher. The trend is expected to continue with the recovery of global economy.

An interesting phenomenon can be observed that the transshipment activates are increasing at a faster speed. The portion of transshipment activities in the total container traffic has increased from 18 % in 1990 to around 31 % in 2011. This can be attributed to the increasing popularity of the "hub-and-spoke" pattern. This is because of the fact that only several major trade lanes have sufficient demand to

Table 1.1 Container vessel fleet. (Source: World Shipping Council)

	1 July 2013 Existing		1 July 2013 On-order	
Size range (TEU)	Ships	TEU	Ships	TEU
10,000–18,500	183	2,357,532	103	1,451,508
7500–9999	355	3,082,735	119	1,075,437
5100–7499	482	2,964,138	27	173,629
4000–5099	758	3,428,711	49	231,423
3000–3999	275	946,502	47	173,396
2000–2999	671	1,708,860	31	75,848
1500–1999	565	962,650	38	66,061
1000–1499	683	799,639	24	25,664
500–999	774	576,324	7	5580
100–499	222	71,134	0	0
Total	4968	16,898,225	445	3,278,546
Average		3401		7368

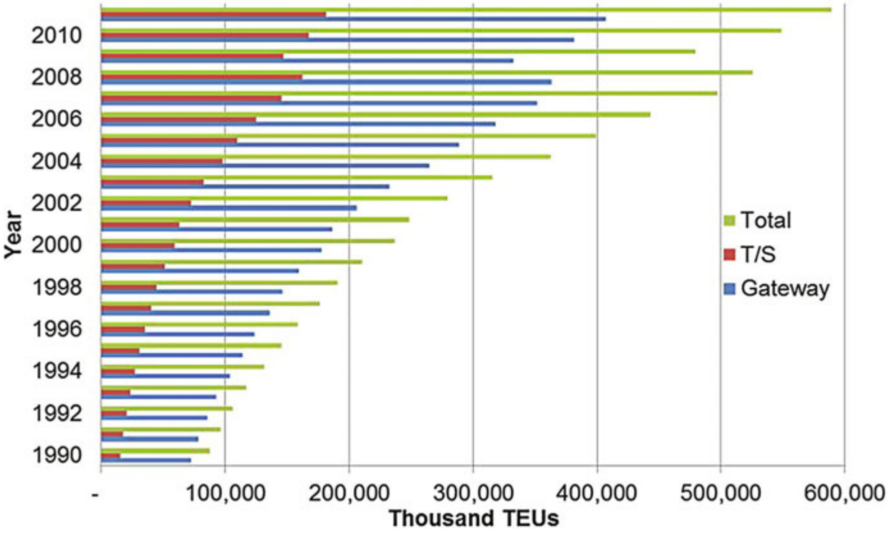


Fig. 1.2 Growth of annual world throughput level. (Source: Drewry Shipping Consultants)

support these mega vessels. Furthermore, the mega vessels can only dock in a small number of ports because of their deep drafts. It is thus natural for carriers to adopt a hub-and-spoke network structure in order to take advantage of the economy of scale by consolidating the demand from smaller trade lanes and ports. Consequently,

Table 1.2 World container terminal ownership ranking. (Source: Drewry Maritime Research)

2010				2011			
Rank	Operator	Million TEU	% Share	Rank	Operator	Million TEU	% Share
1	PSA	51.3	9.4	1	PSA	47.6	8.1
2	HPH	36.0	6.6	2	HPH	43.4	7.4
3	DPW	32.6	6.0	3	DPW	33.1	5.6
4	APMT	31.6	5.8	4	APMT	32.0	5.4
5	SIPG	19.5	3.6	5	Cosco	15.4	2.6
6	CMHI	17.3	3.2	6	TIL	12.1	2.1
7	Cosco	13.6	2.5	7	CSTD	7.8	1.3
8	MSC	9.9	1.8	8	Evergreen	6.9	1.2
9	SSA Marine	8.6	1.6	9	Eurogate	6.6	1.1
10	Modern Terminals	8.3	1.5	10	HHLA	6.4	1.1

the demand for transshipment operation has also increased due to the shift towards larger vessels and the shipping alliance. With more mega-vessels putting into use, this trend is expected to continue. Besides, the "hub-and-spoke" pattern also increases the flexibility of the container transport chains. It allows the transport network reach more ports at lower cost with the proper combination transshipment hub and spoke terminals.

3. Intense port competition

The recent market share of global terminal operators can be revealed by a report from Drewry Shipping Consultants. The world container terminal ownership ranking in 2010 and 2011 can be shown in Table 1.2. The big four terminal operators (PSA, HPH, DPW, APMT) are the dominant players in the global container transport chains, in both 2010 and 2011. There are quite a few challengers, like Cosco, Evergreen, and CSTD (China Shipping Terminal Development). According to the report, these top four market leaders enjoyed around 27.8 % share of world throughput in 2010 and around 25.5 % share in 2011. The influence of these four companies within the container terminal market may continue to strengthen but the effect of two new arising challengers from China cannot be ignored.

On the other hand, the competition among major container ports is becoming more intense, especially in the Asia region. Among the container terminals with highest throughput around the world, most of them are located in the Asia region, as shown in Fig. 1.3. In 2010, the port of Shanghai has surpassed Singapore to be the world's busiest container ports. Moreover, current dominant ports are being challenged by

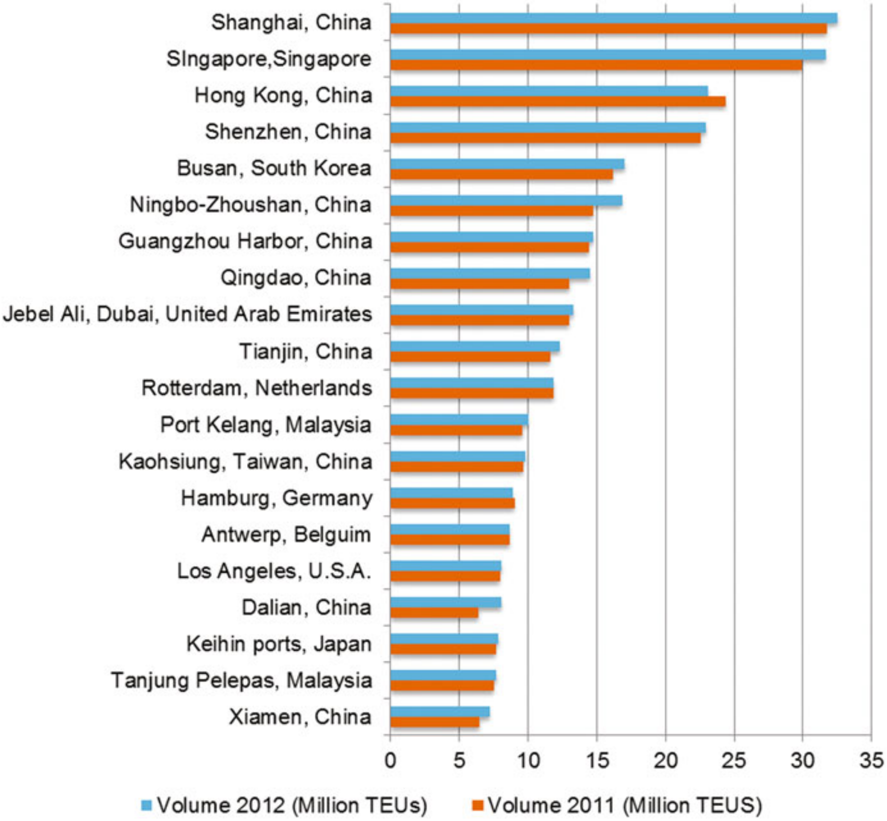


Fig. 1.3 Annual throughputs of top terminals around the world. (Source: The Journal of Commerce, August 20, 201and August 19, 2013 and ports)

the emergence of other fast-developing ports that are located within proximity of the region, such as Shenzhen and Ningbo.

This competition is most crucial for the stevedore-based terminal operators, especially PSA and HPH. Among the top four operators, APM Terminals is a hybrid terminal operator which focusing on both shipping and terminal operation. The main activity of such company is container shipping, but separate terminal-operating divisions have been established. These terminal-operating divisions are expected to handle a significant amount of third-party traffic besides serving the core liner shipping business of the parent companies. They are often designed to operate as independent profit centers. Opposite to them, HPH, PSA and DP World are the leading stevedore-based operators whose core business is terminal operation. They view terminals as profit centers. However, DP World is one of the top two global container terminal operators who have the most geographically balanced spread of container terminal activities. On the other hand, PSA and HPH remain heavily dependent on traffic generated by terminals in the Far East and South East Asia. These two regions account for more than 81 % of the PSA's equity TEU throughput, and 59 % of HPH. To better position itself in the global container transport chain, PSA purchased substantial shareholdings in terminals in Kandla and Kolkata in India. It also entered into a joint venture with Pacific International Lines (PIL), one of the largest shipowners in Southeast Asia, to develop and operate a terminal in Singapore. Although the landfill projects are constant in Singapore, the scarcity of land space is still a major challenge to PSA.

1.1.2 Challenges for Transshipment Ports

With all the trends discussed in the previous, challenges are now faced by container terminals worldwide, especially for transshipment ports. With the increasing vessel size and shipping alliance, the "hub-and-spoke" pattern is becoming more and more popular. Combined with the growth of container transportation, the increasing importance of transshipment activities is inevitable. The annual transshipment throughput has grown 11.7 times, from 15.504 million TEUs in 1990 to 181.596 million TEUs in 2011. The transshipment throughput as a proportion of total volume handled has grown from 18 % in 1990 to 31 % in 2011. As a result, any thorough study of port competitiveness should include the transshipment capability of a container terminal. On the other hand, the competition among port operators is becoming more intense, especially in the Asia region. Among the leading ports in Asia, the dominance of many ports depends on transshipment activities, such as Singapore and Hong Kong. To provide some helpful insights for addressing these challenges, we focus the discussion on transshipment ports.

Our first topic for discussion is the challenge of port competition in the Asia region. There are many factors known to affect the decision of port selection for shipping liners, while port connectivity is one of the key considerations for transshipment hubs. In order to help major ports understand their particular strengths and weaknesses to recognize the potential threats and opportunities, the following questions become crucial.

- What makes a port transshipment hub? How to measure the connectivity of each terminal?

After addressing this question, we will focus on the means to improve the performance of transshipment ports among the global container transport chains. The innovative container terminals are discussed from two perspectives, namely "innovative management methods" and "innovative terminal designs".

- Innovative management methods: How to manage a transshipment port? To be more specific, what are the innovative management ideas which have been used or proposed for transshipment ports?
- Innovative terminal designs: What are the state-of-the-art systems which are potential to improve the transshipment ports? How will the trends of container transportation affect the design of container terminals?

1.2 What Makes A Port Transshipment Hub?

As a result of globalization, mergers and acquisitions among shipping lines, the container traffic is being combined in volume and controlled by a single line or a shipping alliance. This implies that the capacity of shipping line to influence the business of a port is much greater than it has been in the past. This warrants a need for the port operators to understand the underlying factors of port competitiveness from the shipping lines' perspective.

In 2000, Maersk Sealand relocated its major transshipment operations from the Port of Singapore (PSA) to the Port of Tanjung Pelepas (PTP) in Malaysia. The impact of this relocation on the regional transshipment market structure was significant. Maersk Sealand was then the largest shipping operator in Singapore. Its shift to PTP resulted in a decline of approximately 11 % in PSA's overall business. In 2001, PSA's total container throughput fell from 17.9 million TEUs to 15.52 million TEUs (Tongzon 2006), marking a year-on-year drop of 8.9 %. In the same period, PTP's container throughput had increased nearly 5 folds, from 0.42 to 2.05 million TEUs. The shipping industry in Singapore region grew concerned about Maersk Sealand's relocation and the potential ripple effect on other shipping lines' decisions as well as the related business activities (Allison 2000; Kleywegt et al. 2002). As shipping lines form strategic alliances to achieve economies of scale, the interdependency among alliance members and small- and medium-size shipping lines heightens. Consequently, Maersk Sealand's decision on changing its transshipment port-of-call could well induce similar decisions among affiliating carriers. In 2002, Evergreen and its subsidiary Uniglory also shifted most of their container operations, amounting to 1–1.2 million TEUs of annual throughput, from PSA to PTP. Since then, other shipping lines have also started to provide direct services to PTP. For example, APL had chosen PTP for its West Asia Express service between Asia and the Middle East (Kleywegt et al. 2002).

However, the port of Singapore has never lost its importance in transshipment activities in the Asia region. In 2006, "Emma Maersk", the largest container vessel at that time, called at the port of Singapore on her historic maiden voyage. During the same year, the port of Singapore surpassed Hong Kong to be the biggest container port, with the annual throughput of 24.8 million TEUs. In 2011, Singapore was still the world's busiest transshipment port with the annual throughput of 29.9 million TEUs, which is around 4 times the PTP. In 2013, the third of Maersk Line's latest Triple-E class of vessels has been named "Mary Maersk" at a ceremony at PSA's Pasir Panjang terminal in Singapore.

Nowadays, the major relocation of transshipment operations from PSA to PTP is still a hot research topic in the maritime area, while the nice recovery of PSA as a major transshipment hub also catches our attention. There are many interesting questions to be answered. Firstly, what caused this relocation of transshipment activities? Secondly, what helped Singapore successfully recover? From a more general perspective, what makes a port transshipment hub?

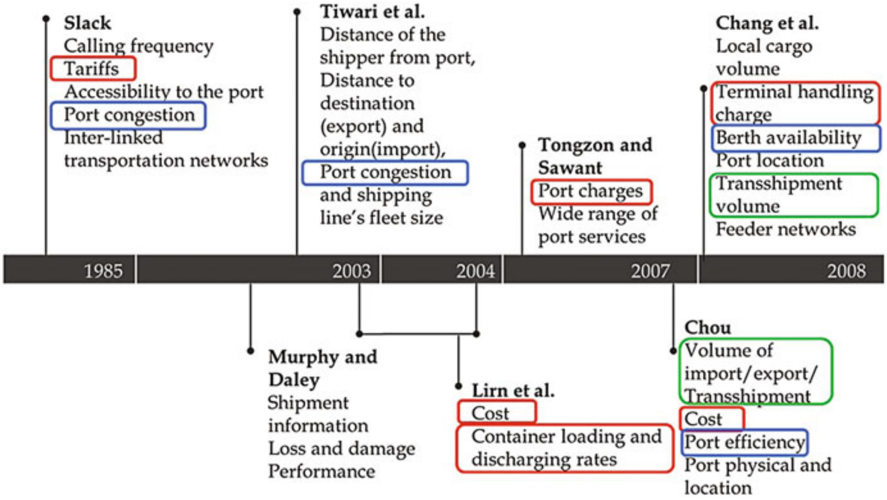


Fig. 1.4 Port selection criteria found in port selection literatures

1.2.1 *Factors of Port Selection*

Many port selection studies aim to find the key factors that affect the decision for selecting a port. To identify the key factors, the first thing we shall clarify is who make the decision of port selection, shippers or shipping lines? Although Slack (1985) conducted his study from the shipper's perspective, he mentioned in his analysis of survey results that shipping lines are the key actors in the port selection process. Also, D'Este et al. (1992a, b) suggested that as shipping lines increased their scale of operations and shippers began soliciting prices for door-to-door service rather than individual segments, the port selection shifted from the shipper to the shipping lines. The port selection criteria in the existing port selection literatures can now be summarized in Fig. 1.4.

In particular, with increasing importance of the port function as a transshipment facility, recent port selection studies have paid attention to transshipment port selection problem. Lirn et al. (2003, 2004) found in their studies that the transshipment port selection depends mainly on port competitiveness and efficiency, as represented respectively by the cost and the container loading and discharging rates. Similarly, Chou (2007) suggested port manager that if they want to become a transshipment hub port, it will be the most efficient way to attract ocean carriers by increasing the volume of import/export/transshipment containers and decreasing port charge. Chang et al. (2008) considered port selection problem from the trunk liners' and feeder service providers' perspective. Authors concluded that local cargo volume, terminal handling charge, berth availability, port location, transshipment volume and feeder network are the most important factors.

Among all these identified factors, the major relocations of transshipment activities introduced at the beginning of this section can be attributed to three major factors, namely “cost”, “efficiency” and “connectivity”.

The “cost” factor is the major fuse to the relocation of transshipment activities from PSA to PTP. As global customers exert increasing pressure on shipping lines to lower their prices, the competition to reduce costs among shipping lines inevitably intensifies. Shipping lines are forced to explore options which give the most cost-saving. With these drivers in mind, the attractiveness of PTP’s port price, which was some 30 % lower than that of PSA’s at that time, becomes apparent. In fact, Evergreen had estimated that their shift to PTP would save them between US$ 5.7 million and US$ 30 million per annum (Kleywegt et al. 2002).

However, it is the “efficiency” and “connectivity” that helps the port of Singapore retained its position as a transshipment hub. Although the port of Singapore costs higher, it is well known for its operation efficiency and cutting edge technologies applied in port operations. It is equipped with the deep draft and up-to-date quay cranes, which enable the port to serve the largest container vessels, such as “Emma Maersk” and “Mary Maersk”. The operation efficiency is the basis to become a transshipment hub or even for the success of a general container terminal. On the other hand, “connectivity” is the most important factor that enhanced the position as a transshipment hub. As Singapore is subjected to stringent growth limitations as a gateway port but possess excellent locations along the Strait of Malacca, the transshipment role of Singapore can date back to centuries ago. The transshipment activities offer Singapore a good opportunity to expand beyond the demands of their respective catchment economies and more importantly, tap into the international cargo flows to enjoy superior profits. Such a long history of transshipment operations have equipped Singapore with advanced feeder line networks which serve to transport containers to/from tributary ports. These networks give Singapore the good port connectivity. As PTP only started operation in 1999, it lacks the operation efficiency and, most importantly, the advanced port connectivity.

The “cost” and “efficiency” will be discussed in the Sect. 3, “How to manage a transshipment port”. In this section, the discussion is focused on the “port connectivity”. Firstly, the port connectivity analysis methods will be introduced in Sect. 2.2. To focus on the application of the method, we will not go into the detailed models, but demonstrate the basic idea and the index for measurement. With the basic knowledge of the method, a case study of port connectivity analysis will be presented in Sect. 2.3. The role of a transshipment port in a transportation chain will be revealed through the connectivity analysis.

1.2.2 Connectivity Analysis for Transshipment Ports

Transshipment operations have rapidly grown in importance in the past few decades. According to a report from Drewry Shipping Consultants, transshipment volume as a proportion of total volume handled has grown from 18 % in 1990 to 31 % in 2011. As such, it is no surprise that any thorough study of port competitiveness should include their transshipment capability. A major factor influencing transshipment capability is the concept of port connectivity. Intuitively, port connectivity indicates how well one port connects to others in a maritime transportation network and its ease of accessibility by liner services. A port with a high level of connectivity is likely to have a great advantage for transshipment operations, so much so that connectivity could be used as a direct proxy for a port's competitiveness in terms of transshipment. In general, the higher connectivity level a port has, the more attractive it will be to liner companies in the sense of facilitating the transportation of cargo and reducing transportation cost and time, leading to the port being more competitive than its peers.

However, it is not immediately evident how to determine a port's connectivity as it is a concept that can have different interpretations and meanings in different cases. To date, this concept of port connectivity has not been well defined despite several papers on the topic. In this section, we propose a connectivity analysis framework from a network perspective that can be used to benchmark the competitiveness of various ports in the network in terms of the impact of their transshipment operations.

At first glance, a possible way to measure the connectivity of a port is to simply count the number of direct connections it has to other ports in the network, i.e., the number of incoming and outgoing shipping services to/from the port in question. Despite providing a reasonable estimate, this simplistic method can leave out some important details. Three such important factors are discussed as follows for an origin port A and destination port B:

Responsiveness: The average waiting time that a supplier has to wait for a service to ship his goods from port A to port B. This factor is related to the frequency of shipping between port A and port B. The lower the frequency, the longer the expected waiting time for the supplier. Vice versa, the higher the frequency, the more trips from A to B is made per time period and thus the shorter the expected waiting time.

Capacity: Even if there are frequent services between port A and port B, it could be that the ships serving port A and port B have low capacities, which limit the amount of cargo that the supplier can ship at any time. On the other hand, if services between port A and port B have low frequencies but high capacities, a large amount of cargo can be shipped in a single shipment even though the average waiting time might be long. Therefore, capacity should also be considered as a factor in port connectivity.

Network structure: Besides direct links, there may also be services connecting ports A and B that comprise of multiple stops. In this case, the structure of the transportation network, such as the existence of bottlenecks or the ease of accessibility to large hubs, can also play a very large part in the viability of transshipment services and hence affect a port's connectivity.

In order to account for these factors, a new definition of connectivity is proposed as follows. Considering that transshipment has become a major port service, the connectivity of a port can be defined as the impact on the transportation network's performance when transshipment services are not available at this port. Intuitively, such impact is proportional to the connectivity of the corresponding port. This methodology also accounts for any network effects that may influence services between ports that are not directly connected. If a port has a high degree of connectivity, then many carriers will come to this port to transship their cargoes. If said port's transshipment capabilities are now disabled, then many of these carriers will have no choice but to select other ports at which they can transship their cargoes, which will likely result in greater transportation cost or longer shipping time. Thus, ports with a high degree of connectivity will result in a greater impact on the transportation network's performance than those with low connectivity when transshipment services are not available.

Such an impact on the network can take many forms. For example, the model can measure the impact on the transportation flow capacity of the system, the impact on the transportation time of cargoes and so on. It is necessary to examine the results from different perspectives in order to obtain a thorough and comprehensive understanding of this concept and provide meaningful benchmarking results.

1.2.3 Implications of Connectivity Analysis—A Case Study

This subsection focuses on the effects of transshipment on transportation capacity and shipping time of major ports in the Asia Pacific region. By comparing various scenarios in which transshipment is disabled for certain ports against a base case, we can rank the ports analyzed according to the network benefits provided by transshipment services. The network was constructed using data from the top 10 largest liner companies in 2008 according to CI-Online.

When considering the capacity model, the disabling of transshipment services at Singapore, Shanghai, Dalian and Qingdao has the greatest impact in descending order (shown in Table 1.3). Singapore is the largest by a wide margin as it serves as the primary transshipment hub for all traffic to Oceania and between Asia and Africa/Europe. Shanghai is close behind due to its share of transshipment traffic to the USA, while Dalian and Qingdao serve as gateway ports to the North China region. The disabling of transshipment services at Qingdao and Dalian has a significant impact on the flow of traffic to Tianjin and the Beijing area. In comparison, Hong Kong and Shenzhen have a relatively small impact on the network, especially considering that they have a very large number of linking services.

The waiting time model provides some different insights as seen in Table 1.4. Disabling transshipment services at Singapore still has the greatest impact by a wide margin due to the very large number of transshipment services available. Singapore also serves as a consolidation hub for South East Asia ports to major markets such as the USA and Europe, thus there is a large increase in waiting time when these smaller services are not able to transship onto larger international liner services. Busan, Qingdao and Hong Kong are ranked second, third and fourth after Singapore in terms of impact on waiting time when transshipment is disabled. This is likely

Table 1.3 Connectivity results of major ports considering transportation capacity

Port	Weekly per OD*	TEU Change in capacity	% change	Rank
Base case	2,35,346.5			
Singapore	2,30,133.2	– 5213.32	– 2.22	1
Shanghai	2,31,861.3	– 3485.21	– 1.48	2
Dalian	2,32,224.6	– 3121.98	– 1.33	3
Qingdao	2,32,406	– 2940.54	– 1.25	4
Yokohama	2,33,248.1	– 2098.47	– 0.89	5
Ningbo	2,33,434.4	– 1912.17	– 0.81	6
Busan	2,33,475	– 1871.53	– 0.80	7
Kaohsiung	2,33,947.3	– 1399.24	– 0.59	8
Port Klang	2,34,010	– 1336.52	– 0.57	9
Hong Kong	2,34,389.6	– 956.95	– 0.41	10
Tianjin	2,34,395	– 951.49	– 0.40	11
Shenzhen	2,34,419.7	– 926.86	– 0.39	12
Tanjung Pelepas	2,34,487.7	– 858.88	– 0.36	13
Guangzhou	2,34,594	– 752.51	– 0.32	14
Yingkou	2,34,983.7	– 362.84	– 0.15	15
Laem Chabang	2,35,112.6	– 233.98	– 0.10	16

*OD is defined as Origin and Destination ports

due to their geographical location as gateway ports to the North China region, forming a bottleneck through which all services must pass. Tianjin has no impact on waiting time due to its position deep in the bay of Bohai. This means that services will not be routed via Tianjin as there will be some backtracking and hence time wasted.

The results provide an indication of the relative impact of these ports on the transportation network as a whole, which can be seen as a reflection of their connectivity. Further analysis can be performed using the same framework to obtain different insights, which can then be combined as part of the benchmarking process. For example, the results from the capacity model and the waiting time model can be integrated using the Pareto graph in Fig. 1.5 below.

The grouping of ports shows that Singapore clearly has the highest connectivity in terms of both impact on capacity and impact on waiting time. Shanghai, Qingdao and Busan form a second group of ports that have similar connectivity rankings, with some tradeoffs between capacity and waiting time within the group. Dalian has a large impact on capacity, but a small impact on waiting time, which makes its overall connectivity relative lower than the second group.

Table 1.4 Connectivity results of major ports considering waiting time

Port	Weekly per OD*	Change in waiting time	% change	Rank
Base case	7.2803			
Singapore	7.8865	0.6062	8.33	1
Busan	7.6767	0.3964	5.44	2
Qingdao	7.4596	0.1793	2.46	3
Hong Kong	7.4148	0.1345	1.85	4
Port Klang	7.3681	0.0878	1.21	5
Shenzhen	7.3548	0.0745	1.02	6
Laem Chabang	7.3444	0.0641	0.88	7
Shanghai	7.3412	0.0609	0.84	8
Kaohsiung	7.3357	0.0554	0.76	9
Dalian	7.3146	0.0343	0.47	10
Ningbo	7.3126	0.0323	0.44	11
Tanjung Pelepas	7.3038	0.0235	0.32	12
Guangzhou	7.2852	0.0049	0.07	13
Yokohama	7.2816	0.0013	0.02	14
Yingkou	7.2809	0.0006	0.01	15
Tianjin	7.2803	0	0	16

*OD is defined as Origin and Destination ports

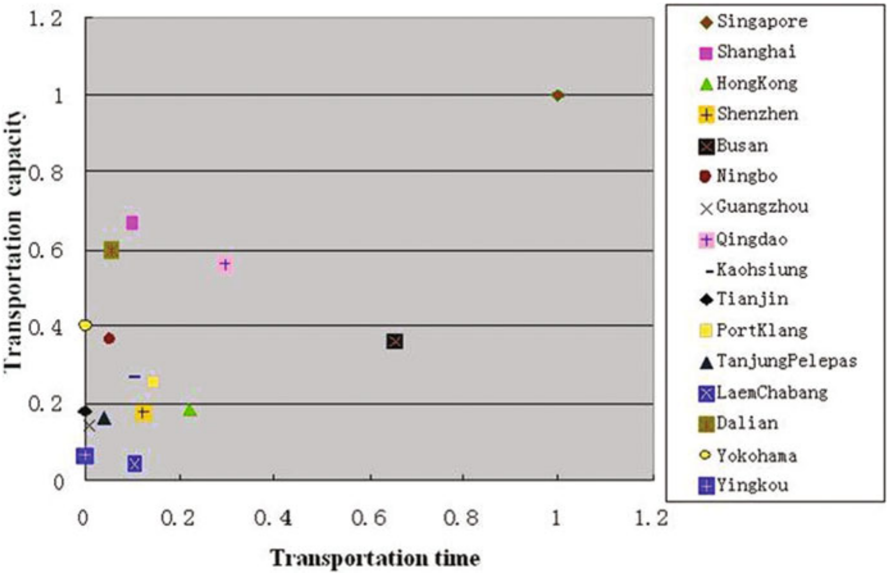


Fig. 1.5 Pareto analysis of capacity vs. waiting time

1.3 How to Manage A Transshipment Port?

1.3.1 Major Challenges in Transshipment Ports

1. BOA pressure to yard operation

The performance of quayside operations can be measured in terms of the "berth on arrival (BOA)" for ships. The BOA is typically defined as the probability that ships can successfully berth at the quay within 2 h of arrival at the port. To achieve a high BOA level, it is crucial to keep the vessel turnaround time at the minimum level. Both practitioners and researchers have focused much attention on quay-side operations to improve the loading and discharging of containers with new technologies, such as the dual-cycle quay cranes and tandem lift. However, the storage yard is now becoming the new bottleneck in terminal operation, especially for transshipment ports. The overall terminal performance will not benefit much from faster quay-side operations without the effective storage and retrieval of containers in the storage yard.

2. The congested yard operations

Transshipment hub ports usually handle a high volume of containers and most containers unloaded from one vessel will be eventually loaded onto other vessels in the port. To avoid too much double handling, containers are usually stored in the same location until being loaded. However, if a container at the bottom need to be retrieved first, the containers on top need to be repositioned. The extra moves to retrieve a container are called "reshuffles". Reducing reshuffles is one of the key considerations in yard management, as reshuffles increase the retrieval time. On the other hand, the loading and unloading activities are very heavy and are performed at the same time. This leads to many traffic movements potentially crossing one another due to space allocation in the same yard. Reducing congestion is another key consideration in storage yard management in transshipment ports. Thus systematic planning is important for transshipment ports to reduce reshuffles and congestions.

3. The scarcity of land

With the growing container traffic, more and more containers will be handled and temporarily stored in the ports. Simple physical expansion of a port is often constrained by the scarcity of land, especially for ports located in or near urban areas, such as Singapore and Shanghai. Conventional capacity expansions are often limited by competing land usage, availability of initial investment, and environmental concerns (Le-Griffin and Murphy 2006). Port operators have to look for innovative measures to increase container terminal capacity and productivity in order to meet the ever growing demand despite limited investment and terminal space.

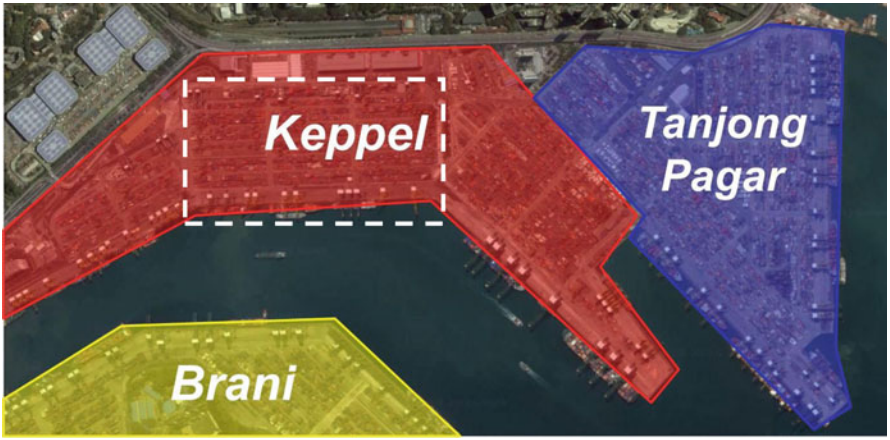


Fig. 1.6 Container terminals in Singapore. (Source: https://maps.google.com)

1.3.2 Existing Management Concepts for Transshipment Ports

To achieve the operation efficiency, many unique yard management concepts can be applied in transshipment ports. The storage yard management in transshipment ports can generally be divided into three phases, namely "yard sectioning", "template planning" and "space allocation". The first phase divides the whole storage yard into sections, each corresponding to several berths. A yard section usually serves a group of vessels. Within each section, the storage space is managed as small storage locations, each reserved for a destination vessel. The reservation of storage locations among the destination vessel is called the "yard template", which is planned in the second phase. In the third phase, the storage space is allocated to the incoming containers according to the given yard template. The three phases are discussed in more details as follows.

1. Yard sectioning

Take Singapore Port as an example, PSA operates five container terminals at Tanjong Pagar, Keppel, Brani and Pasir Panjang, with a total of 52 container berths. They operate as one seamless and integrated facility. Pasir Panjang Terminal is PSA's most advanced terminal. It is equipped with berths up to 16 m deep and with quay cranes able to reach across 22 rows of containers to accommodate the world's largest container ships. The Tanjong Pagar, Keppel, Brani are more close to each other, which can be shown in Fig. 1.6.

To provide more flexibility during operation, a terminal can be divided into sections. Vessels are assigned to sections, each corresponding to several berths, rather than the exact berth locations. The important planning issues at this phase include how to allocate arriving container vessels to each section considering berth allocation. For example, the port operators will try to group or allocate vessels to yard section such that movement of containers between yard sections can be minimized. A typical section in Keppel terminal can be shown in the dashed square in Fig. 1.6. Using this section as an example, we will demonstrate the considerations and innovative management concepts in phase 2 and 3.

2. Template planning

In transshipment ports, the "consignment" concept is generally applied to store containers to same destination vessel together. This is to facilitate faster loading process as it reduces reshuffles as well as long distance movements of yard cranes (Han et al. 2008). Different storage strategies have been proposed to achieve consignment. Commonly, the whole storage yard is divided into small storage locations reserved for different destination vessels. The reservation of storage locations is known as the "yard template". As shown in Fig. 1.7, all different sections of the storage yard are composed of some common basic modules: "sub-block" and "block". The smallest unit for the consignment strategy in yard storage allocation process is a "sub-block". Under the current consignment concept, the containers can only be assigned to the sub-blocks reserved for their own destination vessel.

Within each block, the containers are stacked on top of another by the yard cranes. As shown in Fig. 1.8, a typical block can be described in three dimensions, namely "bay", "row" and "tier". The configuration of a block depends on the yard cranes used for container stacking. The basic unit of the storage space is "slot", which can fit one TEU (20-foot equivalent unit). Several containers stacked on top of each other form a "stack".

The depth of each sub-block is six rows of containers, and the length of each sub-block is eight slots (each slot can accommodate one 20 ft container length-wise). The stacking height is five containers high (which we call tier). A certain number of sub-blocks in a row form a bigger unit, called a "block". There is a dedicated lane for the movement of prime movers (the "truck path") and a separate "passing" lane strictly to allow trucks to pass each other when required. The passing lane is only wide enough for one prime mover and it is shared between two neighboring container blocks.

3. Space allocation

The main purpose of storage yard management is to decide where to store the incoming containers and how to deploy the yard cranes and prime movers to handle the containers. Once the space is allocated to the incoming containers, they will be transported to the corresponding storage locations by the prime movers and stacked by the yard cranes. Thus, the space allocation to incoming containers determines the workload for prime movers and yard cranes at each storage location. The efficiency of storage yard management depends greatly on the space allocation to incoming containers (Lee et al. 2006).

To avoid double handling, the incoming containers are usually stored at the same storage location until being retrieved. The loading activity at each storage location is just a result of the space allocation to incoming containers. Thus, the space allocation

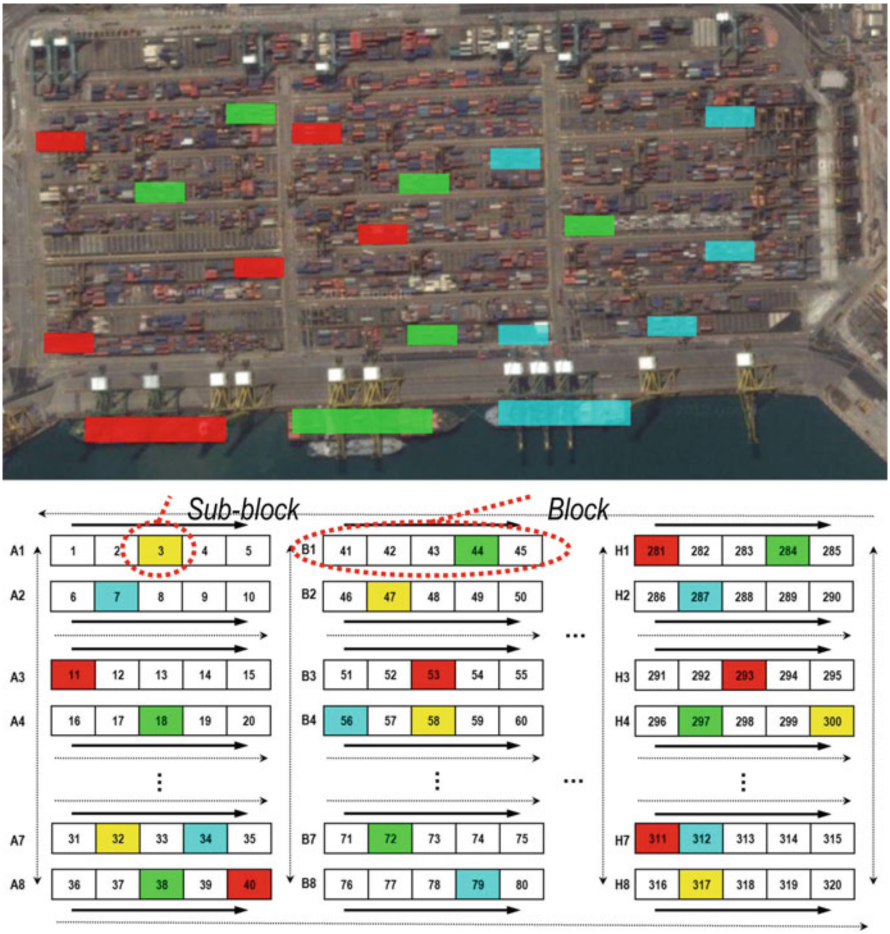


Fig. 1.7 An example of yard template

plan not only needs to consider the discharging of incoming containers, but also has to take into account the loading activities.

Traffic congestion may happen when too much workload needs to be handled within a small area at the same time. For example, if there are a lot of container movements in sub-blocks 7 and 12 (see Fig. 1.7), then there will be many prime movers waiting or moving nearby. This will cause traffic congestion. Similarly, if the workload in sub-blocks 6 and 7 is high, then the prime movers waiting at sub-block 7 may block other prime movers from going to sub-block 6 as they share the same path.

To ensure smooth flow of traffic, the port operator has imposed several restrictions during the planning stage. Among them are:

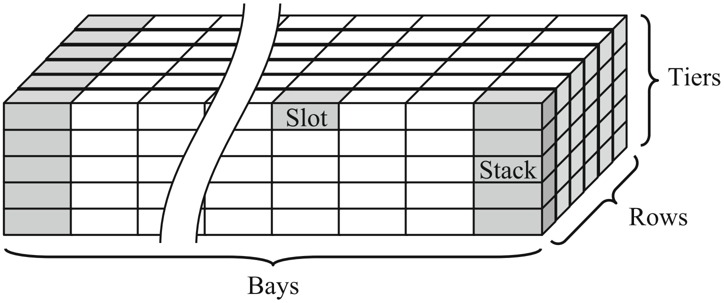


Fig. 1.8 Container stacking in a block

- When a sub-block is in the loading process, its neighboring sub-blocks should not have any loading or unloading activities.
- There should not be two or more neighboring sub-blocks which are having high unloading activities.

To incorporate these restrictions into the planning model, we introduce a "high-low workload" rule/protocol and a vicinity matrix.

The protocol of high-low workload is to ensure that at any given time, many yard cranes will be highly utilized as the jobs are concentrated as they do not need to move around frequently to other sub-blocks to perform jobs. The ranges of high workload and low workload do not overlap. For example, the range of high workload is set between 50 and 100 containers per shift, while the low workload is set between 0 and 20 containers per shift by the port operator.

To capture the possible traffic congestion between sub-blocks, we use a vicinity matrix to represent the neighborhood structure between different sub-blocks. A sub-block is a neighbor if it is adjacent. Adjacency of sub-blocks inherently implies that trucks must use the same path. For example, sub-block seven is said to be a neighbor of sub-blocks 6 and 12. Sub-block 7 is not a neighbor of sub-block 2 even though they are back to back. In a vicinity matrix, a value of one means that the sub-blocks are neighbors to each other, and zero means that they are not. For the layout shown in Fig. 1.9, the vicinity matrix is given in Table 1.5. As the vicinity matrix is symmetric, only the top right half is shown. The workload of a neighboring sub-block should be low if the neighbor has been assigned a high workload.

1.3.3 Developing Operation Strategies for Transshipment Ports

With the increasing volume of transshipment container handling, the scarcity of storage space is urging new studies to improve the land utilization under the complex requirements of transshipment ports. Although the consignment strategy used is an effective way to reduce reshuffles, the prior reservation of storage spaces for each

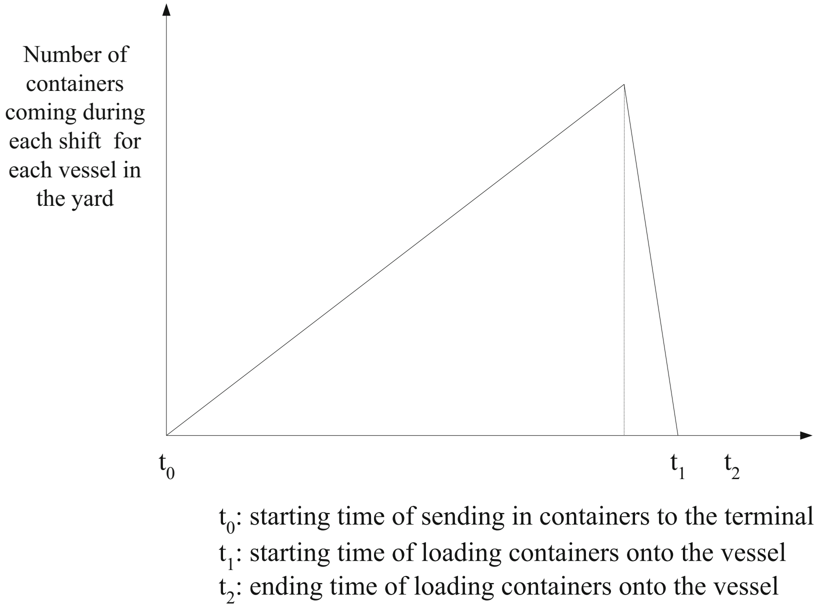
Table 1.5 Part of the vicinity matrix for the yard configuration shown in Fig. 1.5

Rii'	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2		0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3			0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4				0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5					0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6						0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
7							0	1	0	0	0	1	0	0	0	0	0	0	0	0
8								0	1	0	0	0	1	0	0	0	0	0	0	0
9									0	1	0	0	0	1	0	0	0	0	0	0
10										0	0	0	0	0	1	0	0	0	0	0
11											0	1	0	0	0	0	0	0	0	0
12												0	1	0	0	0	0	0	0	0
13													0	1	0	0	0	0	0	0
14														0	1	0	0	0	0	0
15															0	0	0	0	0	0
16																0	1	0	0	0
17																	0	1	0	0
18																		0	1	0
19																			0	1
20																				0

destination vessel causes under-utilization of land due to the fact that the majority of containers usually come in during the time close to their departure date.

According to the workload patterns provided by the port operator, there are two important characteristics of the incoming containers. The characteristics are that the higher incoming workload always happens near to their departure date, while the very low activity happens right after the vessel departs. Hence, it is a common practice for them to use triangular workload profile to do the planning.

For the static yard template (as in Lee et al. (2006) and Han et al. (2008)), all the sub-blocks in each block have a fixed space capacity. This means the maximum amount of space needed at the peak time will be exclusively assigned to each vessel during the whole planning horizon. As much space is only occupied for a short period, it clearly leads to under-utilization of the space. In this section, we propose two space-sharing concepts which aim at improving the use of storage space while ensuring the efficiency of yard operations.



t_0: starting time of sending in containers to the terminal
t_1: starting time of loading containers onto the vessel
t_2: ending time of loading containers onto the vessel

Fig. 1.9 The buildup pattern of the coming workload for one vessel

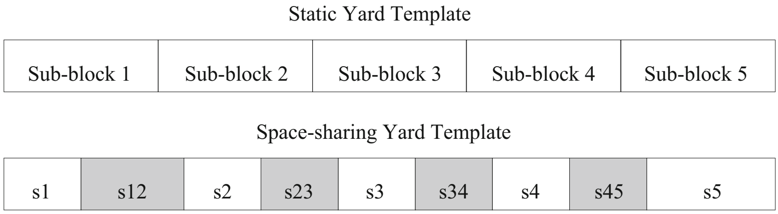


Fig. 1.10 Schematic diagram of one block for the static yard template and the space-sharing yard template

1. The partial space-sharing strategy

To enjoy the benefit of consignment while increasing the land utilization, we propose a space-sharing method which allows some space to be shared between adjacent neighbors. Essentially it will help to reduce the original space needed for a given workload. As shown in Fig. 1.8, for the space-sharing yard template, each sub-block has certain amount of storage space for sharing. For example, s12 is the part that can be shared between sub-blocks 1 and 2 Fig. 1.10.

As very few space is needed during the period right after the loading process, the sharing space of one sub-block can be lent to its neighbors. It will then be returned from its neighbors, before the major workload comes into this sub-block. Since the major workload arrives at different periods for different vessels, they will also need

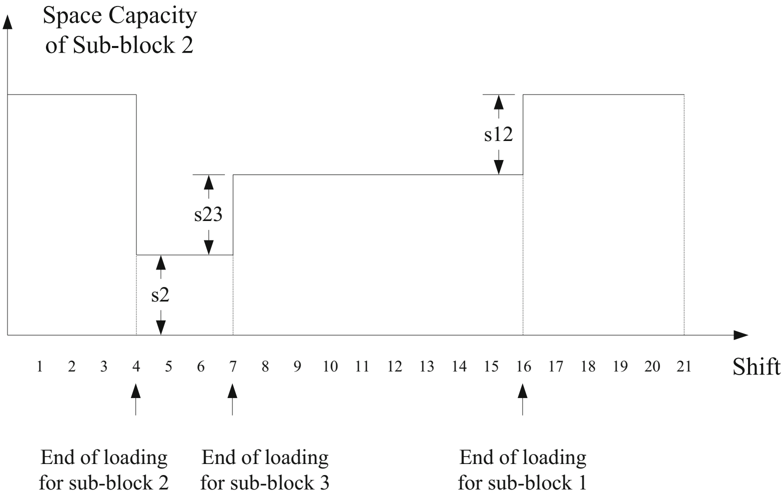


Fig. 1.11 A schematic diagram for space capacity of one sub-block

the sharing space during different shifts. We can take the sub-block 2 as an example to demonstrate how its space can change over time. Suppose that Sub-blocks 1, 2, and 3 have been assigned to different departing vessels, and the starting times of their loading operations are Shifts 14, 2 and 5 respectively. Then, the starting times of sharing the spaces to the neighbors for Sub-blocks 1, 2, and 3 are Shifts 16, 4, and 7 respectively, assuming that the loading operations last for 2 shifts. Since Sub-block 2 has Sub-blocks 1 and 3 as neighbors, the change of its space capacity over the 21 shifts can be plotted as in Fig. 1.11. Similarly, the storage space of all the sub-blocks in one block changing over the 21 shifts is shown in Fig. 1.12. In other words, the space capacity of one sub-block will decrease after the sub-block's loading process, while it increases when its neighbors finish loading. However, the sum of a non-sharing space and its neighboring sharing spaces should be not more than the standard size of a sub-block given by the port operator.

To implement this space-sharing concept, three key issues should be resolved; namely, yard template, size of sharing space and workload assignment.

Since the yard cranes and transporters handle one container at a time, the number of loading and unloading containers in each sub-block can be used to indicate the potential traffic. To ensure a smooth flow of traffic, we adopt the high-low workload balancing protocol and the vicinity matrix from Lee et al. (2006) and Han et al. (2008). The vicinity matrix is used to capture the neighborhood relationship among sub-blocks, while the high-low workload balancing protocol is implemented to avoid potential traffic congestion and to ensure high utilization of yard cranes.

2. The flexible space-sharing strategy

In the previous section, a "partial space-sharing strategy" is proposed to improve the space utilization while retaining the advantage of consignment. Although the partial

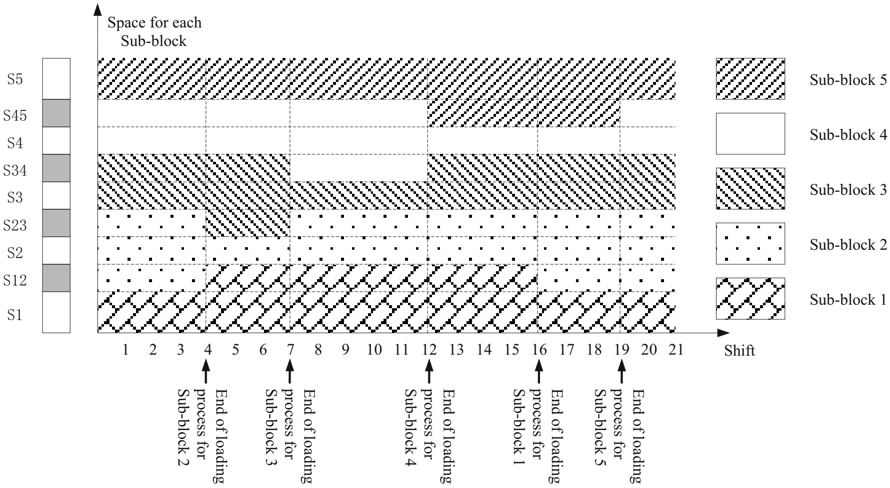


Fig. 1.12 A schematic diagram for the space capacity of each sub-block in one block

space-sharing strategy outperforms the non-sharing strategy, it has some limitations. Firstly, in the partial space-sharing strategy, there is always a clear boundary between the spaces reserved for different destination vessels. The performance of the partial space-sharing strategy depends on the size of sharing and non-sharing spaces. Secondly, once the size of sharing and non-sharing spaces is fixed at the planning stage, it will limit the flexibility of space allocation during operation.

To address this challenge while retaining the feature of the consignment, we propose a new approach named the “flexible space-sharing strategy”. The idea is that the container space can be shared by two different vessels as long as their containers do not occupy the space at the same time. When less space is needed by one vessel, more space can be allocated to another vessel. In this way, the space allocated to each vessel can vary according to the amount of containers stored in the yard. The mechanism of this strategy is described as follows Fig. 1.13.

- Each sub-block is reserved for two different destination vessels, while two adjacent sub-blocks have one vessel in common. (The reservation of sub-blocks for vessels is known as the “yard template”, as shown in Fig. 1.14)
- In each sub-block, one vessel fills from the left corner with incoming containers, while the other vessel fills from the right corner.
- When a vessel is common for two adjacent sub-blocks, the containers to this vessel fills one sub-block from the right corner and the other sub-block from the left corner as shown in Fig. 1.14.

Under this strategy, the space occupation by each vessel in a block can be shown with the example in Fig. 1.14. The containers to the same vessel in two adjacent sub-blocks form a cluster named the “container group”. As a block is managed in five sub-blocks, there will be six container groups (G1–G6) corresponding to the six

(V1)	(V2)	(V3)	(V4)	(V5)
S1	S2	S3	S4	S5

(V1, V2)	(V2, V3)	(V3, V4)	(V4, V5)	(V5, V6)
S1	S2	S3	S4	S5

Fig. 1.13 Yard template for non-sharing strategy (*up*) and flexible space-sharing strategy (*down*)

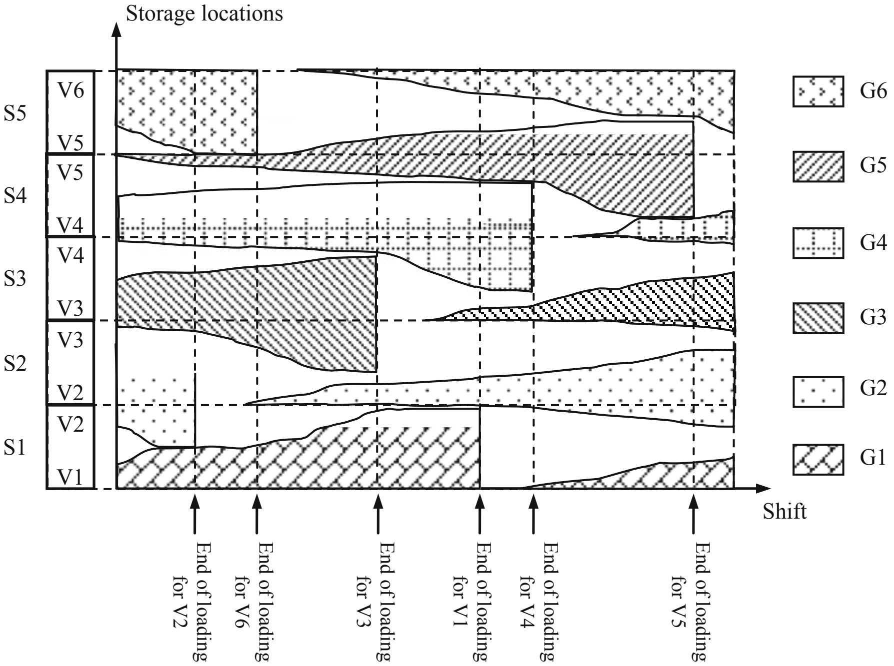


Fig. 1.14 An example of space occupation under flexible space-sharing strategy

vessels in each block. A yard crane will be dedicated to a container group during loading for a faster vessel turnaround time.

1.4 Innovative Terminal Designs

1.4.1 Trend of Future Terminal Designs

Since its introduction in the 1960s, the container has been rapidly dominating the intercontinental cargo transportation scene. During the last two decades, the number of containers transported via seaport container terminals has dramatically increased. Although the global container trade suffered a heavy blow during the 2007–2010 financial crises, the container shipping industry is expected to grow in the long term. Continuous growth is expected in the coming years, especially between Asia and Europe. For cost-effective transportation, mega containerships have been used to transport cargos between container terminals. The loading capacities of containerships have recently been increased up to 15,000 TEUs and even larger. Asian governments have recognized the challenges and opportunities as well as the importance of the development and improvement of seaport container terminals to meet the increasing demand of transferring containers.

However, the expansion of a container terminal is often challenged by the scarcity of land, availability of initial investment, and environmental concerns. To meet the growing demands, port operators need sustainable solutions and we are beginning to see more port operators have the intention to adopt automated container terminals (ACTs) with green considerations. Among them are port operators like PSA (Singapore), Kawasaki (Japan), Antwerp (Belgium), etc.

The first automated container terminal was implemented by Europe Container Terminals (ECT) in 1990s at the Delta Dedicated North Terminal, automated stacking cranes (ASCs) and automated guided vehicles (AGVs) was used. The AGVs are used to transport containers between the quayside and yard side while the ASCs are used to stack containers in yard. Later, Container Terminal Altenwerder (CTA) also introduced ASCs and AGVs to the terminal in Hamburg.

Another commercially used automatic equipment is called automatic straddle carrier, which can both transport and lift/release containers. It allows for decoupling the work flow of transport and crane activities by using buffers at the respective interfaces. The automated straddle carrier system was operated in Brisbane port in 2005. They are also implemented in Maersk APM Shipping Container Terminal Port in Portsmouth Virginia. A study on the efficiency of the transportation equipment (AGV and automatic straddle carrier) was conducted by Park et al. (2007).

In the last decades, several new concepts terminals are proposed but have not been commercially implemented in ports, like Linear Motor Conveyance Systems (LMCS), Automated Storage/Retrieval Systems (AS/RS), Grail Rail (GR). The comparison of the productivity among these ACTs can be found in the paper by Liu et al. (2002).

Most of the existing automated container terminals either use AGVs or ALVs to transport the containers. In this section, we will discuss some innovative designs of container terminals which can be a potential solution for transshipment ports. A million-dollar winning design in a recent “Next Generation Container Port (NGCP)” challenge will also be discussed. We hope these innovative terminal designs can provide some solutions for transshipment ports to fundamentally address the aforementioned challenges, bring container shipping to a whole new level, and satisfy future increasing demand for speedier handling of containers and increased number of containers per unit area of terminal space.

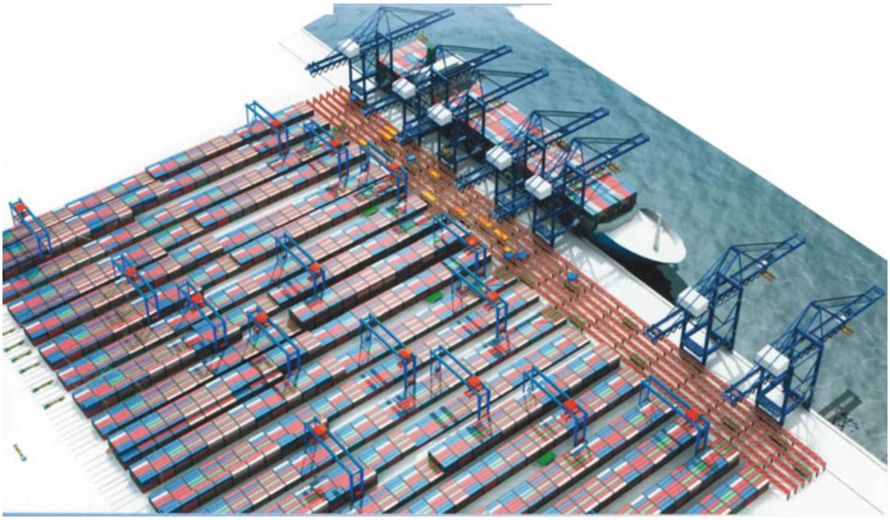


Fig. 1.15 Frame bridge system designed by ZPMC

1.4.2 Innovative Terminal Designs

1. Frame bridge system

Shanghai Zhenhua Heavy Industries Co Ltd has recently designed a real size prototype automated container terminal which utilizes frame bridges, rail mounted frame trolleys and ground trolleys to transport containers between quay side and yard side. A top view of terminal design can be shown in Fig. 1.15, and the major components of the system are shown in Fig. 1.16

We can see that the new designed ACT is composed of three major parts: quayside operation area, transportation area and storage area. Rail bridges are built to transport containers which can be categorized into two types: one is laid parallel to the berth and interfaced with quay cranes, denoted as berth rails; the other one is laid parallel to the yard block and interfaced with yard cranes, denoted as yard rails. The berth rails are constructed above the ground and the yard rails are laid at ground level. These two parts of rail bridges crosses each other perpendicularly. The transfer platform (TP) that sits on the berth rail provides an interface between the berth rail and the yard rail. The trolleys mounted on the berth rails are called frame trolleys (FTs). They are used to transport containers between QCs and TPs. The trolleys mounted on the yard rails are called ground trolleys (GTs). They are used to transport containers between yard cranes and TPs. Since these two parts of rails are perpendicular to each other, the TP is used to rotate the containers 90° during the handover of containers between the two kinds of trolleys. The TP and the two kinds of trolleys cooperate with each other to complete the transportation of a container between quay crane and

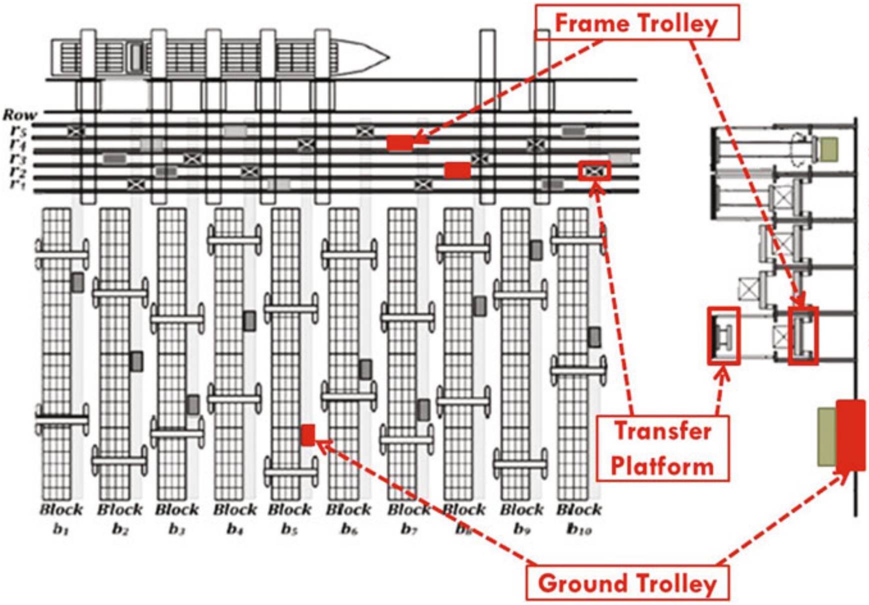


Fig. 1.16 Low frame bridge system designed by ZPMC

yard crane. Since the terminal is covered by frame bridge rails, it is called Frame bridge based ACT (FB-ACT).

Comparing FB-ACT to other designs of container terminals, the advantages can be summary as follows:

- The trolleys mounted on the rails can move at a high speed. It can reach a speed up to 14 mph while an AGV can travel up to 5 mph when they are fully loaded.
- The productivity of yards is increased. In AGV-ACT system, YC needs to travel on average a longer distance to pick up/store containers as the handover of containers between the yard crane and AGV occurs at the end of the block while in FB-ACT, the handover of containers can occur at the side of the block and the GT speed is much faster than the YC.
- This system is a flexible because the capacity can be increase by adding additional layer of rails below or on top of the original rails.
- This system is green and requires less labor. The equipment likes FTs, GTs and TPs are powered by electricity instead of diesel and they are less expensive compared an AGV.

Since the frame bridge system is originally designed for a gateway port, there are also some potential challenges to apply it to a transshipment port. The operation of the FB-ACT for transshipment activities can be challenging. The handshakes among all the devices are one of the major issues to address because more equipment are involved in this system. For a container handled by QC, the handshake is between

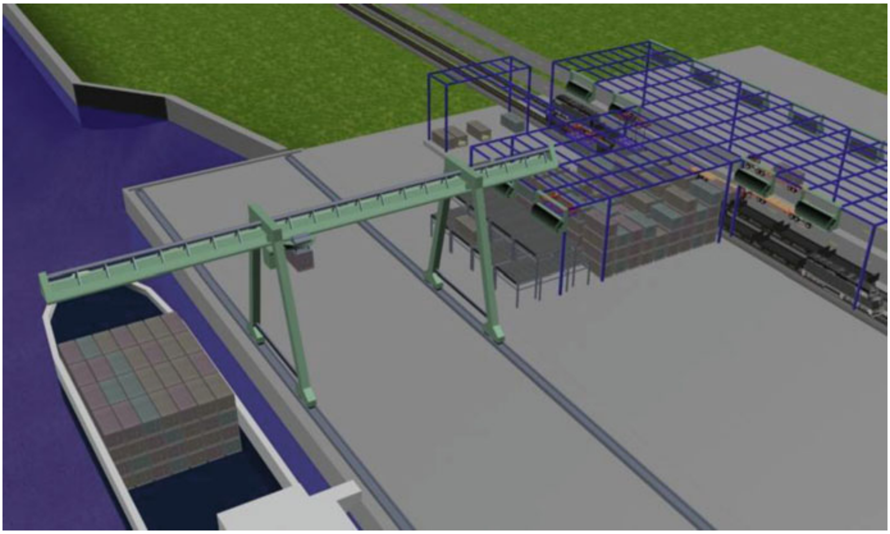


Fig. 1.17 GRID system designed by BEC

QC and FT. For a container handled by TP, the handshake is among FT, GT and TP. Moreover the system needs good traffic control, especially for FTs. Since the trolleys are mounted on the rails, they cannot cross over each. The conflict of the trolleys greatly affects the performance of this system.

2. GRID system technology

A new automated container terminal concept, known as the GRID system, introduced by BEC Industries, could be a potential solution. In the new concept, the container terminal is highly automated with an overhead grid rail structure which covers a wide area of the terminal, directly interfaced with the quay crane, gate buffers and inspection area. The container-handling devices, called Transfer Units (TU), travel on the overhead grid rails and have access to any part of the container yard, thus eliminating the need of ground vehicles. In addition to high automation, the grid system based ACT stands out in maximizing land utilization, two-three times the storage density of typical port layout, which is a vital advantage in consideration of the scarcity of yard land like Hong Kong and Singapore. An example of GRID system layout can be shown in Fig. 1.17.

The advantages of GRID system can be summarized as follows.

- Two-three times the storage density of typical port layout
- Faster ship-to-shore crane cycle times
- Faster simultaneous ship load/off-load
- Eliminate quayside & land side handshakes
- Greater efficiency by parallel loading of trucks and rail
- Powered exclusively by electricity

- Maximizes land utilization
- Reduces operating costs
- Flexible module for different layouts

However, the GRID system technology is still under development. Currently, the traveling speed of the transfer unit is not fast enough and the construction cost of the frames shall be further reduced. There are also some potential challenges to apply such system in a transshipment port. The performance of the GRID System depends on how fast a transfer can handle one container. When both loading and unloading activities are concentrated in the same yard, the conflicts among transfer units will greatly affect the throughput. Effective conflict prevention methods or rules are needed for better performance of the GRID system in a transshipment port.

1.4.3 An Innovative Design of Next Generation Container Port—SINGA Port

1.4.3.1 Background of NGCP Challenge

The Next Generation Container Port (NGCP) Challenge is an international public competition jointly organized by the Singapore Maritime Institute (SMI) and the Maritime and Port Authority of Singapore (MPA) in 2012. Through this competition, the SMI and MPA also aim to raise awareness of the maritime sector and encourage the industry to challenge conventional designs and thinking. It seeks revolutionary new ideas in the planning, design and operations of the next generation of container ports that will achieve a quantum leap in handling efficiency and productivity to support future shipping in an economically and environmentally sustainable manner. The KPIs of the NGCP challenge can be summarized in the following figure. The NGCP challenge welcomes all interested individuals, companies, tertiary institutions and research institutions from around the world. It received a total of 56 submissions from 25 countries Fig. 1.18

Given the operating specifications within limited square land ($2.5 \times 1.0\,\mathrm{km}^2$), several challenges can be identified through analyzing the quay and yard sides requirements on the assumption that the port uses existing systems. From the quay side analysis, it is observed that quay side operations need to have at least 75 quay cranes to attain a productivity of 35 move/hour. From the yard side analysis, it is found that the conventional yard layout would not be able to provide sufficient storage space to accommodate the containers given the annual throughput target underlying the practical storage space allocation strategies. A new design need to address the following issues:

1. New quay crane technology is needed to meet the throughput requirement of more than 35 move/hour.
2. The supporting operations at the yard side need to be fast enough to provide a seamless operation for quay cranes.

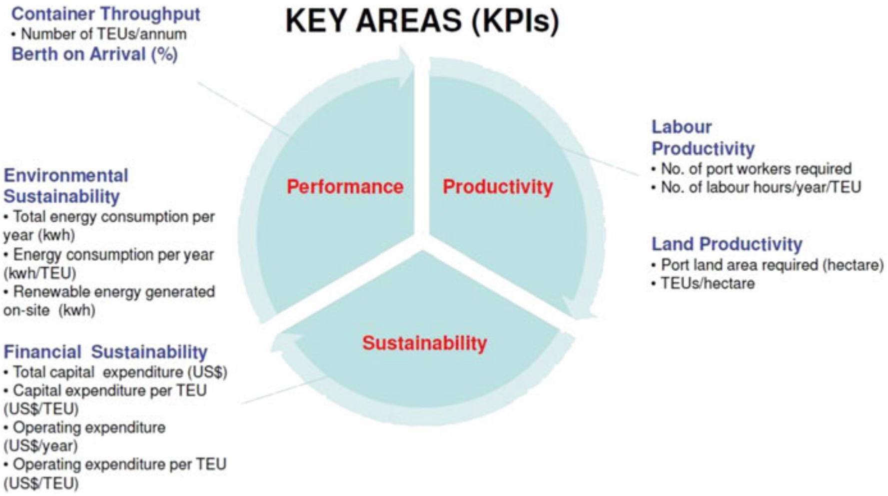


Fig. 1.18 KPIs in NGCP challenge

3. Handshakes between different types of equipment need to be addressed to reduce the waiting times for quay cranes.
4. Sufficient buffer spaces for vehicles are needed to support the quay crane's operation efficiency.
5. Multiple access locations for quay cranes to load/unload containers are needed in order for quay cranes to work with less interruption.
6. An innovative layout design concept to increase the land productivity for storage is needed to ensure a seamless operation at the yard side.

1.4.3.2 Winning Design—SINGA Port

The winning proposal, "SINGA Port", that took the US $1 million grand prize, was a joint work of the National University of Singapore, Shanghai Maritime University and Shanghai Zhenhua Heavy Industries Company Limited (NUS-SHMU-ZPMC). To address the challenges both at the quay side and yard side, the team of "NUS-SHMU-ZPMC" proposed an innovative double storey container terminal design concept, which is illustrated in Fig. 1.19. The double storey terminal will not only able to provide more space for storage and transportation, but also offers more access points for quay cranes because the containers can be loaded and unloaded at both the upper and lower levels. There are also cutouts in the double storey structure, which allow necessary container movements between the first and second storeys. Moreover the second floor provides a natural shelter for the first floor and some storage locations on the first floor can be used to house refrigerated containers and this helps to reduce energy usage.

Fig. 1.19 Double storey container terminal

To reduce the construction cost of the second floor, a creatively designed indented storage yard is proposed for the first floor. This will help to reduce the height of the second floor. The columns which support the second floor will also form a natural support for the possible use of overhead bridge cranes (OBC) on the first floor. With the double storey structure, the terminal will have a surplus of land which can be used to develop an integrated logistics center within the port premises.

At the quay side, the design uses a triple hoist quay crane with tandem lift capability to improve the productivity for the quay side. This triple hoist quay crane is specially designed for the double storey structure which allows the quay cranes to load/unload containers on the first and second floors simultaneously. Moreover, the proposed quay crane can be further improved to allow different quay cranes or the same quay crane to work on consecutive bays simultaneously and this can further improve productivity by reducing interference between quay cranes. In addition, to reduce the waiting time for both quay cranes and yard cranes, the design uses the automated lifting vehicle (ALV) as the transportation equipment to eliminate the handshakes between the handling equipment and the transportation equipment. The double storey container terminal can be fully automated. Detailed features of the double storey container terminal design concept will be discussed in more details.

1. Double storey feature
    – Multiple access points at quay side

A novel double storey terminal provides multiple access points for quay cranes that allow the quay cranes to work continuously with least interruption since the jobs can be spread over a larger area. Moreover, this design concept will also increase the buffer spaces at handover points which can reduce the dependencies between the quay crane and the Automated Lifting Vehicle (ALV). Figure 1.20 shows the multiple access points located on the first floor and second floor.

Fig. 1.20 Multiple access points for loading/unloading containers by quay cranes

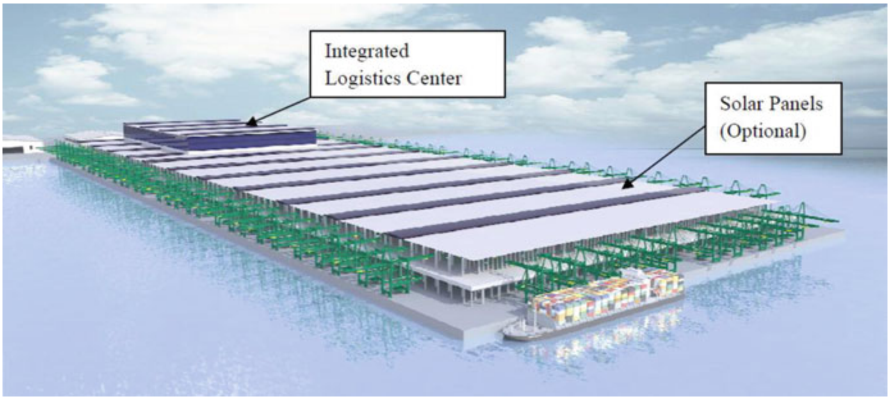


Fig. 1.21 Double storey container terminal with an integrated logistics center

- Additional storage space & Natural shelter

The double storey structure provides additional storage spaces for stacking containers and more road space to ease traffic congestion, thus improving the productivity of the ALV and space flexibility for the construction of an integrated logistics center (see Fig. 1.21) or for future capacity expansion.

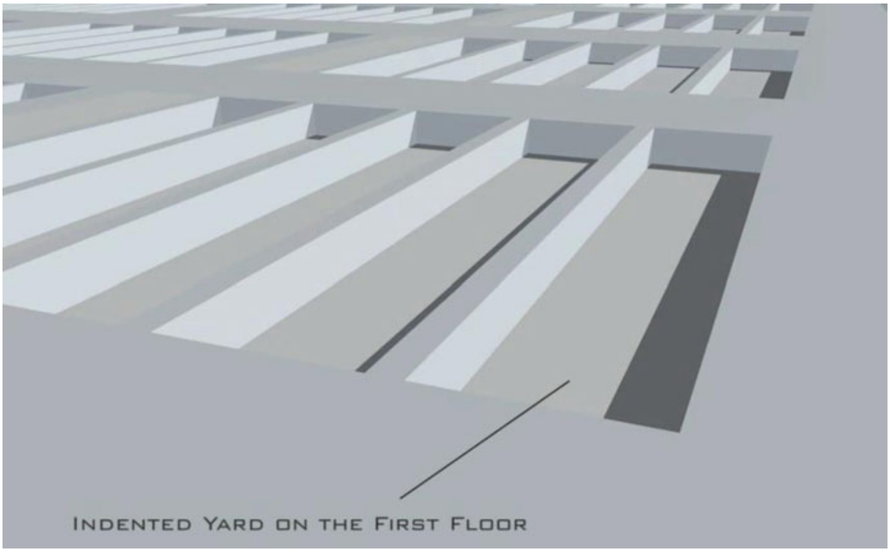


Fig. 1.22 The indented storage yard

The double storey structure provides a natural shelter for the first storey port area. Given the hot weather condition of Singapore, the shelter would help to reduce energy consumption, especially, for the temperature control of refrigerated containers. Moreover, the equipment and vehicles are protected from the harmful UV light, thereby prolonging their life spans.

- Indented storage yard

The second floor of the container terminal needs to be supported by columns. The construction cost increases with the required height and size of the columns. To reduce the construction cost, the design proposed to have the base of the storage area below the ground level which we shall call an "indented yard". If the whole terminal is built on reclaimed land, the indented yard will further reduce the construction cost as less sand is needed to reclaim the land. The design proposed that the depth of the indented yard be three-container high, and so it will reduce the column height by 8.688 m as well as the column cross-sectional area. The reinforced concrete floor and side walls of the indented yard are designed to be impervious to any seawater seepage. Adequate drains, sumps and pumps are installed in the indented yard for water drainage. Moreover, as the transportation lanes for the ALV are on ground level, it will reduce the energy usage for yard crane operations during the loading and unloading processes. Figure 1.22 shows the indented storage yard layout.

- Cutouts between floors

Container handling activities on the first and second floors will be connected by using the concept of "cutouts". This will provide flexibility to the planning because

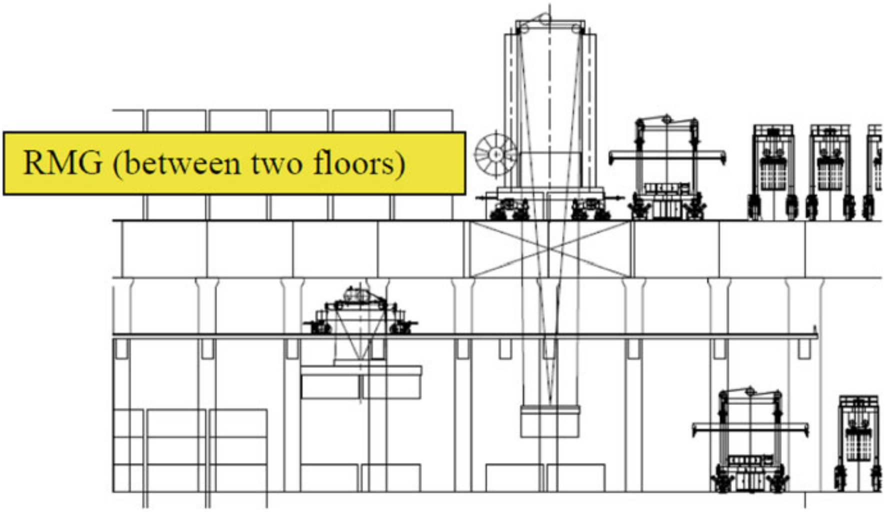


Fig. 1.23 Cutout operations between two floors

containers can be moved between floors and hence the two floors can be treated as one single floor in storage planning. This can help storage planning to be more efficient. In the "cutouts" design, the yard cranes at the second floor can reach down to the first floor to load and unload containers whenever necessary. The construction cost of the "cutouts" will be much lower than building a bridge or ramps for vehicles to move between the two floors. Figure 1.23 shows the "cutout" operations between the two floors.

2. Equipment selection
   – Triple hoist quay crane

To ensure that the aforementioned features can be realized in the container terminal, we need to use the appropriate equipment. The quay crane is the key equipment that determines the maximum productivity of a port. We propose a state-of-the-art quay crane technology which is specially designed for the double storey container terminal. This quay crane is not available in the market at this moment. We shall name this quay crane the triple hoist quay crane with tandem lift. The quay crane can achieve 38 moves per hour and move 152 TEUs per hour. Owing to two separate trolleys that serve the different floors and the transfer platform, the tandem lift technology can now be more practically used because the two containers do not need to be discharged side by side. Another new concept that is currently explored by quay crane manufacturers can further enhance our proposed quay crane system. This concept allows different quay cranes or the same quay crane to work on consecutive bays at the vessel simultaneously and thereby reducing the interference between quay cranes. With this capability, we believe that our proposed triple hoist quay crane will be able to turn the vessel faster. This is especially important for ports serving mega

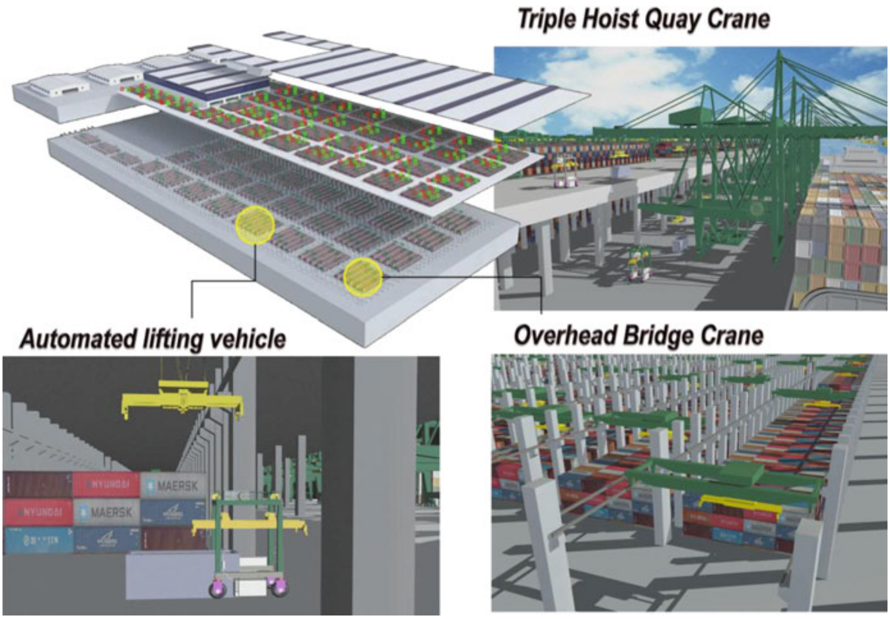


Fig. 1.24 Equipment selection for SINGA Port

vessels as we can reduce the port stay of the vessels significantly by assigning more quay cranes to serve them.

- Yard crane combination—OBC+RMG

For the yard cranes, we use the OBCs for the first floor while RMGs are deployed on the second floor. The second floor is supported by precast reinforced concrete columns. The columns become a natural support for the OBCs. As there are cutouts between the two floors, a tailored made RMG is used to move containers between the two floors besides its normal activity in handling containers in its respective block. Counter weight technology is introduced in the RMG to reduce the energy consumption when lifting the containers.

The weight of the loading unit of the yard crane is usually around 11 t. On the other hand, the average weight of a 20 ft container is around 11 t, which means the loading unit's weight is almost half of the total loading weight. This requires more energy than necessary to lift the containers because of the weight of the loading unit. Hence in order to eliminate this extra weight, a counter weight is introduced to balance the weight of the loading unit during the lifting operations. Connected by a tight wire, the balance weight and the loading unit form a mechanical balance system. When the yard crane lifts a container, it does not consume any extra energy except for the friction force. Thus, it consumes less energy and is more environmental friendly Fig. 1.24.

– ALVs to eliminate handshakes

ALVs are self-lifting vehicles that decouple the handshake between different types of equipment. Quay cranes and yard cranes can directly load or discharge containers on the ground and do not need to wait for the ALVs which help to eliminate handshake between equipment. The proposed ALV can load one container and pass through another container on the ground. The design proposes to use a hybrid ALV or a fully electric powered ALV. For the fully electric powered ALV, there are two potential ways to charge electricity. One is to use a charging station specialized for ALV. Another way is automatically charging the ALV by using the induction coil which is embedded in the traveling path but this charging technology is not feasible now as the charging power is too small. However, we believe that it will be feasible within the next decade due to the rapid development of wireless charging technology.

3. Peripheral support systems

Besides the physical layout designs, it is also important for the port to have a peripheral supporting system that ensures the seamless operations of the port and enhances the competitiveness of the port. In particular, the design considers the smart electric power management, solar energy, the integrated logistics service system and the terminal operating systems (TOS).

– Sustainability and green port

It is important that we use a smart electric power management. In essence, grid-supplied electricity will transform crane operation from a stand-alone mode to a collective mode. Regenerative energy recovered from braking cranes can be shared among other cranes through the grid for achieving common economical and technical benefits as well as better overall energy efficiency.

As Singapore enjoys good sunlight throughout the year, it makes sense to harvest solar photovoltaic energy by mounting solar panels on the terminal roof. Although the estimated amount of harvested solar photovoltaic energy may not be sufficient to meet the entire power demand of the container terminal, it can be used to meet the harmonic power requirements. Based on our computations, if solar panels are installed at the whole plot of land, one should be able to get 405,882 MWh per year.

It should be reiterated that shelter provided by the second floor helps reduce energy usage for refrigerated containers. The indented yard also helps to reduce energy usage when containers are loaded onto or discharged from the ALV. The use of the counter weight balance helps the RMG to further reduce energy usage when it lifts the container.

– Integrated logistics center

The integrated logistics centre provides value-added services, which ideally could be combined together through the logistics information platform so as to achieve value differentiation. Also transportation energy and time are saved by having the logistics center in the port.

- Terminal operating system (TOS)

The TOS is designed to provide a set of computerized procedures to manage the containers, equipment and people within the facility to enable a seamless link to efficiently and effectively manage the facility. The different simulation supporting tools are presented in case studies for integrated planning and detailed operations respectively. Both simulation tools can be incorporated into the TOS.

1.5 Conclusions

With the increasing vessel size and shipping alliance, the "hub-and-spoke" pattern is becoming more popular in the container transport chains. This leads to the increasing importance of transshipment activities and the major ports which provide transshipment services. This chapter provides an in-depth discussion focused on transshipment ports.

Our first topic discusses a major factor influencing transshipment capability, which is known as the port connectivity. The traditional performance measures mainly focus on the local information at each port, such as "number of port calls", "annual container throughput", "berth on arrival" etc. However, such measures though useful they do not measure the larger dimension on how a port impacts the overall network shipping services. To address this, a new connectivity framework is proposed from a network perspective. The proposed framework enables us to possibly sum up the complex interaction among shipping capacities, frequencies and the network structure to generate an intuitive and quantifiable measure of port connectivity.

The second topic discusses the innovative management concepts for transshipment ports. The storage yard management in transshipment ports can generally be divided into three phases, namely "yard sectioning", "template planning" and "space allocation". Generally speaking, the "consignment" concept is applied in transshipment ports to store containers to same destination vessel together. This is to facilitate faster loading process as it reduces reshuffles as well as long distance movements of yard cranes. The whole storage yard is divided into small storage locations reserved for different destination vessels. The reservation of storage locations is known as the "yard template". Three different storage strategies can be used to achieve consignment, namely the "non-sharing strategy", the "partial space-sharing strategy" and the "flexible space-sharing strategy". All these strategies can be used in space allocation to improve container storage and retrieval efficiency in transshipment ports.

The third topic discusses innovation terminal designs that can be potential solutions for transshipment ports. To meet the growing demands, port operators need sustainable solutions and we are beginning to see more port operators have the intention to adopt automated container terminals (ACTs) with green considerations. However, most of the existing automated container terminals either use AGVs or ALVs to transport the containers. In this section, we introduce two innovative ACT systems, namely the "Frame bridge system" designed by Shanghai Zhenhua Heavy Industries Co Ltd, and the "GRID system", designed by BEC Industries, could be a potential solution. Besides, an international port design competition was jointly organized by the Singapore Maritime Institute (SMI) and the Maritime and Port Authority of Singapore (MPA) in 2012. It seeks revolutionary new ideas in the planning, design and operations of the next generation of container ports that will achieve a quantum leap in handling efficiency and productivity to support future shipping in an economically and environmentally sustainable manner. The challenges highlighted by the competition and winning design are also discussed in this section. We hope these innovative terminal designs can provide some insights for transshipment ports to fundamentally address the future challenges.

References

Allison, Tony. (2000). A new era in Asian shipping, Asia Times Online: Southeast Asia, 2 September 2000. http://www.atimes.com/se-asia/BI02Ae06.html. Accessed 3 April 2010.

Chang, Y. T., Lee, S. Y., & Tongzon, J. L. (2008). Port selection factors by shipping lines: Different perspectives between trunk liners and feeder service providers. *Marine Policy, 32*(6), 877–885.

Chou, C. C. (2007). A fuzzy MCDM method for solving marine transshipment container port selection problems. *Applied Mathematics and Computation, 186,* 435–444.

D'Este, G. M., & Meyrick, S. (1992a). Carrier selection in a Ro/Ro ferry trade. Part1: Decision factors and attitudes. *Maritime Policy & Management, 19,* 115–126.

D'Este, G. M., & Meyrick, V. (1992b). Carrier selection in a Ro/Ro ferry trade. Part2: Conceptual framework for the decision process. *Maritime Policy & Management, 19,* 127–138.

Han, Y., Lee, L. H., Chew, E. P., & Tan, K. C. (2008). A yard storage strategy for minimizing traffic congestion in a marine container transshipment hub. OR Spectrum, 30, pp. 697–720.

Kleywegt, A. T., Goh, M. L., Wu, G. Y., & Zhang, H. W. (2002). Competition between the ports of Singapore and Malaysia, Technical Report, The Logistics Institute, Georgia Tech, and The Logistics Institute-Asia Pacific, National University of Singapore.

Lee, L. H., Chew, E. P., Tan, K. C., & Han, Y. (2006). An optimization model for storage yard management in transshipment hubs. OR Spectrum, 28, pp. 539–561.

Le-Griffin, H. D., & Murphy, M. (2006). Container terminal productivity: Experiences at the Ports of Los Angeles and Long Beach. National Urban Freight Conference, Long Beach.

Lirn, T. C., Thanopoulou, H. A., & Beresford, A. K. C. (2003). Transshipment port selection and decision-making behaviour: Analysing the Taiwanese case. *International Journal of Logistics: Research and Applications*, *6*(4), 229–244.

Lirn, T. C., Thanopoulou, H. A., & Beresford, A. K. C. (2004). An application of AHP on transshipment port selection: A global perspective. *Maritime Economics & Logistics, 6,* 70–91.

Liu, C. C. (2008). Evaluating the operational efficiency of major ports in the Asia-Pacific region using data envelopment analysis. *Applied Economics, 40,* 1737–1743

Liu, C.-I., Jula H., et al. (2002). Design, simulation, and evaluation of automated container terminals. *IEEE Transactions on Intelligent Transportation Systems, 3*(1), 12–26.

Park, N. K., Choi H. R., et al. (2007). A study on the efficiency of transportation equipments at automated container terminals. In *Proceedings of the 2007 Annual Conference on International Conference on Computer Engineering and Applications*, World Scientific and Engineering Academy and Society (WSEAS).

Slack, B. (1985). Containerization, inter-port competition and port selection. *Maritime Policy & Management, 12*(4), 293–303.

Tongzon, J. L. (2006). Privatization: The port of Singapore experience. Project Paper of Asia Research Centre, Murdoch University.

Chapter 2
Container Terminal Operation: Current Trends and Future Challenges

Kap Hwan Kim and Hoon Lee

Abstract This study reviews various planning and control activities in container terminals. Decision-making problems for operation planning and control are defined and new trends in the technological development for each decision-making process are discussed. Relevant research directions and open questions are proposed. The functions of the Terminal Operating System (TOS), which is the software used to implement the decision-making processes, are discussed and commercial TOSs are introduced and compared.

2.1 Introduction

As a result of globalization, international trade has greatly increased and containerships have become considerably due to economy of scale. By 2011, more than 100 container vessels larger than 10,000 TEU were in operation and a further 150 were on order. Vessels of 18,000 TEU began to call at Busan from April 2013. High oil prices and labor costs are other important motivations driving changes in the maritime industry. After 9/11, various security measures have been implemented in maritime and port transportation. Carriers and port operators are improving their equipment and operation strategies in order to satisfy the regulations for environmental protection. The logistics market has changed from a supplier-oriented one to a customer-oriented one because the supply of logistics resources has exceeded the demand. Consequently, shipping liners have gained stronger negotiation power over port operators. In some cases, shipping liners demand a high performance level from terminals as part of the contract conditions, and this can include the throughput rate per berth or the turnaround time of a vessel or road trucks.

K. H. Kim (✉)
Department of Industrial Engineering, Pusan National University, Jangjeon-dong,
Gumjeong-gu, Busan 609-735, Republic of Korea
e-mail: kapkim@pusan.ac.kr

H. Lee
Institute of Logistics Systems, Total Soft Bank Ltd., 66-39, 513 Bansong-ro,
Haeundae-gu, Busan, 612-070, Republic of Korea

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_2

By deploying mega vessels on main routes, the requirements for hub ports have also changed. Handling the expected 9,000 moves within 24 h for a vessel of 15,000 TEU calling at a port necessitates about 350 moves per hour per berth, which is more than twice the current productivity in Busan. Such a doubling of productivity will require dramatic innovation in the handling systems or operational methods.

This paper addresses decision-making problems for the operation of container terminals. There have been useful papers which reviewed publications on this issue (Vis and de Koster 2003; Stahlbock and Voss 2008; Schwarze et al. 2012). The main objective of this paper is to introduce current trends of and new challenges to researches in this field. Section 2 discusses the necessary optimization of operational decisions during the operation planning stage and the real-time operation stage. Section 3 introduces the current status of Terminal Operating System (TOS) and suggests potential improvements in TOS. Concluding remarks are presented in the final section.

2.2 Optimizing Operation Plans of Container Terminals

Operation planning is performed for the efficient utilization of key resources during critical operations, which are those closely related to the key performance indices of a container terminal. Examples of operation planning are berth planning, quay crane (QC) scheduling, loading/unload sequencing, and space planning (Crainic and Kim 2007; Kim 2007; Böse 2011). Some resources are classified as key resources because of their high cost and the consequent expense in increasing their capacity. Key resources may include berths, QCs, and storage spaces in most container terminals.

2.2.1 Berth Planning

The berth planning process schedules the usage of the quay by vessels. For the berth planning process, the information on vessel calls (ship ID of each call, the route, ports of the call, etc.), vessel specifications (length, width, tonnage, etc.), and hatch cover structure are transferred from a corresponding shipping line to the terminal. The information is then registered into the berth planning system of the terminal. The berthing positions for some vessels are pre-allocated at dedicated berths which are based on the contracts between shipping lines and the terminal.

Berth planning is the process of determining the berthing position and time of each vessel and the deployment of QCs to the vessel in a way of maximizing the service level for container vessels. It is desirable that vessel operations are completed within an operation time pre-specified by a mutual agreement between the corresponding ship carrier and the terminal operator. The QC deployment that determines the start and the end times for a QC serves a vessel and must satisfy the limitation in the total number of available QCs. Berth planning and QC deployment are inter-related because the number of QCs to be assigned to a vessel affects the berthing duration of the vessel. In addition, when the outbound containers for a vessel have already arrived at the yard, the vessel berths should be close to the storage area with the outbound containers.

A popular objective function is to minimize the total tardiness of the departures of vessels beyond their committed departure times and each vessel has a different importance to the terminal operator depending on the bargaining power of the corresponding carrier. The second popular objective is to minimize the total flow time of vessels, which means the total turnaround time of vessels (Park and Kim 2003). In addition, there are different types of constraints that must be considered when determining the berthing positions of vessels such as the depth of water along the quay and the maximum outreach of QCs installed at specific positions on the quay. Further issues for consideration are presented below.

Continuous Quay Assumption Berth planning is a well-defined problem much discussed in the literature. The quay may be considered to be the set of multiple discrete berths or a continuous line on which a vessel can berth at any position. The berth planning problem had been considered to be an assignment problem of each vessel to a berth under the assumption of discrete berths, whereas some researchers have recently started to consider the problem of determining the exact position of each vessel on a continuous quay (Imai et al. 2005; Lee et al. 2010).

Dynamic Berth Planning and Re-planning A container terminal makes a contract with shipping lines for regular calling services, weekly in most of cases. Because ships' arrival times, which depend on weather conditions, ships' operating environment, or the departure time from the preceding port, and the working conditions of the current terminal may change at any time, the berthing times and vessel positions need to be continuously changed. Therefore, planning processes and algorithms need to be studied considering these situations. The robustness may be an important property ofa good berth plan (Hendriks et al. 2010).

Considering Traffic in the Quay and the Yard At multi-berth terminals, berth planning is conducted to minimize any interference between docked vessels and berthing vessels, which may happen during the arrival and departure of vessels. When the traffic of containers for two vessels cross in the yard during shipping operation, the interference between the traffic may seriously delay the ship operation. Transshipment containers may be a source of traffic to be considered. These factors need to be considered during berth planning for more efficient operation of terminals.

Considering Tidal Difference Ports with a large tidal difference have a further issue requiring consideration during the berth scheduling for large vessels. Some ports have bridge piers to overcome the large tidal difference. Even so, berth planners must consider the water depth at the vessel arrival and departure times in order to confirm berthing feasibility. The changing water depths of the channels for vessels to approach the terminal also need to be considered in some ports. Many container terminals have similar restrictions in the timing of berthing or de-berthing.

2.2.2 Stowage Planning

Stowage planning is the process of specifying the attributes of containers to be loaded into slots in a ship bay. For some containers already loaded in the vessel, relocations within the vessel or via temporary storage areas at the apron are planned for more efficient ship operations in succeeding ports. The stowage plan, which is usually constructed by vessel carriers, does not specify each individual outbound container to be loaded into each slot.

During the stowage planning process, the rehandling of containers bound for succeeding ports must be considered. Thus, it is necessary to locate containers that are bound for preceding ports in higher tiers and locate those for succeeding ports in lower tiers. In addition, various indices of vessel stabilities and strengths must be checked. The positions of the inbound and outbound containers are preferably distributed as widely and evenly as possible over the entire range of a vessel in order to reduce the possibility of interference among QCs during the ship operation (Imai et al. 2002, 2006; Ambrosino 2006; Sciomachen and Tanfani 2007).

2.2.3 QC Work Scheduling

In order to discuss the loading and unloading operations, we introduce the concept of "container group." Outbound containers of the same size and with the same destination port, which have to be loaded onto the same ship, are categorized under the same container group. Likewise, inbound containers of the same size that have to be unloaded by the same ship are said to be categorized under the same container group. Containers in the same group are usually transferred consecutively by the same QC.

When the discharging and loading operations must be performed at the same ship bay, the discharging operation must precede the loading operation. When the discharging operation is performed in a ship bay, the containers on the deck must be transferred before the containers in the hold are unloaded. Further, the loading operation in the hold must precede the loading operation on the deck of the same ship bay. It should also be noted that the QCs travel on the same track. Thus, certain clusters of slots cannot be transferred simultaneously when the locations of the two clusters are too close to each other, because the two adjacent QCs must be separated by at least a specific number of ship bays so that the transfer operations can be performed simultaneously without interference (Moccia 2006; Sammarra et al. 2007; Lee et al. 2008; Meisel 2009).

In practice, one example of a QC scheduling process may be described as follows: a QC work sequence is decided for tasks divided by hatch cover and hold/deck of a vessel. A basic sequencing rule is to sequence unloading tasks from the stern to the bow, and loading tasks from the bow to the stern. The most popular criterion is to finish all the tasks by multiple assigned QCs at the same time. Thus, the entire loading and unloading tasks are allocated to each QC by splitting the working area with two boundaries of the hatch for both hold and deck so that the amount of work allocated to each QC is as similar as possible among different QCs. More complicated characteristics of the QC scheduling problem are considered below.

Reduction of Planning Lead Time The cutoff time within which outbound containers are allowed to be delivered into the yard is mainly due to the time needed for the ship operation planning. A reduction in planning time would therefore reduce the cutoff time and hence improve the customer service level. Such efficiency gains can be achieved by automating the scheduling process.

Simultaneous Planning of Quay Side and Yard Side Operations If the containers for any two clusters of slots have to be picked up at or delivered to the same location in a yard, the tasks for the two clusters cannot be performed simultaneously due to the resulting interference among the corresponding yard cranes (YCs). Thus, for QC scheduling, any potential interference between YCs needs or congestion in a yard area to be considered simultaneously (Jung and Kim 2006; Choo et al. 2010; Wang and Kim 2011).

Integration with Real-Time Operation Control Function and Load/Unload Sequencing Process The real-time ship operation may not progress as planned in the QC schedule due to unexpected delay of lashing operation, delay of yard operation, and uncertain operation time of QC operators. Thus, the real-time progress needs to be considered in the QC schedule, which should be able to be updated whenever a significant disturbance happens in the QC operation.

In practice, the load/unload sequencing is done with the result of the QC schedule as a constraint to be satisfied. However, there may be cases where a minor modification of a QC schedule can significantly improve the load/unload sequence. Thus, a better schedule will be obtained if the QC scheduling is done together with the load/unload sequencing.

Increasing the Adaptability and the Rescheduling Capability of QC Scheduling Module Generally, multiple QCs are assigned to a ship. When an operation of a specific QC is delayed or a QC is broken down, the workload among QCs becomes unbalanced or the QC schedule may become disturbed. Such unbalance and disturbance may cause unexpected interference between QCs during the operation. Lashing or un-lashing operation can be delayed for containers on board. During the discharging operation, a specific container on board may have an unexpected difficulty during the un-lashing operation. A popular way to overcome these difficulties is for an under-man or a ship planner to change the work schedule adaptively. It would be helpful for a ship planning system to have the capability of automatically changing the work schedule adaptively.

Providing a Planning Process for Multiple Planners for Multiple Vessels Considering the Shared Resources Among them A popular way to support multiple planners in constructing operation plans for multiple vessels is to specify a planning boundary for each planner in the stowage plan of a vessel in order to remove conflicts between planners. In addition, the system provides the function of temporarily locking and unlocking data lists in the data base corresponding to overlapping parts in the stowage plan shared by multiple planners. However, these methods guarantee the optimality of the schedules not from the viewpoint of individual planners but from the system's viewpoint. They are usually sharing the same resource such as storage spaces and handling equipment at many different time periods. However, the sequential or random decision making by planners may not lead to the optimal decisions of the system.

2.2.4 Load/Unload Sequencing

After constructing the QC schedule, the sequence of containers for discharging and loading operations is determined. It specifies the slot in the vessel into which each outbound container should be loaded and the loading sequence of the slots (containers). The loading sequence of individual containers significantly influences the handling cost in the yard. Researchers have focused on the sequencing problem for loading operations compared to discharging operations, because determining the discharging sequence is straightforward and determining the stacking locations of discharging containers is usually done in real time. However, in loading operations, containers to be loaded into the slots in a vessel must satisfy various constraints on the slots, which are pre-specified by a stowage planner. In addition, since the locations of outbound containers may be scattered over a wide area in a marshaling yard, the time required for loading operations depends not only on the transfer time of QCs and but also on that of YCs. The transfer time of a QC depends on the loading sequence of the slots, while the transfer time of a YC is affected by the loading sequence of the containers in the yard (Jung and Kim 2006; Lee et al. 2007).

In practice, the sequencing is done in the following process: when a vessel is berthed starboard against berth, unloading work sequences in a bay profile at deck are sequentially decided from starboard to portside. A container lashing operation is to remove fixation devices (corn, lashing bar, etc.) before unloading operation and to fix them after loading operation. While planning the unload (load) sequence, consideration needs to be given to the removal (fixing) of corns and lashing bars from (to) containers on the board of a vessel, which tends to move in the horizontal direction at tier by tier. The unloading or loading work sequence in a bay profile at hold tends to move in the vertical direction stack by stack. The loading plan should satisfy the general stowage plan, which is received from the shipping line and specifies the port of destination and the weight group of the container to be loaded onto each slot. Of course, the travel distances of trucks during the ship operation and the re-shuffling for picking up the container should also be considered for the load sequencing.

Further issues for consideration are presented below:

*Postponement of Decisions on Sequencing Containers and Assignment of Slots to Containers*Traditionally, the loading sequence plan is constructed so that containers are loaded at fixed cell positions in a fixed sequence. However, to give more flexibility during the loading operation, it would be better if slot positions for outbound containers or the loading sequence for the corresponding slots can be changed adaptively.

The sequencing of slots for loading operation is constrained by some precedence constraints arising from their relative physical positions between two slots. An example is two slots in the same stack, in which the slot in the upper tier must be filled after the one in the lower tier is filled. If two slots in the same sequence list are in two different stacks, then the sequence between them may be changed, which may be finally determined at a latest possible moment. The strategy of utilizing this flexibility of loading sequence is called "flexible loading."

The other strategy to improve the adaptability of the plan is "category loading," in which case the planner creates a category consisting of multiple containers with the same attributes and the assignments of the containers in the same category to specific slots can be changed during the real-time operation. The strategy of "flexible category loading" is the combination of two strategies of "flexible loading" and "category loading," in which the decisions on the loading sequence as well as slot positions of containers in the same category are postponed until the loading operations for the slots are performed. A typical example to apply this strategy is empty containers.

Progressive Planning In principle, the loading and unloading sequence is constructed before the ship operation starts. However, the container terminal may be requested to allow arrivals of containers later than the cargo closing time. To cope with late arrival containers, the ship planning module should be able to construct the schedule incrementally. The loading plan for some part of the stowage plan may be constructed after the part of the discharging operation is performed. Progressive planning is the strategy of constructing operation plans whenever necessary.

Considering Lashing Operations and the Structure of Cell Guides The discharging and loading sequence of containers is heavily influenced by the lashing operation and the locking or unlocking of cones. On the deck, the sequence of loading or discharging tends to proceed in the horizontal direction, while it proceeds in the vertical direction in hold. These operation details need to be considered in the sequencing algorithm.

Supporting Tandem or Twin Lifts Spreaders for QCs have been improved so that they can handle various combinations of different sized containers. Spreaders with the capacity and flexibility to handle all possible combinations of 20-, 40-, and 45 ft containers quickly and efficiently have been developed. Some spreaders can also handle four 20 ft containers simultaneously and separate the two 20 ft containers longitudinally between 0 and 1.5 m.

When the twin lift loading or unloading operations for 2 × 20 ft containers are performed by a QC, it will be efficient if yard trucks (YTs) can perform twin carries with 2 × 20 ft containers. When the tandem lift loading or unloading operations for 4 × 20 ft containers are performed by a QC, it would be helpful if two YTs can be dispatched at the same time, each of which performs a twin carry with 2 × 20 ft containers. These types of operations propose new and challenging problems for the vehicle dispatching process.

Dual Command Cycle Operation Usually, a QC spreader reciprocating motion for the unloading or loading operation handles one container at once. This method will be referred to as the single cycle. Dual command cycle operation handles one loading container and another unloading container in its return path in order to handle a total of two containers in a cycle (Goodchild 2005; Goodchild and Daganzo 2006; Goodchild and Daganzo 2007; Zhang and Kim 2009). Besides QC, this procedure is equally applicable to the operation of YT and YC. Even without additional investments in equipment, this method is a productivity improvement technique that uses the facilities and existing equipment and can be expected to save costs and increase productivity. When a QC performs its operation in a dual command cycle, a YT can receive a discharged container just after delivering an outbound container to the same QC, which enables the YT to perform its operation in a dual command cycle. The same improvement may be possible during the transfer operation between YCs and YTs. When one QC is performing its loading operation in single command cycles and the other QC is doing its unloading operation in single command cycles in an adjacent location, a truck can deliver a loading container to the former QC and then receive a discharged container from the latter QC so that a dual command cycle operation can be implemented. Several studies have attempted to maximize the number of dual command cycles of QCs but relatively fewer studies have examined the dual command cycle operation for YTs and YCs.

2.2.5 Space Planning

Yard planning is the pre-planning of a space for temporarily storing containers discharged from a vessel or that for outbound containers carried in from the gate. A yard management system is operated for efficient operation of handling equipment in the yard, monitoring of the utilization of the yard space, and quick identification of the inventory level of containers. Reefer containers are stacked at an area with power supply equipped racks, and hazardous cargo containers are stored in segregated areas based on IMDG segregation rules. Empty containers are usually stored in a segregated area with reach stackers or top handlers.

Yard planning can be divided into two stages: the space planning stage and the real-time locating stage. In the space planning stage, the storage space is pre-planned and reserved before the containers arrive at the yard. However, the specific storage location of each individual container is determined when each inbound container is discharged from a vessel or when each outbound container arrives at the gate. Storage space for outbound containers is planned in advance. However, the storage location for inbound containers is determined in real time. Thus, the space planning stage for inbound containers does not usually exist in many terminals. The four popular objectives of space planning for outbound containers are: (1) minimizing the travel distance of transporters between the yard and the corresponding vessel, (2) minimizing the movements of YCs, (3) minimizing the congestion of YCs and transporters in the yard, and (4) minimizing the number of relocations.

With regard to the first objective, the outbound containers are usually stacked in positions close to the berthing position of the corresponding vessel. For the second objective, the speed of the transfer operation can be increased if the containers are transferred consecutively at the same yard-bay, which is possible because the gantry travel of the YCs can be minimized. Thus, the outbound containers of the same group are usually located at the same yard-bay (Woo and Kim 2011).

Congestion is another important obstacle which lowers the productivity of the yard operation. Thus, a rule to reduce congestion is to spread the workload over a broad area in the yard (Lee et al. 2006; Bazzazi et al. 2009; Jiang et al. 2012a, b; Won et al. 2012; Sharif and Huynh 2013). Another important objective of locating containers is to minimize the possibility of relocations during retrievals (Wan and Tsai 2009; Dekker 2006). When locating outbound containers, the weights of the containers must be taken into account (Woo and Kim 2014). For maintaining vessel stability, heavy containers are usually placed in low tiers of the holds, must therefore be retrieved earlier than light containers from the yard, and hence must be stacked in higher tiers than light containers so that relocation can be avoided during their retrieval. For inbound containers, more frequent relocations are expected because the retrieval requests are issued in a random order by randomly arriving road trucks (Sauri and Martin 2011). Storage charge may be used to control the inventory level of inbound containers (Lee and Yu 2012).

The decision-making problem related to space allocation is not well-defined compared with other decision-making problems for the operation of container terminals, partially due to the difficulty in evaluating the result of the decision making. Decision-making rules that are used in practice depend highly on the terminal or on the decision-makers and thus differ from one terminal to another. They are difficult to be justified and conflict with each other in many cases. Consequently, this is a research area worth of investigation for researchers.

2.2.6 Potential Improvements in Operation Planning Processes

Integrating Planning Activities Operation plans are usually constructed in a hierarchical way. The plan in the highest hierarchy is the berth plan, followed by the QC schedule and the space plan. The load/unload sequence is determined based on the QC schedule. The load/unload sequence is basic information to construct the real-time schedule for handling equipment. The plans in the higher hierarchies become the constraints to the plans in the lower hierarchies. Because of this hierarchical decision-making structure, some serious problems may arise in the lower hierarchy of a plan, which could be solved by a minor modification of a plan in an upper hierarchy. The integration among planning activities in different hierarchies may improve the quality ofvarious operational plans.

Enhancing Rescheduling Capabilities Situations in the terminal are continuously changing. Thus, plans constructed based on the situation at a certain moment in a previous time may not remain valid throughout the implementation period of each plan. When the progress of the operations deviates too much from a plan, the plan needs to be revised. The revision process should be sufficiently fast and should not disturb the various on-going operations.

Automating the Operation Planning Process The cutoff time for outbound containers, which specifies the latest time when outbound containers can be delivered to the yard, is specified for planning of the ship operation. Normally, it takes 5–6 h for the ship planning for one vessel. Thus, if we can reduce the planning time, then a longer cutoff time may be suggested by the terminal operator to shippers, which is a service level improvement. The planning time may be reduced by automating the operation planning process.

Sharing Information on Resources Among Planners The various kinds of planners have different duties. A vessel planner is in charge of a vessel for the planning of the ship operation for the vessel. A yard planner is in charge of allocating storage space to various inflows of containers. At a first glance, although they are in charge of planning different operations, they share the same resources in many cases. For example, the yard space and the handling capacity of YCs are shared by different vessel planners and the yard planner. That is, if one planner uses more, then the other planners have to use less. However, the information on the usage of the shared resource is not usually transparent to all the planners. Various ways to make the availability of shared resources open to all the related planners need to be developed.

Evaluating Plans in Advance When too many uncertain factors or unexpected events that had not been considered in the plan arise during real-time operation, the gap between the plan and the real progress may be very large, which significantly degrades the quality of the plan. Thus, in many terminals, the plan is evaluated by using a simulation technique.

Collaborating with Outside Partners Possible improvements can be made by collaborating with outside partners including trucking companies, vessel liners, barge operators, rail operators, shippers, and forwarders. The collaborating activities may include information sharing, improving data accuracy, integrated scheduling, and devising economic measures for the collaboration (Lee and Yu 2012).

2.3 Real-Time Control

The plans in the previous section are constructed for critical resources (berths, QCs, and, in some cases, storage spaces) and tasks (loading and unloading operations). However, it is impossible or impractical to plan all the details of handling activities in advance. Thus, for the remaining activities, decisions on the utilization of equipment and the assignment of tasks to each piece of equipment are usually made on a real-time basis. Examples include the assignment of tasks to transporters, the assignment of tasks to YCs, and the assignment of specific storage positions for incoming containers. Two reasons for these activities not being pre-planned are the high uncertainties of the situation and the lower importance of the resources, as compared with the importance of resources like berths or QCs. In decision making, although a schedule can be constructed for the events of the near future (less than 20 min into the future), these decisions are essentially made in response to an event that has occurred at that moment. Further, even the decisions included in the various plans can be modified and updated during the implementation, responding to the deviation of the situation from expectations or forecasts (Kim 2007).

The real-time control function became a critical issue with the increasing trend toward automation in advanced container terminals. Unlike traditional terminals, most real-time decisions need to be made by computer software, which must affect the performance of automated container terminals significantly. Because more than one type of equipment is involved in the terminal operation, coordination and synchronization are crucial for obtaining a high level performance. Furthermore, many unexpected events may arise and the operation time of equipment is not certain. As a result, the application of optimization to the decision-making problem is very complicated.

In spite of the complexity of the decision-making problems during the real-time control, in order to improve the agility of the control decisions, many functions, which had been conducted by operational planning systems, are being transferred to the functions of the real-time control. For example, space allocation tends to be done in a real time rather than an operation planning function. In addition, due to the improvement of information technologies, more real-time information on logistics resources has become available. The real-time control system should be able to utilize the real-time location information which became available from advanced information technologies.

Table 2.1 shows the various functions of a real-time control system. The control functions may be viewed from the perspectives of the operations and of the resources. From the former, the control system monitors and controls the operations at the gate side and the vessel side. The control of the gate side is relatively simple. The system controls the flow of road trucks from the gate and to the storage yard and vice versa.

Congestion in the yard is the most important consideration for trucks with outbound containers. The truck is routed to the block that has the lowest work load at the time of the arrival of containers, if the block has an empty space reserved for the group of containers corresponding to the arriving container. Controlling the flow of trucks for inbound containers is simple because the trucks have no choice in terms of selecting the storage location of the container being carried. The major performance measure for the carry-in and carry-out operation is the turnaround time of trucks in the terminal. However, a lower priority is usually given to the gate side operation

Table 2.1 Various control activities in the operation system

Classification	Functions	Decisions to be made
Ship operation	Berth monitoring	Problem detection, alerting & solving
	Load & discharge control	Operation scheduling
	QC operation control	Equipment scheduling
	Transporter control	
Hinterland operation	Transport monitoring	Problem detection, alerting & solving
	Gate management	
	Barge management	
	Rail operation management	
Yard operation	Yard monitoring	Problem detection, alerting & solving
	Yard positioning	Real-time container positioning
	House-keeping	Re-marshaling & shuffling
	Reefer operation control	YC scheduling
	YC control	
Resource control	Equipment management	Workforce & equipment deployment
	Operator management	

than to the vessel side operation because the control problem of discharging and loading containers is complicated but more important.

The task scheduling problem may be defined as follows: task assignment is conducted in two steps: equipment deployment and task scheduling. The former involves the deployment of a certain group of equipment pieces to specific types of tasks. For example, a group of YCs may be dedicated to delivery and receiving tasks for a certain period of time, and a group of YTs may be assigned to the task of delivering a group of containers from one block to another for a certain period of time. This type of decision must be made before the start of the real-time assignment of tasks to each piece of equipment (Zhang et al. 2002; Linn and Zhang 2003).

Unlike the hinterland operation, the vessel side operation must be carefully scheduled. The discharging and loading tasks are decomposed into the elementary tasks for QCs, transporters, and YCs. These new tasks are then scheduled. The task scheduling problem involves the assignment of tasks to each piece of equipment and the sequencing of the assigned tasks to be carried out. The unloading and loading tasks introduce the following two considerations for the scheduling. Firstly, because the most important objective of the unloading and loading operations is to minimize the turnaround time, the maximum make-span of QCs may be minimized as a primary objective. However, because we are considering only 5–10 tasks among several hundred assigned to each QC, it may be more reasonable to use the total weighted idle time as an objective term instead of the maximum make-span of QCs. Instead, the higher weight can be assigned to the QC whose operation is delayed longer compared with the other QCs. Secondly, because the loading and unloading operations are performed by QCs, YCs, and transporters together, the activities of these types of equipment must be synchronized with each other. During the loading operation, it is important for trucks with containers to arrive at the QC in the right sequence. When QCs are performing their operation in twin or tandem lifting type, then the corresponding multiple transporters should arrive almost simultaneously in order to minimize the waiting time of transporters. This scheduling problem considering handover of a container between different types of equipment have not been paid attention to so much so far (Chen et al. 2007; Lau and Zhao 2008). The transporter scheduling can be integrated with the storage location determination (Lee et al. 2009; Wu et al. 2013).

There have been many researches on dispatching delivery tasks to transporters (Briskorn et al. 2006; Liu and Kulatunga 2004; Ng et al. 2007; Angeloudis 2009; Yuan 2011). Two strategies are used when assigning delivery tasks to transporters: the dedicated assignment strategy and the pooled strategy. In the former strategy, a group of transporters is assigned to a single QC, and they deliver containers only for that QC. In the latter strategy, however, all the transporters are shared by different QCs, so that any transporter can deliver containers for any QC; hence, this is a more flexible strategy for utilizing transporters (Nguyen and Kim 2013).

New and recently introduced equipment capable of moving multiple containers in a single cycle includes twin lift and tandem lift QCs, multi-load transporters, and twin lift YCs. Such equipment upgrades have necessitated new operation methods (Grunow 2004).

Further, the YTs and automated guided vehicles (AGVs) can load or unload containers with the help of cranes, while the straddle carriers (SCs) and shuttle carriers can not only deliver containers but also pick them up from the ground by themselves. Thus, although the containers can be transferred by a QC to a YT or AGV only if the YT or AGV is ready under the QC, the operations of SCs and QCs (or YCs) do not have to be synchronized, which results in a higher performance than that of YTs or AGVs. This difference between the two types of transporters requires operation methods that are different from each other (Yang et al. 2004; Vis and Harika 2004).

When automated guided transporters are used, the traffic control problem becomes a critical issue that must be addressed to ensure the efficiency of operations. Due to the numerous large transporters, special attention must be paid to prevent congestion and deadlocks. The transporters in container terminals are free-ranging vehicles that can move to any position on the apron with the help of GPS, transponders, or microwave radars. Thus, the guide path network must be stored in the memory of the supervisory control computer. Once the guide path network is designed, the route for a travel order can be determined. The guide path network and the algorithm to determine the routes of transporters impact the performance of the transportation system significantly; this is another important issue that should be investigated by researchers (Evers and Koppers 1996; Möhring 2004; Vis 2006).

For the efficient operation of yard cranes, scheduling problems have been studied (Ng and Mak 2005; Lee et al. 2007; Murty 2007; Li et al. 2009; Huang et al. 2009). Further, new conceptual YCs that have recently been introduced include overhead bridge cranes that are being used in Singapore, two non-crossing rail mounted gantry crane (RMGC) in a block, two crossing RMGC in a block, and two non-crossing RMGCs with one additional crossing RMGC. New operational methods must be developed for the efficient operation of these new conceptual YCs (Kemme 2011, 2012). Impacts of different yard layouts on the operational performance of the yard needs to be studies further (Petering 2013; Lee and Kim 2010, 2013).

Some general guidelines for improving real-time control are discussed below.

Planning Principle: Schedule Activities Ahead Most real-time control functions have been performed by human operators or supervisors. For example, the location decision for an arriving container has been done by a human operator and dispatching of internal trucks has been done by a supervisor under each QC. The decision is made for the action to be taken immediately but not for a future action. However, some decisions should be made in advance for preparing future actions. For example, trucks for receiving discharged containers should be sent to the corresponding QC in advance before the QC starts releasing the containers onto the trucks. In this case, the dispatching decisions need to be made in advance a long time before the handover operation between the truck and the corresponding QC happens. Thus, pre-planning is necessary for these activities. As the control function becomes improved, more decisions will be made based on the pre-planning function rather than on myopic decision rules.

Uniform Workload Principle: Avoid Congestions One major cause of low efficiency in a container yard is congestion of trucks or YCs. Even though the real-time operation may not follow the plan, such congestion may be anticipated if operation plans are analyzed carefully. Thus, when the plans are constructed, the workload should be distributed as uniformly as possible over the entire yard space and the planning horizon.

Pooling Principle: Share Resources if Possible Utilization and efficiency must be improved when multiple resources are shared by multiple users. However, the pooling must be supported by complicated operation rules. Thus, it is necessary to develop efficient operational rules for the pooling strategy can be applied to practices.

Postponement Principle: Commit a Decision at the Latest Possible Moment Situations change dynamically during real-time operation. Thus, schedules constructed based on the previous situation become unrealistic soon after the schedules start to be implemented. One popular strategy in logistics is postponing decisions until the latest possible moment in order to overcome the uncertainty in the operation and enable the system to respond quickly to the changing situation. For that purpose, real-time information collected from IT devices needs to be fully utilized

Synchronization Principle: Minimize Waiting Time by Synchronizing Movements of Different Equipment Containers are moved among vessels, yards, hinterland transportation centers, custom offices, and container freight stations and they are transferred from one type of equipment to another. These types of equipment must be synchronized during the handover operation to prevent one type from having to wait for the arrival of the other type. An efficient scheduling method needs to be developed to reduce the waiting time during the handover operation.

Minimum Empty Travel Principle: Minimize the Empty Travels of Equipment The travel distance is directly related to energy consumption and gas emission. For minimizing the travel distance, the layout of the yard needs to be improved and the allocation of tasks to equipment and the sequencing of tasks should be carefully determined. Both the empty travel distance and the loaded travel distance, which depends on the storage locations of containers, need to be reduced.

Flexibility Principle Decision rules should be flexible enough to accommodate the changes in throughput requirement, the changes in the layout, and the introduction of new types of equipment with a minimum modification. Even in these cases, their performance should be maintained at a high level for various situations. The software should be able to be applied to various terminals with different characteristics with minimal modifications.

Adaptability Principle: Easy to Adapt to Continuously Changing Situations Decision rules should be adaptable and capable of responding to changing situations. Considering that the situation may change dynamically and unexpected events may happen, more functions have been moved from planning functions to the function of real-time control.

2.4 Terminal Operating Systems

Many commercial products, called Terminal Operating Systems (TOS), have been developed and applied in practice. This section introduces some typical and popular products. TOS is composed of sub-systems for administration, planning, scheduling, executing and reporting parts. The administration part supports the management of container move orders from shipping lines. Generally, container move orders are transferred to the terminal through electronic data interchange (EDI) or internet access. This information is basic input data for the planning part.

The vessel calls are pre-defined by contracts with shipping lines and these are inputted into the berth planning module. The actual berthing time and position of vessel are scheduled by the berth planning module. The yard planning supports automatic stacking for import, export, and transshipment containers by determining an optimal yard position for a container. The resource planning allocates human resources (crane drivers, vehicle drivers, checkers, etc.) to various handling tasks in order to support the major activities in terminals. The ship planning and rail planning supply a crane split and work programs for unloading or loading containers. Tables 2.2, 2.3, and 2.4 summarize the various features of existing TOSs.

Table 2.2 Common features of the planning system in TOS

Module	Features
Berth planning	Editing calling schedules which come from contracts with shipping lines
	Assigning vessels to berths considering QC allocation
	Supporting berth allocation considering traffic flow of transporters and container yard positions
	Estimating berthing and departure time of each vessel
	Supporting ad-hoc vessel calls which are not included in the regular calling schedule
Yard planning	Defining automatic stacking rules for import, export, and transshipment containers
	Covering inbound containers from vessels and outbound containers from the gate and the rail
	Selecting storage slots of containers considering the efficiency during retrieval operations
	Considering workload distribution over yard areas during vessel loading process
	Forecasting future container inflow, outflow, and inventory for each vessel
	Supporting the space reservation for each vessel at each bay in each block
	Shared reservation of the same space for multiple vessels or multiple container groups
	Supporting the planning and operation of housekeeping of containers
	Visualizing the yard map showing stacks by container groups
Resource Planning	Registering personnel information—skill chart, job rotation, etc.
	Defining time units and calendar information—shift, day, week, and holidays, etc.
	Identifying the workload and available human resources during each time segment
	Allocating operators to shifts and gangs
Ship operation planning	Managing container stowage orders—bay profile, loading list, handling instructions, etc.
	QC split and work scheduling
	Slot sequencing for loading and unloading
	Automatic QC scheduling and slot sequencing
	Real-time rescheduling of QC works and re- sequencing slots to overcome disturbances
	Real-time stability calculations
	Managing vessel specific considerations—vessel stability calculation, stowage restrictions, twist lock handling, hatch covers handling, booming up/down, etc.
	Managing QC specific considerations−operation productivity of each QC, balancing QCs workloads, visualization of crane split, etc.

Table 2.2 (continued)

Module	Features
	Considering operations in the yard - yard workload balancing, avoiding unnecessary moves in blocks, minimizing travel dis-tance between the yard and vessels, etc.
	Considering special handling requirements—IMDG segregation rules, late arrival connections after cargo closing time, twin/tandem lifting, double cycling, etc.
Rail operation planning	Collecting container handling order information including the loading list from rail operation companies or shippers
	Rail crane split & rail crane work scheduling considering crane specifications
	Slot sequencing for loading and unloading
	Wagon composition for each ingoing/outgoing train considering wagon specifications
	Scheduling container transport between the yard of the port container terminal and the rail terminal
	Supporting direct loading of containers from road truck onto wagons or discharging from wagons onto road trucks
	Planning operations considering QC schedules in the port container terminal

During real-time operation, TOS constructs an optimal executing schedule for QCs, vehicles, and YCs to perform the various handling tasks on time. The real-time schedule is a short-term schedule which covers a period shorter than 30 min. TOS also schedules the handover times of containers between different pieces of equipment in order to minimize the waiting of equipment. When equipment becomes available to execute the next job or when a new job requests a schedule, a dispatching decision has to be made for matching the job with a set of resources required to perform the job. The storage locations for arriving containers from a vessel, the gate, or the rail terminal are determined by a yard positioning module which has a rule set. Furthermore, the equipment scheduling and dispatching modules should support various types of operations such as flexible loading, double cycling, and twin carrying. Various features of TOSs related to the real-time scheduling function are summarized in Table 2.3.

Another important group of functions of TOSs is controlling the real-time operations in the terminal. The gate system supports the carry-in/carry-out operations of outbound/inbound containers via road trucks. The TOS identifies a road truck driver, validates the cargo card, optionally inputs the pre-information if it is not received, and inspects a container, and issues a trip card. The truck appointment/pre-advice system receives a booking for carry-in/carry-out operations, which allows fast track checks of containers at the gate.

The TOS maintains job queues for each QC and checker, and jobs are dispatched from the TOS to a crane driver or a checker by using a voice and radio data terminal

Table 2.3 Common features of the real-time decision making in TOS

Real-time operation scheduling	Supporting hierarchical task decomposition of various operations. For example, a loading operation for a container may be decomposed into elementary tasks by a YC, a truck, and a QC
	Prioritizing various tasks for handling
	Real-time monitoring the progress of an operation for a container
	Real-time problem identification for re-scheduling
	Warning for the violation of time constraints by various operation schedules
	Real-time scheduling the yard operation: pre-positioning of containers, re-shuffling containers during idle times, and the prevention of deadlocks and collisions between YCs
	Real-time scheduling transport operations: pre-positioning containers, minimizing empty travel distances of transporters, synchronizing transport operations with operations by QCs and YCs
	Scheduling rail related operations considering departure times of trains
	Scheduling reefer container operations: scheduling YC operations, scheduling temperature checks, scheduling reefer plug connection/disconnection, and scheduling tasks for reefer operators
	Supporting such transport services as dual command cycling or twin carrying
Transporter dispatching	Pooling equipment among different groups of tasks classified by individual vessel, gate, or rail
	Pooling based on actual workload of cranes—mealtime, stoppage, and productivity of cranes
	User configurable priority settings for different groups of tasks
	Automatic generation of transport orders triggered by various events at the terminal
Yard positioning	Determining storage locations for unloading moves, carrying-in moves, and re-shuffling moves
	Decision making considering driving distances of cranes/vehicles and YC workloads
	Space allocation with the capability to scatter containers among multiple blocks or consolidate containers into a single block

(RDT). Crane drivers and checkers receive container handling jobs via RDT, execute jobs, and report results of jobs. When a container terminal uses automated cranes or vehicles, the TOS needs to support an event-driven messaging interface with the control system for the automated equipment. The TOS needs to send container handling orders to each piece of automated equipment, receive feedback about the progress of each order and relate it to the operational status of the corresponding pieces of the equipment. Table 2.4 summarizes the various functions of TOS related to the real-time operation and control.

Table 2.4 Common features of the real-time operation system in TOS

Gate	Identification of the truck driver
	Validation of the cargo card
	Input of information on carry-in or carry-out (pre-advise, pre-booking information)
	Handling documents related to customs
	Managing inspection information for containers
	Creation of temporary trip card indicating the destination in the yard
	Interface to truck appointment/pre-advice system
	Interface to auto gate system—OCR handling, barriers control, etc.
Quay crane (QC)	Container location control on the platform and QC stacks
	Reporting QC position and the status of the container being handled
	Sensing the stack profile
	Registering operation delays—input possible reasons of delays or stoppage codes
	Claiming the movement range of each crane for preventing conflicts between adjacent QCs
Vehicle or Yard Truck (YT)	Receiving a container transport order
	Reporting the progress of a container transport order—vehicle position & task progress status
	Prepositioning a vehicle to receive a container
Yard crane (YC)	Receiving a container handling order in a block
	Managing re-marshaling or re-shuffling operation
	Reporting a container handling order and crane position & status
Rail crane (RC)	Receiving a container handling order in a rail terminal
	Loading/unloading a container onto/from a train
	Reporting a container handling order, the position and status of a crane
	Registering the delay of operation by a stoppage code
	Claiming the range of a crane movement for preventing interference between two cranes
Container checker	Identification of ID, the size, and the type of a container
	Identifying the dimension of an Out-of-Gage container
	Identifying IMO code of a container
	Identifying physical characteristics, seals, damage condition, and door direction of a container
Reefer checker	Controlling the connection or the disconnection of the reefer plugs
	Checking the temperature inside a container periodically
Rail checker	Checking containers before unloading and after loading
	Controlling a wagon composition

To support the various functions of TOS, many commercial TOSs have been developed and used in practice. Table 2.5 introduces some representative products: Navis SPARCS N4, CATOS, Mainsail Vanguard, TOPS, and OPUS.

NAVIS is a company located in Oakland, USA and is the world's first provider of TOS. Their product "Navis SPARCS N4" has been implemented at around 200 container terminals in the world (NAVIS 2013). SPARCS N4 is treated by standard package software. Thus, based on customer needs, the functions of the software are regularly enhanced and the enhanced version is distributed and patched to customers through a version control. Although, it is expensive to modify the software in order to consider the individual local requirements of a specific customer, the system offers customers many options and adjustments, which may be used to adapt the standard system to the unique requirements of individual customers. The selection of options and the adjusting values of control parameters are also complicated tasks and so consulting companies may help the process of option selection and parameter adjustment.

SPARCS N4 includes AutoStow, Prime Route, Expert Decking, and a variety of user-selectable functions that have been used by many customers. SPARCS N4 SDK (System Development Kit) is a system which effectively supports the interface with the 3rd party provider's systems such as gate automation, private EDI, and local billing system. "SPARCS N4 Prime Route" provides a tool to pool prime movers across cranes, while combining yard and equipment constraints with operating business rules aimed at providing efficient work assignments in real-time, shorter travel distances, and fewer un-laden moves. "SPARCS N4 Expert Decking" is a tool for assigning each container to a storage position based on the business rules and constraints of the terminal, and is aimed at providing a high utilization of yard space, reduction in re-handles, and enhanced equipment utilization. "SPARCS N4 AutoStow" selects the next container to load in real-time by using rules obtained from combining stowage factors (e.g., type, weight) with yard constraints and operating strategy aimed at reducing planning time, increasing yard productivity, and raising responsiveness to operational challenges.

TOTAL SOFT BANK (TSB) is located in Busan, Korea and offers the CATOS (Computer Automation TOS) system that has been implemented at around 70 container terminals worldwide, mostly located in Asia (Total Soft Bank 2012). TSB has a marketing strategy of accommodating individual customer's needs as much as possible to satisfy each customer's local demands. Some functions of CATOS for a customer may not be directly applicable to other customers. Because CATOS has different features from a package software and additional development effort may be necessary for the application to a specific customer.

"CATOS Berth Planning" constructs and shows the berthing schedule by using powerful graphics. "CATOS Yard Planning" maximizes the yard stacking capacity while minimizing the planning time by supporting the popular planning process and rules of space planners in practice. "CATOS Ship Planning" supports simultaneous planning for multiple vessels by multiple planners, automatic load/discharge planning, and operation simulation. "CATOS Ship Planning" constructs multiple scenario-based ship plans, one of which is implemented considering the real-time

Table 2.5 Global TOS providers

Company	Name	NAVIS	Total Soft Bank	Tideworks	RBS
	Established	1988	1988	1999	1991
	Location	Oakland, USA	Busan, Korea	Seattle, USA	Parramatta, Australia
	Site	www.navis.com	www.tsb.co.kr	www.tideworks.com	www.rbs-tops.com
Customers	America	81	8	42	2
	Europe	42	7	5	5
	M.E & Af.	49	5	0	1
	Asia	45	47	3	13
	TOTAL	217	67	50	21
Product	Platform	J2EE	J2EE,.NET	Adobe flex	Unix/Linux, NET
	Name	Navis SPARCS N4	CATOS	Mainsail vanguard™	TOPS advance
	Administration	Navis SPARCS N4 platform EDI−EDI management	CATOS operation management CATOS Billing, Web-IP, EDI	Terminal billing Forecast®−customer website Data interchange	TOPS container, billing TOPS Web terminal view, EDI server
	Planning	Vessel—managing vessel activities and berth scheduling AutoStow−vessel and rail planning Yard−optimized control of yard space and container handling equipment Expert decking—yard management Rail—management of rail operations	CATOS –Vessel define – Berth Planning – Ship Planning –Yard planning – Rail planning	Spinnaker® –Vessel berthing –Vessel planning –Yard planning with yard navigator – Rail planning	TOPX Vessel Management—Vessel profile editor, berth planning and scheduling, vessel planning TOPX yard/truck management—Yard layout editor, yard allocation management, truck exception handling TOPX rail management—train schedule, rail planning

Table 2.5 (continued)

Company	Name	NAVIS	Total Soft Bank	Tideworks	RBS
	Executing	Monitor–monitor the performance of terminal operations Quay commander—crane scheduling and monitoring PrimeRoute—container handling equipment control Radio Data Terminal (RDT)—real-time managing jobs of all equipment Marine Telematics Solutions	CATOS Terminal Monitoring CATOS berth monitoring CATOS C3IT server Container Handling Equipment Supervisor—automatic job scheduling and CHE dispatching Supporting RDT CATOS SCADA–external systems integration	Terminal view^{TM} —3D visualization tool Traffi—Control^{TM} –Equipment Dispatching and monitoring Supporting RDT GateVision®—Gate Automation	TOPX equipment control—work queues and work instructions list, monitoring and control equipment by the type of work Supporting RDT Supporting the third party solution
	Advanced functions	ASC Manager–managing automated stacking cranes	ATC Supervisor–controlling and monitoring unmanned RMGC TSB Port Emulator–terminal operation evaluating with TOS		TOPS SimOne –Full 3D emulation of terminals
Company	Name	CyberLogitec	Yantai Huadong Soft-Tech	PSA	HIT
	Established	2000	1993	1964	1969
	Location	Seoul, Korea	Yantai, China	Singapore	Hong Kong
	Site	www.cyberlogitec.com	www.huadong.net	www.singaporepsa.com	www.hit.com.hk
Customers	America	2	–	–	–
	Europe	1	–	–	3
	M.E & Af.	2	–	–	1
	Asia	14	41	5	7
	TOTAL	19	41	5	11
Product	Platform	J2EE, RCP/Flex	Unix/Windows	Unix/Linux	J2EE

Table 2.5 (continued)

Company	Name	NAVIS	Total Soft Bank	Tideworks	RBS
	Name	OPUS Terminal	HD-CiTOS	CITOS	nGen
	Administration	OPUS terminal−management OPUS Terminal−Billing E-Service (Web Portal) OPUS EDI	Business Information subsystem Electronic data interchange subsystem	PortNet TradeNet	Tractor Appointment system, eBilling, EDI, customer plus
	Planning	OPUS terminal Vessel profile –Berth planning –Vessel planning –Yard planning –Rail planning	Intelligent planning subsystem –Berth Plan –Vessel operation plan –Vessel stowage –Yard plan –Train operation plan –Resource plan	Planning systems –Berth Planning & Monitoring –Yard Planning –Vessel planning –Resource planning –Engineering Management	Ship planning system (GUIDER) Yard automation
	Executing	OPUS Terminal –Terminal monitoring –Job scheduling & controller –Equipment pooling –Auto grounding Supporting RDT Providing automation packages (OCR, RFID, DGPS)	Basic operation subsystem –Vessel Dispatching –Train Management –Document Management –Customer Service –Yard Management –Quayside Security Management –Workload Handling	Operations systems –Ship Operation –Yard Operation –Yard Space Manager –Yard Consolidation –PM Tracking –Flow-Through Gate –Equipment PCs	Operations Monitoring System –Tall Pier-side Operations System –Radio Data System –WIFI Paging –Gate Automation, –Mobile Terminal Message System

Table 2.5 (continued)

Company	Name	NAVIS	Total Soft Bank	Tideworks	RBS
	Advanced functions	OPUS Terminal—job controller for automated YCs Eagle EyE—terminal asset and container monitoring & control	Intelligent Operation Subsystem –Wireless Data Terminal Transmission –Truck Auto-dispatching –Smart Gate Control –Quayside Operation Monitoring	Control Centre Systems –Vessel Operations, Monitoring & Control –Yard Operations, Monitoring & Control –Vessel Plan Analyzer	Next Generation Terminal Management System Computer Simulation

operation situation. The auto ship planning module supports various types of handling equipment such as transfer cranes and SCs and various operation types such as double cycling, truck pooling, and category loading. "CATOS C3IT Server" is responsible for decision making on resource allocation, locating containers, and problem alerting and solving in real time. "Container Handling Equipment Supervisor" is used to ensure on-time delivery of containers and reduce un-laden travel distance via container handling equipment (CHE) pooling, job scheduling and automatic CHE dispatching. "ATC Supervisor" controls job orders for unmanned yard equipment in real time and performs advanced automatic job-scheduling. "TSB Port Emulator" is used to simulate various operational scenarios built on various terminal operation parameters and historical operation data.

Mainsail VanguardTM is sold by Tideworks (2013), which is located in Seattle. Mainsail Vanguard$^{\text{TM}}$ has been implemented at around 50 container terminals worldwide, mostly North and South America. To overcome a poor EDI service environment of customers in some regions, Tideworks directly supports 24-h EDI services by reliable data processing through a data/operation center at the headquarters. Tideworks includes 3D visualization modules in Mainsail Vanguard$^{\text{TM}}$. Mainsail Vanguard$^{\text{TM}}$ provides functions such as real-time inventory management, flexible workflow tools, and instantaneous communication with customers and partners. "Active Inventory Control" carries out inventory management of containers, chassis, rolling stock, break-bulk, over-dimensional cargo, and hazardous materials. "Spinnaker Planning Management System®" integrates various planning tools in one workspace to increase cargo throughput capacity and reduce the vessel turnaround time. It includes the following modules: vessel planning module, yard planning module for automatic container location assignment, rail planning module, vessel workflow and scheduling tools for creating bay-by-bay work lists by shift and gang, and berth planning module. Traffic Control$^{\text{TM}}$ provides a dynamic control function for a terminal's container handling equipment and it replaces radio communication and paper instructions with accurate, real-time, electronic dispatching of work instructions to operators. Forecast® is a web portal that enables terminals to communicate more easily with shipping lines, trucking companies, brokers, and other parties.

"TOPS" is a product by REALTIME BUSINESS SOLUTIONS, which is located in Parramatta. "TOPS" has been implemented at around 21 container terminals worldwide (RBS 2013). "TOPS" provides the following various operational capabilities: yard management, vessel management, berth management, crane allocation, container handling equipment management, rail management, gate management, booking and pre-advice of containers, truck management and enquiry, user security and access control, and reports. "TOPS" supports twin lifting, dual cycling of QCs, double moves (inbound and outbound containers) by a truck without exiting the terminal, and automated housekeeping. "TOPS" is a UNIX-based system in which configuring the shared memory affords excellent data synchronization processing speed. Therefore, "TOPS" as a single system can smoothly handle all the transactions for a container terminal of over 10 million TEU. In addition, "TOPS" is based on X-windows which have advantages in graphical user interface. "TOPS" application is provided by two major components: the foundation system (TOPO) and the graphical planning and equipment control system (TOPX).

CyberLogitec (CLT) is a subsidiary company of Han-Jin Shipping Lines. Thus, the experience of the company in those container terminals has been well reflected in OPUS TerminalTM, which has been implemented at around 19 container terminals. OPUS TerminalTM is a recently developed system whose programming language is Java. The planning and operating modules in TOS are not dependent on the operating system (e.g., Windows, UNIX, etc.) (CyberLogitec 2013).

"OPUS Terminal Planning System" allows multiple users to be involved in the planning process by sharing the same part of the data base and it consists of the following three modules. "Berth Planning" covers the long-term schedule, the dedicate berth management, liner's private voyage number management, and berth chart. "Vessel Planning" provides a flexible planning tool for extraordinary circumstances, managing container handling orders, twin/tandem planning, dual cycling operation, multi user planning, evaluating ship plans, checking vessel stability, and handling late cargo arrivals after cargo closing time. "Yard Planning" estimates the workload in the yard in the near future, allocates yard space based on gate-in pattern, and changes stacking rules in accordance with current yard utilization ratio.

"OPUS TOS" allows users to monitor and control terminal operation such as vessel operation, terminal equipment workload or exceptional cases, transfer point congestion in quay, yard and gate site. It includes the following six functions. "Vessel operation" supports global pooling and partial pooling function for prime movers and twin/tandem operation. "Yard operation" offers the functions of balancing workload among yard equipments, minimizing equipment interference, and performing efficient re-marshalling operation based on the dynamic terminal situation. "Terminal job scheduling and controller module" creates job orders just in time based on operation plans. "Terminal Equipment Pooling" dispatches transporters in real time between the storage area to the quay side with the aim of maximizing the utilization of the transporters. "Auto grounding" allows users to dynamically manage yard operation and to change operation policies and yard stacking rules. "Auto housekeeping" searches candidate containers for housekeeping automatically and creates housekeeping orders.

Yantai Huadong Soft-Tech Company was founded in 1993 in China, whose product name is HD-CiTOS (Huadong Computer Intelligent Terminal Operation System) (Yantai Huadong Soft-Tech 2014)). It is applied to more than 40 container terminals which are located along the eastern coast and rivers in China and whose total throughput amounts to 7 million TEUs per year. Basic functions of the software include system initialization, base material maintenance, vessel dispatching, train dispatching, comprehensive inquires, etc. Intelligent planning subsystem is a core of CiTOS which consists of vessel handling plan, container stockpiling plan, train handling plan, various material plan, and so on. Decision support subsystem supports decision makers through historical data analysis.

PSA introduced business-to-business port logistics portal services (PortNet) in 1984 and a terminal operating system (CITOS, Computer Integrated Terminal Operations System) is launched in 1998 ((PSA 2014). CITOS is managing 52 berths and

188 quay cranes at 5 container terminals in Singapore. PortNET and CITOS both systems are integrated seamless to improve an efficiency of port logistics and container handling service. The PortNet is a web based portal service and supports many kinds of services: slot management, space booking (EZShip), global equipment management system (GEMS), electronic billing of charges (EZBill), cargo booking support (CargoD2D), throughput analysis, vessel information system (TRAVIS), and preplan container stowage on board the vessel (COPLANS). The planning systems includes berth planning & monitoring system (BPMS), yard planning systems (YPS), vessel planning systems (VPS), resource planning system (RPS) and engineering management systems (EAMS). And, operations systems includes ship operation system (SOS), yard operation system (YOS), yard space manager (YSM), yard consolidation system (YCS), PM tracking system (PMTS), flow-through gate system and equipment PCs (QCPC, YCPC, QCOPC, PMPC).

Hong Kong International Terminals introduced Next Generation Terminal Management System (nGen) in 2005, which adopted industry-standard and open-platform technologies such as Java and XML that make scalable across all non-proprietary computer hardware and operating system (Hong Kong International Terminals 2014). nGen is a modular system that offers a flexible architecture for plug-and-play options to sub systems. Operations monitoring system (OMS) visualizes terminal operations and container stacking information. Ship Planning System optimizes sequences of discharging and loading operations. Radio Data System (RDS) provides container movement's information to mobile computers. Yard Automation provides a variety of enquiry, reporting and analysis facilities to assist in the management of container inventory. Tractor Appointment System supports scheduling & collecting inbound containers. Mobile Terminal Message System delivers container handling information to user's mobile phones and Computer Simulation supports properly integrated and optimized operation plans before the deployment.

The four new challenges to TOS are automation, optimization by using IT, evaluation and analysis, and web and mobile. Automation is a global trend in container terminals. A control system for automated stacking cranes (ASCs) or automated RMGC (ARMGC) in cases of automated container terminals is generally now included in terminal operating systems. However, unmanned vehicle control systems (include AGV) have been provided by third party providers. A single terminal operating system, into which an unmanned vehicle control system fully is integrated, is expected to enter the market in the near future.

Optimization is another effective tool to improve the productivity in container terminals. An optimization technique could be effective through the support of real-time information technologies. Examples of the information technology applications are an equipment identification technologies using RFID/IoT (Radio-Frequency Identification, Internet of Things), improved reliability of wireless communication using mesh network, and sensor devices that can collect a variety of real-time information of equipment and work sites. By using the collected real-time information, decision making for job scheduling and equipment dispatching will become more realistic and effective.

Before TOS is deployed to real operations, it will need to be evaluated and tested. Because of its numerous operation parameters, the evaluation and testing of TOS will require a lot of money and time. Evaluation tools for this purpose have been developed from the mid 2000s and have been mainly used in some TOS implementation projects. Such evaluation tools can be widely used to support a process improvement after the operating system is installed.

The rapidly increasing demand for smart phones and tablets has boosted the cloud service market and altered the market leaders in ERP products; later it will incur the same changes in the market of TOS products. The next generation TOS is expected to incorporate some features of open architecture and standard web-based systems to support a variety of mobile devices.

2.5 Conclusions

This paper has reviewed various decision-making problems in container terminals. Potential research issues and directions were proposed for operation planning and real-time control activities. Extensive areas requiring further research were identified. The various functions offered by popular Terminal Operating Systems (TOSs) were introduced. In addition, the most popular TOSs in the present market were introduced, along with their key features. Finally, recent trends of TOSs responding to changes in the technological and market environment were highlighted.

Acknowledgement This research was a part of the project titled 'Technological Development of Low-carbon Automated Container Terminals', funded by the Ministry of Oceans and Fisheries, Korea."(201309550001)"

References

Ambrosino, D., Sciomachen, A., & Tanfani, E. (2006). A decomposition heuristics for the container ship stowage problem. *Journal of Heuristics, 12,* 211–233.

Angeloudis, P., & Bell, M. G. H. (2009). An uncertainty-aware AGV assignment algorithm for automated container terminals. *Transportation Research Part E, 46*(3), 354–366.

Bazzazi, M., Safaei, N., & Javadian, N. (2009). A genetic algorithm to solve the storage space allocation problem. *Computers & Industrial Engineering, 56,* 44–52.

Briskorn, D., Drexl, A., & Hartmann, S. (2006). Inventory-based dispatching of automated guided vehicles on container terminals. *OR Spectrum, 28,* 611–630.

Böse, J. W. (2011) Handbook of terminal planning. Operations research/computer science interface series (Vol. 49). New York: Springer.

Chen, L., Bostel, N., Dejax, P., Cai, J., & Xi, L. (2007). A tabu search algorithm for integrated scheduling problem of container handling systems in a maritime terminal. *European Journal of Operational Research, 181,* 40–58.

Choo, S., Klabjan, D., & Simchi-Levi, D. (2010). Multiship crane sequencing with yard congestion constraints. *Transportation Science, 44*(1), 98–115.

Crainic, T. G., & Kim, K. H. (2007) Intermodal transportation. In C. Barnhart & G. Laporte (Eds.), *Handbook in OR & MS* (Vol. 14, pp. 467–537). Amsterdam: Elsevier.

CyberLogitec (2013) OPUS Terminal. http://www.cyberlogitec.com/terminal. Accessed Oct 2013.

Dekker, R., Voogd, P., & van Asperen E. (2006). Advanced methods for container stacking. *OR Spectrum, 28,* 563–586.

Evers, J. M., & Koppers, S. A. (1996). Automated guided vehicle traffic control at a container terminal. *Transportation Research Part A, 30*(1), 21–34.

Goodchild, A. V. (2005) Crane double cycling in container ports: algorithms, evaluation, and planning, PhD dissertation, University of California, Berkeley.

Goodchild, A. V., & Daganzo, C. F. (2006). Double-cycling strategies for container ships and their effect on ship loading and unloading operations. *Transportation Science, 40*(4), 473–483.

Goodchild, A. V., & Daganzo, C. F. (2007). Crane double cycling in container ports: Planning methods and evaluation. *Transportation Research Part B, 41*(8), 875–891.

Grunow, M., Günther, H.-O., & Lehmann, M. (2004). Dispatching multi-load AGVs in highly automated seaport container terminals. *OR Spectrum, 26*(2), 211–236.

Hendriks, M., Laumanns, M., Lefeber, E., & Udding, J. T. (2010). Robust cyclic berth planning of container vessels. *OR Spectrum, 32,* 501–517.

Hong Kong International Terminals (HIT) (2014). nGen. http://www.hit.com.hk/en/Innovation. Accessed Jan 2014

Huang, Y., Liang, C., & Yang, Y. (2009). The optimum route problem by genetic algorithm for load/unloading of yard cranes. *Computers & Industrial Engineering, 56,* 993–1001.

Imai, A., Sun, X., Nishimura, E., Papadimitriou, S., & Sasaki, K. (2002). The containership loading problem. *International Journal of Maritime Economics, 4,* 126–148.

Imai, A., Sun, X., Nishimura, E., & Papadimitriou, S. (2005). Berth allocation in a container port: Using a continuous location space approach. *Transportation Research part B, 39,* 199–221.

Imai, A., Sasaki, K., Sun, X., Nishimura, E., & Papadimitriou, S. (2006). Multi-objective simultaneous stowage and load planning for a container ship with container rehandles in yard stacks. *European Journal of Operational Research, 171,* 373–389.

Jiang, X., Chew, E. P., Lee, L. H., & Tan, K. C. (2012a). Flexible space-sharing strategy for storage yard management in a transshipment hub port. *OR Spectrum, 35*(2), 417–439.

Jiang, X., Chew, E. P., Lee, L. H., & Tan, K. C. (2012b). A container yard storage strategy for improving land utilization and operation efficiency in a transshipment hub port. *European Journal of Operational Research, 221,* 64–73.

Jung, S. H., & Kim, K. H. (2006). Load scheduling for multiple quay cranes in port container terminals. *Journal of Intelligent Manufacturing, 17,* 479–492.

Kemme, N. (2011). RMG crane scheduling and stacking. In J. W. Bose (Ed.), *Handbook of terminal planning*. New York: Springer.

Kemme, N. (2012). Effects of storage block layout and automated yard cranes systems on the performance of seaport container terminals. *OR Spectrum, 34,* 563–591.

Kim, K. H. (2007). Decision-making problems for the operation of container terminals. *Journal of the Korean Institute of Industrial Engineers, 33*(3), 290–302.

Lau, H. Y. K., & Zhao, Y. (2008). Integrated scheduling of handling equipment at automated container terminals. *International Journal of Production Economics, 112*(2), 665–682.

Lee, B. K., & Kim, K. H. (2010). Optimizing the block size in container yards. *Transportation Research Part E, 46*(1), 120–135.

Lee, B. K., & Kim, K. H. (2013). Optimizing the yard layout in container terminals. *OR Spectrum, 35*(2), 363–398.

Lee, C. Y., & Yu, M. Z. (2012) Inbound container storage price competition between the container terminal and a remote container yard. *Flexible Services and Manufacturing Journal,* 24(3), 320–348.

Lee, L. H., Chew, E. P., Tan, K. C., & Han, Y. B. (2006). A yard storage strategy for minimizing traffic management in transshipment hubs. *OR Spectrum, 28,* 539–561.

Lee, D. H., Cao, Z., & Meng, Q. (2007). Scheduling of two-transtainer systems for loading outbound containers in port container terminals with simulated annealing algorithm. *International Journal of Production Economics, 107,* 115–124.

Lee, D. H., Wang, H. Q., & Miao, L. (2008). Quay crane scheduling with non-interference constraints in port container terminals. *Transportation Research Part E, 44,* 124–135.

Lee, D. H., Cao, J. X., & Chen, J. H. (2009). A heuristic algorithm for yard truck scheduling and storage allocation problems. *Transportation Research Part E, 45*(5), 810–820.

Lee, D. H., Chen, J. H., & Cao, J. X. (2010). The continuous berth allocation problem: A greedy randomized adaptive search solution. *Transportation Research Part E, 46*(6), 1017–1029.

Li, W., Petering, M. E. H., Goh, M., & de Souza R. (2009). Discrete time model and algorithms for container yard crane scheduling. *European Journal of Operational Research, 198*(1), 165–172.

Linn, R., & Zhang, C. (2003). A heuristic for dynamic yard crane deployment in a container terminal. *IIE Transactions, 35,* 161–174.

Liu D. K., & Kulatunga (2007). Simultaneous planning and scheduling for multi-autonomous vehicles. *Studies in Computational Intelligence, 49*, 437–464.

Meisel, F. (2009). *Seaside operations planning in container terminals*. Berlin: Physica-Verlag.

Moccia, L., Cordeau, J.-F., Gaudioso, M., & Laporte, G. (2006). A branch-and-cut algorithm for the quay crane scheduling problem in a container terminal. *Naval Research Logistics, 53,* 45–59.

Möhring, R. H., Köhler, E., Gawrilow, E., & Stenzel, B. (2004) Conflict-free real-time AGV routing. Preprint 026-2004, Technical University Berlin, Institute of Mathematics

Murty, K. G. (2007). Yard crane pools and optimum layouts for storage yards of container terminals. *Journal of Indutrial and Systems Engineering, 1*(3), 190–199.

NAVIS. (2013) Container terminal operation. http://www.navis.com/solutions/container. Accessed Sept 2013

Ng, W. C., & Mak, K. L. (2005). An effective heuristic for scheduling a yard crane to handle jobs with different ready times. *Engineering Optimization, 37*(8), 867–877.

Ng, W. C., Mak, K. L., & Zhang, Y. X. (2007). Scheduling trucks in container terminals using a genetic algorithm. *Engineering Optimization, 39*(1), 33–47.

Nguyen, V. D., & Kim, K. H. (2013). Heuristic algorithms for constructing transporter pools in container terminals. *IEEE Transactions on Intelligent Transportation Systems, 14*(2), 517–526.

Park, Y. M., & Kim, K. H. (2003). A scheduling method for berth and quay cranes. *OR Spectrum, 25*(1), 1–23.

Petering, M. H. (2009). Effect of block width and storage yard layout on marine container terminal performance. *Transportation Research Part E, 45*(4), 591–610.

PSA (2014) CITOS. https://www.singaporepsa.com/our-commitment/innovation. Accessed in Jan 2014

Realtime B. Solutions (R. B. S.). (2013) Terminal Operation Package System (TOPS). http://rbs-tops.com/product-details/. Accessed Sept 2013

Tideworks. (2013) Mainsail Vanguard^{TM} Marine terminal operating system. http://www.tideworks.com/products/mainsail/. Accessed Sept 2013.

Total Soft Bank. (2013) CATOS. http://www.tsb.co.kr/RBS/Fn/FreeForm/View.php?RBIdx=Ver1_38. Accessed Sept 2013.

Sammarra, M., Cordeau, J. F., Laporte, G., & Monaco, M. F. (2007). A tabu search heuristic for the quay crane scheduling problem. *Journal of Scheduling, 10,* 327–336.

Sauri, S., & Martin, E. (2011). Space allocating strategies for improving import yard performance at marine terminals. *Transportation Research Part E, 47,* 1038–1057.

Schwarze, S., Voss, S., Zhou, G., & Zhou, G. (2012). Scientometric analysis of container terminals and ports literature and interaction with publications on distribution networks. In H. Hu, et al. (Eds.), Computational logistics, volume 7555 of lecture notes in computer science (pp. 33–52). Berlin: Springer.

Sciomachen, A., & Tanfani, E. (2007). A 3D-BPP approaoch for optimizing stowage plans and terminal productivity. *European Journal of Operational Research, 183*(3), 1433–1446.

Sharif, Q., & Huynh, N. (2013). Storage space allocation at marine container terminals using ant-based control. *Expert Systems with Applications, 40,* 2323–2330.

Stahlbock, R., & Voss, S. (2008). Operations research at container terminals: A literature update. *OR Spectrum, 30,* 1–52.

Vis, I. F. A. (2006). Survey of research in the design and control of automated guided vehicle systems. *European Journal of Operational Research, 170*(3), 677–709.

Vis, I. F. A., & de Koster R. (2003). Trasshipment of containers at a container terminal: An overview. *European Journal of Operational Research, 147,* 1–16.

Vis, I. F. A., & Harika, I. (2004). Comparison of vehicle types at an automated container terminal. *OR Spectrum, 26,* 117–143.

Wan, Y., & Tsai, P. C. (2009). The assignment of storage locations to containers for a container stack. *Naval Research Logistics, 56,* 699–713.

Wang, Y., & Kim, K. H. (2011). A quay crane scheduling algorithm considering the workload of yard cranes in a container yard. *Journal of Intelligent Manufacturing, 22,* 459–470.

Won, S. H., Zhang, X., & Kim, K. H. (2012). Workload-based yard-planning system in container terminals. *Journal of Intelligent Manufacturing, 23*(6), 2193–2206.

Woo, Y. J., & Kim, K. H. (2011). Estimating the space requirement for outbound container inventories in port container terminals. *International Journal of Production Economics, 133,* 293–301.

Wu, Y., Luo, J., Zhang, D., & Dong, M. (2013). An integrated programming model for storage management and vehicle scheduling at container terminals. *Research in Transportation Economics, 42,* 13–27.

Yang, C. H., Choi, Y. S., & Ha, T. Y. (2004). Simulation-based performance evaluation of transport vehicles at automated container terminals. *OR Spectrum, 26*(2), 149–170.

Yantai Huadong Soft-Tech. (2014) HD-CiTOS. http://www.huadong.net/UploadFile/datum/CiTOS.pdf. Accessed in Jan 2014.

Yuan, S., Skinner, B. T., Huang, S., Liu, D. K., Dissanayake, G., Lau, H., & Pagac, D. (2011). A job grouping approach for planning container transfers at automated seaport container terminals. *Advanced Engineering Informatics, 25,* 413–426.

Zhang, H. P., & Kim, K. H. (2009). Maximizing the number of dual-cycle operations of quay cranes in container terminals. *Computers & Industrial Engineering, 56*(3), 979–992.

Zhang, C., Wan, Y., Liu, J., & Linn, R. (2002). Dynamic crane deployment in container storage yards. *Transportation Research, B 36,* 537–555.

Chapter 3
Container Port Competition in Europe

Theo E. Notteboom and Peter W. de Langen

Abstract Port competition has become a complex and multi-faceted concept due to changes in the market environment of ports and the resulting intensification of rivalry between operators in the same port, between neighbouring ports, between multi-port gateway regions and between entire port ranges. This chapter discusses port competition in Europe with a main focus on container ports and terminals. It provides an in-depth theoretical and empirical description of port competition in the second most important container port system in the world after Asia. The chapter aims to provide the reader with a clear insight on the current status, drivers and issues in European container port competition.

3.1 Introduction

The globalization of production and consumption, the emergence of a global transport network together with changes in inter-port relations, port-hinterland relationships and logistics have created greater competition among ports. Shippers, logistics service providers and shipping lines do not necessarily choose a port, but they select a chain in which a port is merely a node. In order to respond to the requirements of trade and international supply chains, ports need to accommodate and handle more and larger ships and hinterland transport modes faster. Therefore, construction, expansion, planning and management of ports are increasingly complex and costly. These trends and the expansion of the role of the private sector in port activities have forced ports to become more market-oriented, more innovative and more responsive to the needs of all actors involved in the trades which pass through the port. Seaport managing bodies have to play an active role in the marketing sense (in encouraging

---

T. E. Notteboom (✉)
Transportation Management College - Dalian Maritime University, China and ITMMA - University of Antwerp, Belgium and Antwerp Maritime Academy, Belgium,
e-mail: theo.notteboom@gmail.com

P. W. de Langen
Port and Logistics Advisory (PLA), Eindhoven University of Technology,
Eindhoven, The Netherlands
e-mail: P.W.d.Langen@tue.nl

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_3

ships and cargo to use the port), in contributing to efficient supply chains and in enabling trade and regional development.

Against this background, port competition has become a complex and multi-faceted concept. The nature and characteristics of competition depend among others upon the type of the competing ports (e.g. feeder port, hub port) and the cargo group (e.g. containers, liquid bulk, dry bulk, non-containerized general cargo). For example, the excellent maritime accessibility of the port of Rotterdam for ULCCs (Ultra Large Crude Carriers) has rendered this port a very strong position in North-Western Europe for the handling of crude oil. In container transport a distinction has to be made between the large load centres or main/hub ports and the smaller regional or feeder ports. The load centres are primarily competing for deep sea intercontinental liner services, in which large ships of up to 18,000 TEU are being deployed. Regional ports are striving for connections to as many nearby load centres as possible and for having a good regional hinterland connectivity.

Heaver (1995) points out that terminals are the major focus of competitive strategy, not ports. Along these lines, we can define port competition as competition for trades, with terminals as the competing units, logistics, transport and/or industrial enterprises as the chain managers of the respective trades and port authorities and port policy makers as co-developers of the broadly defined port complex. Competition between ports has increasingly been replaced by competition among market players who often are present in more than one port (cf. global terminal operators such as PSA, DP World, Hutchison Port Holdings and APM Terminals, see Notteboom and Rodrigue (2012) for a detailed analysis) or multimodal logistics and transport service providers who, in addition to operating various transport modes, have also combined stevedoring, storage, forwarding and other activities in one ‘bundle’ for shippers. Port competition can also involve rivalry among port authorities in view of offering the best facilities (both material and non-material) to all actors involved in the supply chains of the various trades (e.g. stevedoring companies, shipping companies, shippers and multimodal operators).

This chapter discusses port competition in Europe with a main focus on container ports and terminals. This chapter provides an in depth theoretical and empirical description of port competition, and does not review or present economic models of port competition (see De Borger et al 2008 and Luo et al 2012, for two papers that take such an approach).

After Asia, Europe features the second most important container port system in the world in throughput terms. Close to 50 European container ports have regular intercontinental services to the rest of the world, while another 80–90 (smaller) container ports play a more regional intra-European role. The aim of this chapter is to provide the reader with a clear insight on the current status, drivers and issues in European container port competition.

This chapter is structured as follows. In a second section we further develop the concept of port competition by looking at the different geographical and functional levels port competition can take place. Then, we analyze recent dynamics for all cargoes handled in the European port system. The fourth section focuses on container ports. A last section brings the discussion on container port competition to a more strategic level and presents conclusions.

3.2 Different Geographical and Functional Levels of Port Competition in Europe

The complex nature of seaport competition manifests itself in the number of levels that can be distinguished in relation to competition (Verhoeff 1981).

A first level concerns intra-port competition which has been extensively discussed in De Langen and Pallis (2006). Private port companies in a certain port, often compete, for handling cargoes (terminals) and for providing other port services (e.g. towage, bunkering). For the port authority or the port itself as a whole, such competition can serve as a management method to improve the efficiency of port activities. Competition between operators or providers of facilities within the same port can generally increase port efficiency and improve services. However, exceptionally the need to realize economies of scale (and reach the so-called Minimum Efficient Scale or MES, see Kaselimi et al. 2011), to offer modern technologies and the existence of enough competition from operators in other ports may justify an operational monopoly of port activities in one port (World Bank 1992). The ‘playing field’ for intra-port competition is very often influenced by basic infrastructural investment decisions by port or regional authorities. For example, in many larger ports not all container terminals have the same draft conditions thereby reducing the choice for very large container vessels. So even if there are a handful of terminals in a large port area, the largest container vessels might be able to call at one terminal only.

The second level is that of the multi-port gateway region, a term coined by Notteboom (2009; 2010) and later also applied by e.g. Feng and Notteboom (2013) and Liu et al (2013). The locational relationship to nearby identical traffic hinterlands is one of the criteria that can be used to group adjacent seaports. Also the port calling patterns in the liner service networks of shipping and hinterland connectivity profile can help to group ports to a multi-port gateway port (Notteboom 2009). When the ports concerned have a separate management, the neighboring gateways are vying for the same hinterland cargo flows. Later in the chapter, we will deploy the concept of ‘multi-port gateway regions’ as one way of looking at container port competition in Europe. Many stevedoring companies are expanding their activities over more than one port of such a port region. The increasing ‘footloose’ character of shipping companies, pushes port authorities and port companies into fierce competition. On port authority level, the battle is mainly focused on offering the best basic infrastructural (docks, quays) and ‘infostructural’ (IT) facilities, the best logistic/distribution facilities and the lowest port user costs. On the terminal level, the competition mainly focuses on price, handling time and productivity. When ports in the same gateway region do not fall under the same national government, as is for example the case for the Rhine-Scheldt delta ports in the Low Countries, government policies can have an impact on the conditions and level of competition among sub-groups (e.g. Dutch ports against Flemish ports).

Regional authorities often aim to secure complementary product-market developments in neighbouring (rival) ports of the same port region (see Notteboom et al. 2009, for an edited book on issues regarding ‘ports in proximity’). As such regional authorities try to provide a framework in which each seaport can operate i.e. manage its port-specific advantages, while at the same time encouraging cooperation.

The third level of competition is the port range which can be defined as a group of ports situated at the same seashore and sharing more or less the same hinterland. Within port ranges one can generally observe fierce intra-range competition. Within the Hamburg-Le Havre range, the most important port range in Europe, the initiatives in the field of port cooperation taken so far, are primarily based on exchange of information (aimed at improving mutual understanding) instead of real structural cooperation. The vigorous intra-range competition in the Hamburg-Le Havre range is reinforced by the fact that the ports are spread over different countries (Belgium, Germany, the Netherlands and France), each following their own port policy. Several mainports situated in different countries are competing for the status of nodal point of nodal points within a European transport network. The Netherlands for example is sparing no effort in trying to consolidate the position of Rotterdam as a European gateway i.e. mainport. The same is true for Belgian ports like Antwerp and Zeebrugge. On the other hand, the ports of countries such as Germany, France and Italy, whose industries generate considerable import-export trade, try to channel goods transport as much as possible to their own port infrastructure. For example, the ports of Hamburg and Bremen are striving to keep the competitive edge over Rotterdam and Antwerp, particularly as regards the goods produced by the industries situated in the Ruhr area in the western part of Germany.

The fourth and last level of port competition involves the rivalry between port ranges. As will be demonstrated later, the gradual completion of one European transport network and the increasing hub-feeder port relations has intensified inter-range competition in Europe, e.g. between the Hamburg-Le Havre range and the Mediterranean ports. Intra-range competition requires a common approach to port development, as different policies distort the ‘playing field’ and thus lead to inefficient freight flow patterns. However, such a common approach is often complicated even when supra-national authorities, like the European Commission, are involved. Cooperation between ports in different ranges exists, for example, the port authorities of Zeebrugge (Belgium) and Götenborg (Sweden) and a terminal operator together formed ‘Gothenbrugge Ltd.’, a company focused on the coordination of investments in roro-facilities and the realization of fast transport connections between both ports.

In the next sections we will analyse port competition in Europe at all levels, except the intra-port competition level.

3.3 Total Throughput in the European Port System

With a total throughput of an estimated 3.79 billion t in 2012, the European port system ranks among the busiest port systems in the world. Growth was particularly strong in the pre-crisis period between 2000 and 2008, partly driven by fast growing container throughput, i.e. an average annual growth rate of 10.5 % in the period 2005–2008 and 7.7 % in the period 2000–2005. The economic crisis which started

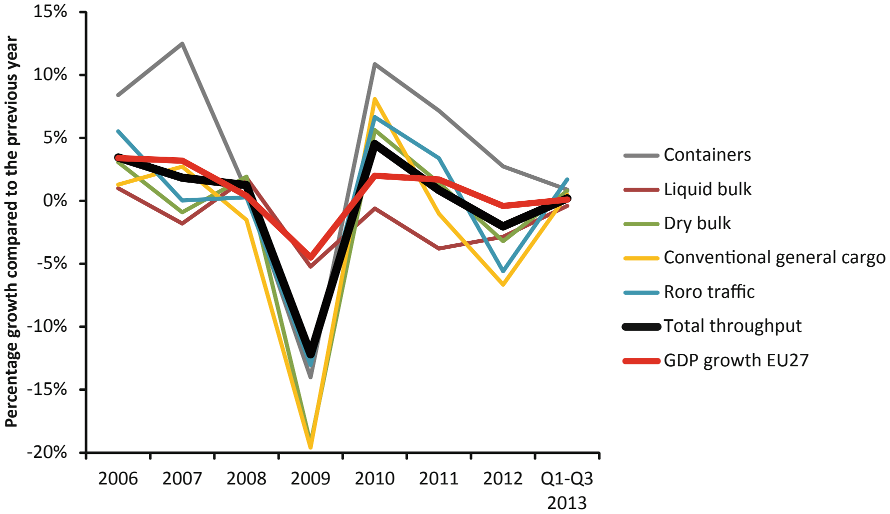


Fig. 3.1 Year-on-year growth in total EU port traffic (basis = ton) and EU27 GDP (Note: growth figures for Q1-Q3 2013 are estimates based on a sample of about 60 European ports included in the Rapid Exchange System database)

to have its full effect in late 2008 made an end to the volume growth in the European seaport system. Total cargo volumes handled by European ports decreased 12.2 % in 2009, from 4.18 billion t in 2008 to 3.67 billion t in 2009. The throughput figures bounced back in 2010 to 3.84 billion t (+ 4.5 % compared to 2009), but more recent years did not bring further throughput recovery to pre-crisis levels (Fig. 3.1). In 2011 growth was merely 0.8 % and in 2012 the European port system recorded a mild drop of 2 % in cargo handlings. The first three quarters of 2013 brought a very modest growth of only 0.2 % compared to the figures of the first nine months of 2012. Important to note is that the European port system is still not back at pre-crisis cargo volumes. Total cargo throughput in European ports in 2012 was still 10 % below the 2008 volumes. 2013 points to a very small change compared to 2012. Next to dry bulk and conventional general cargo, liquid bulk flows seem to face a hard time to turn the tide. Only container traffic in European ports has managed to rise above the 2008 level (by a modest 6 %).

Figure 3.1 also provides more detail on the traffic evolution for five cargo groups: liquid bulk (mainly oil and oil products), dry bulk (major bulks such as iron ore, coal and grain, but also minor bulks such as minerals and fertilizers), containers, roll-on/roll-off cargo and conventional general cargo (steel, forest products, heavy lift, etc.). The latter two cargo groups were initially affected the most by the crisis with a volume drop of nearly 20 % in 2009. The recovery in 2010 was too weak to undo the 2009 effect. The year 2012 brought volume losses, after a stagnation in 2011. Container traffic was also heavily affected in 2009, but since 2010 the European container port system shows some growth again, be it at a much lower rate than

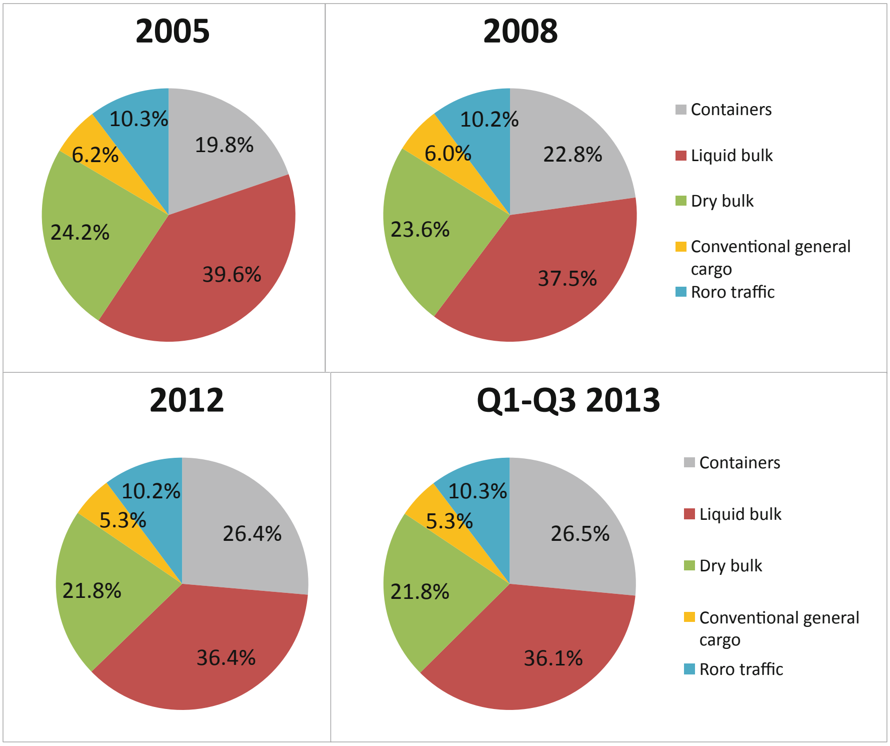


Fig. 3.2 Distribution of cargo flows in the EU port system

before. Liquid bulk volumes initially recorded a rather modest decline in 2009, but growth figures have remained negative since. The first three quarters suggest that the year 2013 will be a year of overall stabilization with almost zero growth in all cargo groups. The growth pattern per individual port can look very different from the overall pattern. For example, in the first nine months of 2013 liquid bulk volumes in the European port system decreased by an estimated 0.4 % (based in the RES sample of ports) with quite diverging growth figures for some of the main liquid bulk ports. Rotterdam recorded a small decline of 2.4 %, Antwerp showed a massive increase of 32 % (mainly due to recent large scale investments in tank storage facilities) and Le Havre remained fairly stable at + 0.8 %. In the same period, Nantes-St-Nazaire saw a 13.3 % drop in volumes and Marseille of – 11.2 %, while Sines grew by 20 % and Bilbao by 14.7 %.

The differences between the growth paths of the respective cargo groups changed the cargo type distribution in the European port system (Fig. 3.2). Liquid bulk still accounts for the largest share, but its relative importance has dropped from about 40 % in 2005 to 36.4 % in 2012. The share of container traffic continues to grow.

A comparison of the year-on-year growth figures in the European port system with the GDP growth figures for the EU27 reveals that ports overreact to swings in

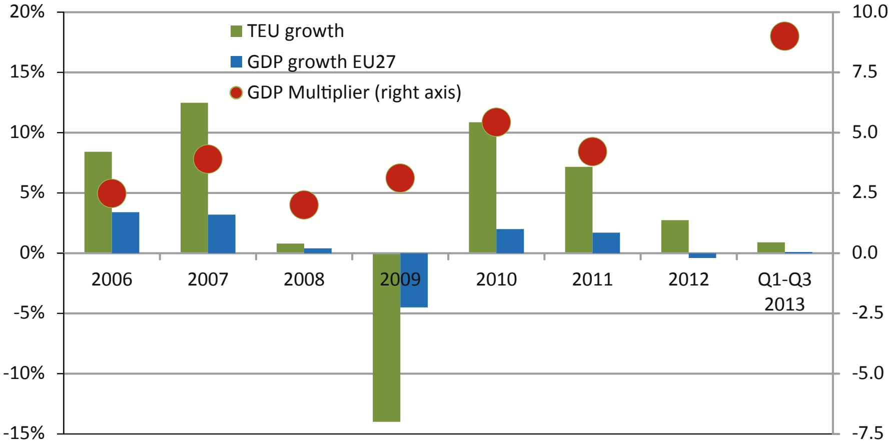


Fig. 3.3 The GDP multiplier in the EU container port system

economic growth. When the economy booms, seaports typically show high to very high growth figures. However, an economic crisis has a very pronounced negative effect on cargo volumes in seaports. The year 2013 seems to be a year of stabilization with almost zero growth in both GDP figures and cargo throughput. Figure 3.3 shows the evolution in the European GDP multiplier, i.e. here the ratio between world TEU growth and world GDP growth. The results point to a complex relationship between container port traffic and economic growth in Europe. This phenomenon is further illustrated in Fig. 3.4 which shows the container growth in a number of container ports around Europe in the period 2008–2012. The highest growers can be found all over Europe, including countries such as Greece, Portugal, Spain and Italy which have been severely affected by the government debt crisis. The weakest performers in terms of growth are also found all over Europe, including in countries with the best economic status in the Eurozone (such as Germany). In other words, seaports in countries with the weakest economies of Europe do not necessarily underperform compared to seaports in stronger countries. The main reason underlying this observation is that quite a few ports rely heavily on container flows which are not related to the immediate hinterland, but on flows that are distantly generated.

3.4 Dynamics in the European Container Port System

This section discusses recent developments in the European container port system. We are particularly interested in the impact of the crisis on the port hierarchy in Europe. Are new container ports and port regions emerging as challengers of established ports and regions? Are some port regions in Europe gradually losing their significance? How is the balance between north and south evolving?

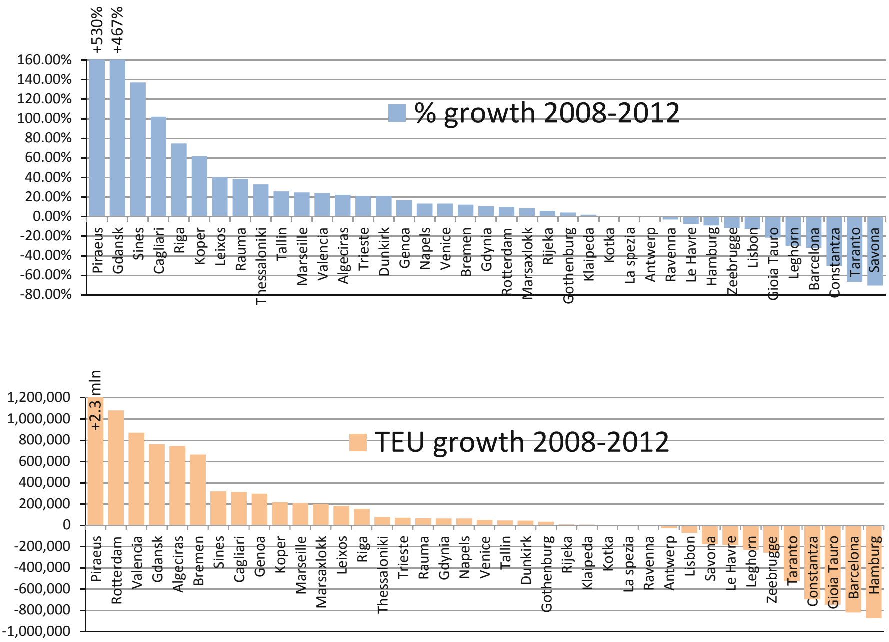


Fig. 3.4 Strong growth differences between individual ports—TEU traffic

3.4.1 Total Port System

With a total maritime container throughput of an estimated 95.2 million TEU in 2012, the European container port system ranks as the second busiest container port system in the world. Growth has been particularly strong in the period 2005–2007 with an average annual growth rate of 10.5 %, compared to 6.8 % in the period 1985–1995, 8.9 % in 1995–2000 and 7.7 % in 2000–2005. The economic crisis made an end to the steep growth curve. Total container throughput increased from 90.7 million TEU in 2008 to 95.2 million TEU in 2012 or an average annual growth of 'only' 1.26 %. The year 2009 is at the root of this slow pace given a y-o-y drop in container volumes of about 14 % in 2009. Between 2009 and 2012 traffic volumes have recovered at a rate of 6.87 % per year. An overall growth of 0.9 % in TEU was realized in the first nine months of 2013. For 2014 most sources predict a revival of container volumes in Europe. For example, the 'North Europe Global Port Tracker' of Hacket Associates and the Institute of Shipping Economics and Logistics (ISL) in Bremen expects a growth for North-Europe in incoming container traffic of 16 %. For the entire European port system import growth would reach 9 %. At the export side, the forecasted growth in North Europe amounts to 11 % (mainly driven by Asia and North America).

Table 3.1 provides an overview of the fifteen largest container ports in the European Union. Saint-Petersburg, which handled 2.52 million TEU in 2012 and has

Table 3.1 The top 15 European container ports (1985–2013, in 1000 TEU). (Source: Updated from Notteboom (2010) based on statistics individual port authorities)

in 1000 TEU

R		1985		1995		2005		2008		2009		2012	2013	%	Note 2013	*R*
1	Rotterdam	2655	Rotterdam	4787	Rotterdam	9287	Rotterdam	10784	Rotterdam	9743	Rotterdam	11866	11621	-2.1%	final	*1*
2	Antwerp	1243	Hamburg	2890	Hamburg	8088	Hamburg	9737	Antwerp	7310	Hamburg	8864	9254	4.4%	final	*2*
3	Hamburg	1159	Antwerp	2329	Antwerp	6488	Antwerp	8664	Hamburg	7008	Antwerp	8635	8578	-0.7%	final	*3*
4	Bremen	986	Felixstowe	1924	Bremen	3736	Bremen	5448	Bremen	4565	Bremen	6115	5859	-4.2%	estimate	*4*
5	Felixstowe	726	Bremen	1518	Gioia Tauro	3161	Valencia	3597	Valencia	3654	Valencia	4470	4328	-3.2%	final	*5*
6	Le Havre	566	Algeciras	1155	Algeciras	2937	Gioia Tauro	3468	Algeciras	3043	Algeciras	4071	4338	6.6%	final	*6*
7	Marseille	488	Le Havre	970	Felixstowe	2700	Algeciras	3324	Felixstowe (*)	3021	Felixstowe (*)	*3700*	-			*7*
8	Leghorn	475	La spezia	965	Le Havre	2287	Felixstowe (*)	3200	Gioia Tauro	2857	Piraeus	2734	3163	15.7%	final	*8*
9	Tilbury	387	Barcelona	689	Valencia	2100	Barcelona	2569	Marsaxlokk	2330	Gioia Tauro	2721	-			*9*
10	Barcelona	353	Southampton	683	Barcelona	2096	Le Havre	2502	Zeebrugge	2328	Marsaxlokk	2540	-			*10*
11	Algeciras	351	Valencia	672	Genoa	1625	Marsaxlokk	2337	Le Havre	2234	Le Havre	2304	2463	6.9%	growth 9m	*11*
12	Genoa	324	Genoa	615	Piraeus	1450	Zeebrugge	2210	Barcelona	1801	Genoa	2065	1999	-3.2%	growth 9m	*12*
13	Valencia	305	Piraeus	600	Marsaxlokk	1408	Genoa	1767	Southampton (*)	1600	Zeebrugge	1953	2000	2.4%	estimate	*13*
14	Zeebrugge	218	Zeebrugge	528	Southampton	1395	Southampton (*)	1710	Genoa	1534	Barcelona	1750	1719	-1.8%	final	*14*
15	Southhampton	214	Marsaxlokk	515	Zeebrugge	1309	Constanza	1380	La spezia	1046	Southamption (*)	*1600*	-			*15*
	TOP 15	10450	TOP 15	20841	TOP 15	50067	TOP 15	62697	TOP 15	54072	TOP 15	65388				
	TOTAL Europe	17172	TOTAL Europe	33280	TOTAL Europe	73729	TOTAL Europe	90710	TOTAL Europe	78011	TOTAL Europe (est.)	95220				
	Share R'dam	*15.5%*	*Share R'dam*	*14.4%*	*Share R'dam*	*12.6%*	*Share R'dam*	*11.9%*	*Share R'dam*	*12.5%*	*Share R'dam*	*12.5%*				
	Share top 3	*29.4%*	*Share top 3*	*30.1%*	*Share top 3*	*32.4%*	*Share top 3*	*32.2%*	*Share top 3*	*30.8%*	*Share top 3*	*30.8%*				
	Share top 10	*52.6%*	*Share top 10*	*53.8%*	*Share top 10*	*58.2%*	*Share top 10*	*58.8%*	*Share top 10*	*58.8%*	*Share top 10*	*58.5%*				
	Share top 15	*60.9%*	*Share top 15*	*62.6%*	*Share top 15*	*67.9%*	*Share top 15*	*69.1%*	*Share top 15*	*69.3%*	*Share top 15*	*68.7%*				

(*) Estimate

witnessed strong growth in the past few years is not included in the ranking (as Russia is not an EU member). A number of the listed ports act as almost pure transhipment hubs with a transhipment incidence of 75 % or more (i.e. Gioia Tauro, Marsaxlokk, Algeciras) while other load centres can be considered as almost pure gateways (e.g. Genoa and Barcelona to name but a few) or a combination of a dominant gateway function with sea-sea transhipment activities (e.g. Hamburg, Rotterdam, Le Havre, Antwerp).

About 68 % of the total container throughput in the European port system passes through the top fifteen ports, compared to 61 % in 1985. Since 2008 no major shifts have taken place in the traffic shares of the top 3, top 10 and top 15 ports, although the top 3 ports have lost some ground. Nearly one third of all containers are handled by the top three ports. Worth mentioning is that the dominance of market leader Rotterdam weakened in the late 1990s, but in the past decade the port's position has remained quite stable. Overall, the figures suggest a continued high concentration of cargo in only a dozen large container ports. While the crisis has not significantly altered the rankings, a number of ports lost some positions while others gained. For example, the Belgian port of Zeebrugge initially overcame the crisis very well by climbing to the ninth position in 2010 but afterwards lost traffic and now is in position 13. The Greek port of Piraeus showed the most volatile traffic evolution. Piraeus' volume peaked at 1.6 million TEU in 2003, but strikes and unrest led to a throughput of only 433,000 TEU in 2008. In 2010, the container port started a remarkable recovery path partly pushed by the arrival of Cosco Pacific as operator of the Pier 2 facility. Piraeus reappeared in the top 15 ranking in 2011 and held position 8 in 2012 with a total volume of 2.7 million TEU. In 2013 COSCO Pacific has announced to further expand. Under the terms of the agreement, COSCO will spend 230 million € to increase Piraeus's cargo handling capacity by two thirds over the next seven years to an annual capacity of 6.2 million TEU.

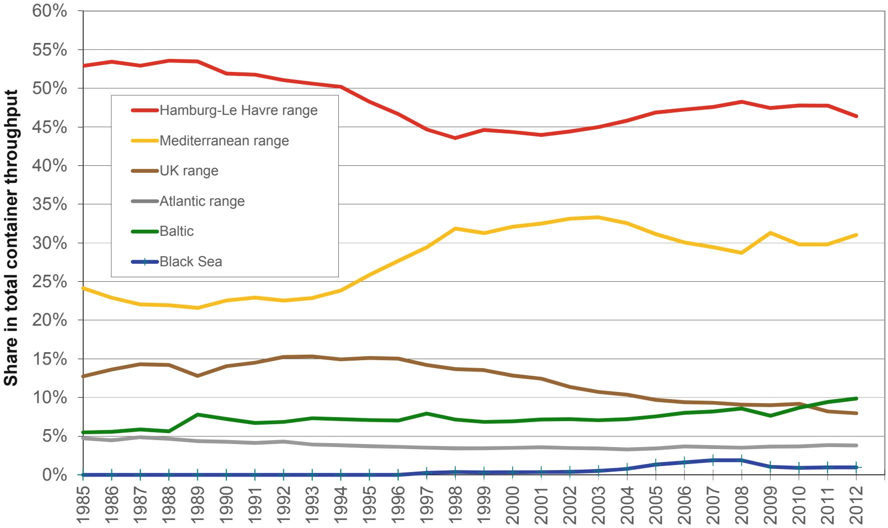


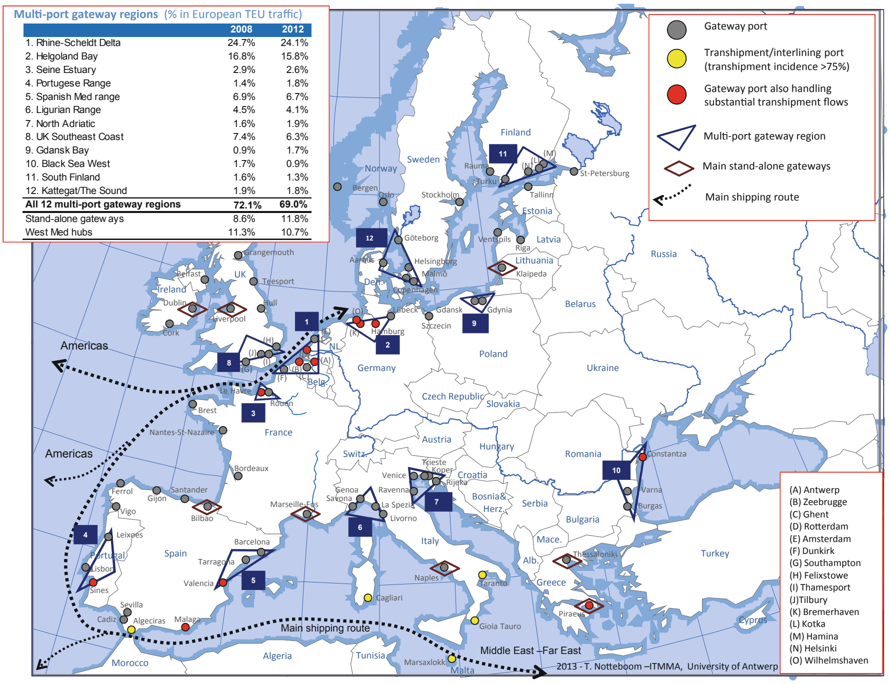
Fig. 3.5 Traffic shares of port ranges in the European container port system (source: updated from Notteboom 2010)

3.4.2 *Port Ranges*

At port range level, the container ports in the Hamburg-Le Havre range (which includes all ports along the coastline between Le Havre in France and Hamburg in Germany) handle about half of the total European container throughput (Fig. 3.5). The share of the Mediterranean ports grew significantly between the late 1980s and the late 1990s at the expense of the ports in the Hamburg-Le Havre range. The significant improvement of the share of the Med was mainly the result of the insertion of transhipment hubs in the region since the mid-1990s (Gioia Tauro, Marsaxlokk, Cagliari, Taranto to new but a few). At the start of the new millennium, the position of the northern range gradually improved while the Med ports and the UK port system lost ground. The crisis seems to have stopped this trend as from 2009 the traffic balance between the Med and the Hamburg-Le Havre range remained quite stable. However, the position of the UK ports (Southeast and South coast only) continued to weaken. The Baltic port region has clearly strengthened its traffic position in the past few years. The strong growth path of European ports in the Black Sea area (Romania and Bulgaria) suddenly stopped in crisis year 2009.

3.4.3 *Multi-Port Gateway Regions*

When we group seaports within the same gateway region together to form so-called multi-port gateway regions some interesting intra- and inter-regional dynamics can be



Multi-port gateway regions (% in European TEU traffic)

	2008	2012
1. Rhine-Scheldt Delta	24.7%	24.1%
2. Helgoland Bay	16.8%	15.8%
3. Seine Estuary	2.9%	2.6%
4. Portugese Range	1.4%	1.8%
5. Spanish Med range	6.9%	6.7%
6. Ligurian Range	4.5%	4.1%
7. North Adriatic	1.6%	1.9%
8. UK Southeast Coast	7.4%	6.3%
9. Gdansk Bay	0.9%	1.7%
10. Black Sea West	1.7%	0.9%
11. South Finland	1.6%	1.3%
12. Kattegat/The Sound	1.9%	1.8%
All 12 multi-port gateway regions	**72.1%**	**69.0%**
Stand-alone gatew ays	8.6%	11.8%
West Med hubs	11.3%	10.7%

Fig. 3.6 Multi-port gateway regions in the European container port system. (Source: updated from Notteboom 2010)

unveiled. Figure 3.6 provides an overview of the main multi-port gateway regions in Europe as well as transhipment hubs and stand-alone gateways. Stand-alone gateways are somewhat isolated in the broader port system, as they have less strong functional interactions with adjacent ports than ports of the same multi-port gateway region. In the next sections we will draw some conclusions based on the changing positions of the port regions between 2008 and 2012 and some preliminary figures for 2013.

3.4.3.1 The Rhine-Scheldt Delta: The Largest European Container Port Region

The Rhine-Scheldt Delta and the Helgoland Bay ports, both part of the Le Havre-Hamburg range, together represent some 40 % of the total European container throughput in 2012. The market share of the Rhine-Scheldt Delta shows moderate fluctuations since 2008 with 24.7 % in 2008, 25.5 % in 2009, 26 % in 2010, 25 % in 2011 and 24.1 % in 2012. In 2013 the Rhine-Scheldt Delta saw a TEU decline of 0.9 % (Rotterdam: – 2.1 % and Antwerp: – 0.7 %). The year 2014 promises to be a key year to the ports given new capacity coming on stream (e.g. Maasvlakte 2) and the full impact of the schedules announced by the P3 Network (MSC, Maersk

Line and CMA CGM), the G6 Alliance and other shipping lines and groups. The Rhine-Scheldt Delta port region has one of the largest terminal capacity reserves in Europe. The massive Deurganck dock in the port of Antwerp, which opened in 2005, provides ample room for traffic growth. The PSA terminal and the Antwerp Gateway Terminal at the dock together handled less than 2 million TEU in 2012 while the design capacity of the dock amounts to some 9 million TEU. Recently a decision was taken to move MSC's volumes (some 4.5 million TEU per year) from the Delwaide dock on the right bank to the Deurganckdock and to concentrate all future P3 Network traffic on the left bank. A deepening program of the river Scheldt was completed a few years ago in view of guaranteeing access to the largest container vessels such as the triple 'E' class within an acceptable tidal window. The current Maasvlakte 2 developments in Rotterdam include the construction of two large scale container facilities, each with a capacity of between 4 and 5 million TEU: a terminal for APM terminals and the Rotterdam World Gateway which will be operated by a consortium led by DP World. The first phases of both terminals will come on stream in 2014. ECT, part of Hong Kong based Hutchison Port Holdings, has room for further capacity growth by extending the current 1.5 km quay of its Euromax terminal. The terminal capacity in Zeebrugge includes PSA's new and still heavy underutilised Zeebrugge International Port (ZIP) facility and spare capacity at the APM Terminals facility in the outer harbour. The strong hinterland ambitions of the Rhine-Scheldt Delta ports are supported by a range of hinterland concepts and products such as a strong orientation on barge transport, rail shuttles into the distant hinterland, ECT's European Gateway Services network (see Veenstra et al, 2012 and Rodrigue and Notteboom 2009 for more details) and similar efforts by terminal operators DP World and PSA, and a dense network of inland terminals and European distribution zones in or in the vicinity of the ports. To secure growth in the future, the ports are actively targeting transshipment markets in the Baltic, the UK and southern Europe and hinterland areas in southern Germany, Italy, South France (cf. Lyon area) and Eastern and Central Europe, next to a continued focus on their cargo rich core service areas (the Benelux, western Germany and northern France).

3.4.3.2 German Ports back on Their Feet After a Dramatic 2009

The North-German ports in the Helgoland bay gained traffic share in Europe from 13 % in the late 1990s to 16.8 % in 2008. Bremerhaven's volume surge and Hamburg's pivotal role in feeder flows to the Baltic and rail-based flows to the developing economies in East and Central Europe were the main causes. However, sharp volume drops in 2009, i.e. minus 28 % in Hamburg mainly due to a loss of transshipment flows to Rotterdam and minus 16 % in Bremerhaven, brought the traffic share below 15 %. In the past three years their position recovered to 15.8 %. In 2013 Hamburg recorded a healthy growth of 4.4 % (mainly attributed to regaining part of the Baltic T/S market back from Rotterdam) while Bremerhaven witnessed a volume drop of 4.2 %. The deepening of the Elbe river is high on the agenda in Hamburg as the port is currently facing some restrictions to accommodate the largest container vessels.

The region welcomed newcomer Wilhelmshaven in 2012 when the JadeWeserPort was opened for business. With a volume of about 76,265 TEU in 2013, the new large scale terminal facility clearly has to make its mark. Short-term prospects to attract new business have improved by the announcement that the P3 Network would include the port on two of its services on the Europe-Far East trade. Newcomer Wilhelmshaven is actively pursuing transshipment business, given that it can attract volumes more quickly than gateway traffic, which is more difficult to attract, mainly because of the time required to develop intermodal services. Note that rail services have been established primarily using in-house rail/intermodal firms, and prices to/from Wilhelmshaven and inland points have been matched with those to/from Hamburg and Bremerhaven to the same inland destinations.

3.4.3.3 'Renaissance' of the Seine Estuary

The Seine Estuary, the third region in the Le Havre-Hamburg range, suffered from a gradual decline in its market share from 5.5 % in 1989 to 2.9 % in 2008. The 'Port 2000' terminals in Le Havre, a new hinterland strategy, the completed port reform process and the HAROPA initiative aimed at closer cooperation between Le Havre, Rouen and the inland port of Paris should support a 'renaissance' of Le Havre. These initiatives did not have their full effect in 2012 as the region's share in European container traffic declined further to 2.6 %. However, the year 2013 reversed this trend with an impressive growth of 6.9 % in the first nine months of 2013. Several shipping lines (such as MSC) and shippers have committed new volumes to this port area. The port also hopes to benefit from the P3 alliance.

3.4.3.4 The Portuguese Port System Aims for hub Status

Portuguese ports Lisbon, Leixoes and Sines are trying very hard to expand business by developing a transhipment role as well as tapping into the Spanish market (particularly the Madrid area) through rail corridor formation and dry port development. After a long period of declining market shares, the Portuguese port system succeeded to lift its European share to 1.8 % in 2012. The port of Sines recorded the strongest traffic growth mainly due to increasing volume commitments of MSC and a further extension of the PSA/MSC operated terminal facility. Sines more than doubled throughput since 2008 to reach 553,063 TEU in 2012. In the first nine months of 2013 traffic grew by a staggering 76.5 %, thereby surpassing the two other ports which each have a cargo base of around 500,000 to 600,000 TEU.

3.4.3.5 Spanish Med Ports show a Diverging Growth path

Among the major winners before the crisis, we find the Spanish Med ports with a growth of the European share from 4 % in 1993 to 6.9 % in 2008. While the share remained rather stable the past few years, the growth path of the individual ports is quite different. Barcelona was hit hard by the crisis with a volume drop from 2.57 million TEU in 2008 to 1.8 million TEU in 2009. Container activities (particularly sea-sea transshipment) did not recover after 2009, mainly due to lower feeder volumes and the Catalan port closed 2013 at 1.72 million TEU. Barcelona continues to aim for better connectivity with central European hinterlands (see Van den Berg and De Langen 2011). At the other extreme, Valencia recorded a spectacular and consistent growth (also during 2009) from 3.6 million TEU in 2008 to 4.47 million TEU in 2012. However, the 2013 throughput saw a decline of 3.2 %. MSC's choice to use the port as a hub for the region boosted transshipment volumes and consolidated the port's fifth position in the European ranking. While Tarragona remains a smaller player in the region, the port saw strong growth in 2008 when DP World and ZIM Lines took over the Contarsa terminal. Since then throughput amounts to some 200,000 to 250,000 TEU.

3.4.3.6 Ligurian Ports Challenged to Outgrow the Italian Hinterland

The Ligurian ports have difficulties in keeping up with other regions in Europe. The ports jointly represent some 4.5 % of the total European port volume, a decline compared to 6–7 % throughout the 1980s and 1990s. In the first nine months of 2013 Genoa recorded a traffic drop of 3.2 % while La Spezia saw a growth of 2.6 %. The Ligurian ports rely heavily on the cargo rich economic centres in northern Italy. While they also aim at attracting business from the Alpine region, the southeast of France and southern Germany, success in these areas has been limited so far partly because of intense competition from northern ports supported by a strong multimodal offer in terms of rail and barge shuttles.

3.4.3.7 North Adriatic to Become a Southern Gateway to Europe?

Just like the Ligurian ports, the North-Adriatic ports have been facing lower than average growth rates. However, since the crisis year 2009 the tide seems to have turned. The cooperation agreement NAPA (North Adriatic Ports Association) underlines the ambition of the region to develop a gateway function to Eastern and Central Europe and the Alpine region. The strategy should also enable the region to develop larger scale container operations. The NAPA ports are determined to lure trade from northern ports via upgraded rail links and shorter transit times from Asia. For example, Trieste has a harbor that's 18 m deep and able to handle the largest container ships at full load. The Italian port offers shuttle train services to destinations in Germany, Austria, Hungary, Slovakia and the Czech Republic, and is targeting countries as distant as Poland, one of the main markets for Hamburg. Still the Adriatic ports are facing scale differences with the northern hub ports which affect the possibilities to develop a vast intermodal hinterland network. With 'only' 1.8 million TEU in 2012 the Adriatic ports only handle a fraction of the volumes of the two leading multi-port gateway regions of the Hamburg-Le Havre range (i.e. 22.9 million TEU in the Rhine-Scheldt Delta and 15.1 million TEU in northern Germany).

3.4.3.8 The Direct call vs. Feeder Challenge in Ports of the UK Southeast Coast

The UK ports witnessed a rather significant decrease in market share. Many of the load centres along the southeast coast of the United Kingdom faced capacity shortages in the early 2000s while new capacity became available only gradually. Quite a number of shipping lines opted for the transhipment of UK flows in mainland European ports (mainly Rhine-Scheldt Delta and Le Havre) instead of calling at UK ports directly. With the prospect of new capacity getting on stream there is hope for more direct calls and potentially an increase in market share.

Since mid-2013 the combination of bigger ships, larger alliances and the new London Gateway terminal are affecting the UK container port system. Thamesport has lost virtually all deep sea services partly because of draft restrictions in the River Medway approach channel. Evergreen moved its UK cargo from Thamesport to Felixstowe while other lines such as Hapag-Lloyd, OOCL and NYK moved their transatlantic services from Thamesport to Southampton. The volume drop in Thamesport started already earlier with ‘only’ 300,000 TEU handled in 2012, compared to close to 800,000 TEU in 2008. Also Tilbury’s traffic is likely to be affected negatively by larger ships sizes and the opening of DP World’s London Gateway terminal. Both Thamesport and Tilbury, as well as other smaller container ports such as Great Yarmouth, will likely focus more on niche and short sea intra-European services.

The new London Gateway terminal complex will face competition from UK ports Felixstowe and Southampton, but also from mainland European ports such as Rotterdam, Zeebrugge, Antwerp and Le Havre which offer competitive feeder services to the UK. The large scale London Gateway terminal of DP World can be regarded as the embodiment of the UK ambitions to attract more direct calls. The terminal was developed on an old Shell site along the Thames. The port will add 3.5 million TEU to the UK’s port capacity and will help to meet the demand for extra capacity in the UK. The full impact of London Gateway on competitive dynamics between mainland European ports and UK ports will become clear in the coming years. It remains to be seen how DP World is going to balance its many stakes in large scale terminals across the region: the company is investing heavily in the Rotterdam World Gateway facility on Maasvlakte 2 and has a vested interest in filling the Antwerp Gateway terminal. London Gateway received its first vessel in November 2013. The terminal can accommodate vessels with a draft of up to 17 m at any state of the tide. Maersk, MOL and Deutsche Afrika Linien already decided to shift their UK port of call on the South Africa service from Tilbury to London Gateway. Rail links are already in place connecting the terminal with the big centres, with DB Schenker Rail UK taking a lead role in the provision of those services. In June 2013, Marks & Spencer confirmed to invest in a new distribution centre within the terminal area to open in 2016.

3.4.3.9 The Gdansk bay: Attracting Direct Deep sea Calls in the Baltic

In the last couple of years, the ports in the Bay of Gdansk are witnessing a healthy growth and an increasing traffic share in Europe (now 1.7 % compared to 0.9 % in 2008 and 0.5 % in 2004). For a long time, the Polish load centres were bound by their feeder port status, competing with main port Hamburg for the Polish hinterland. However, in the last decade the Polish port reform process gave impetus to the development of new container handling facilities. While Gdynia has benefited from volume gains, Gdansk attracted most attention as volumes increased from 163,704 TEU in 2008 to 928,905 TEU in 2012. Growth remained very strong in the first nine months of 2013, i.e. 30.2 % more volume compared to the same period in 2012. The DCT facility in Gdansk serves as a port of call on one of the main Europe-Far East services of Maersk Line. Emma Maersk class vessels with a capacity of 15,500 TEU not only bring Asian cargo, but also pick up North American container flows via other European ports of call before heading to Gdansk. Since August 2013 the 18,000 TEU Triple E vessels of Maersk Line call at DCT Gdansk in Poland.

The Gdansk case provides empirical evidence that deep sea calls in the Baltic can be viable despite the existence of competitive hub-feeder networks linked to Hamburg and other major northern ports. The port is determined to become a hub for Central and Eastern Europe and Russia. With a throughput of well over 1 million TEU in 2013 (note that St-Petersburg remains the largest container port in the Baltic with 2.52 TEU handled in 2012), the port has ambitious plans to ultimately expand the terminal's annual capacity to around 4 million TEU by 2016. The port is even challenging the established notion of 'Hamburg-Le Havre range' by proposing the notion of 'Gdansk-Le Havre range'.

3.4.3.10 The Rise and Fall of European Black sea Ports?

The Black Sea ports, Constantza in particular, were on the rise in the early 2000s from virtually no traffic to a European share of 1.7 % in 2008. Constantza attracted terminal investments given its potential to serve as a gateway to Eastern Europe and a transshipment hub for the Black Sea area. The crisis abruptly ended this unfolding success story: Constantza's container throughput fell sharply from 1.38 million TEU in 2008 to 594,299 TEU in 2009. In the following years the port could only present a modest growth to reach 684,059 TEU in 2012, still far from the record of 1.4 million TEU in 2007. Early on in its development, Constanta was very much seen as the transshipment gateway for the Black Sea and reached a transshipment incidence of some 75 % in 2008. However, times have changed quite significantly as traffic patterns in the region have evolved. When the crisis hit many container lines changed their liner services in search of cost-efficient logistic solutions. A number of direct services from the Far East into the Black Sea region were cancelled, negatively affecting transshipment volumes. As a result, in 2012 almost three-quarters of the volumes handled at the port consisted of local import and export containers, with the remaining quarter being transshipment. Still, Constanza handles the largest vessels operated in the Black Sea (some 8,000 TEU). Terminal productivity plays an important role in the future development of container terminals in the Black Sea region, where operators in both Ukraine and Russia such as Odessa and Novorossiysk are trying to attract both transshipment and import/export business. The Bulgarian ports of Varna and Burgas remain small players in the container market. The traffic decline in Black sea ports is in sharp contrast to strong growth witnessed by Piraeus and Turkish deep sea ports near the Sea of Marmara. This development demonstrates shipping lines for the time being prefer a hub-feeder model in the Med to service the Black Sea area instead of direct deep sea calls in the Black Sea.

3.4.3.11 Scandinavian Ports

The ports at the entrance of the Baltic and South Finland show a moderate growth path, both losing some ground in a European context. However, the relative decline in their European shares is smaller than in the five years prior to the start of the economic crisis. Scandinavian ports remain highly dynamic players in the market and are European pioneers in far-reaching port cooperation schemes. The ports of Malmo in Sweden and Copenhagen in Denmark were merged in 2001 to form a single company, Copenhagen Malmö Port. It still serves as a successful case in cross-border mergers of two ports. In 2011, the City Councils in Kotka and Hamina on Finland's south coast approved a port merger. The port of Gothenburg in Sweden serves as a good practice in intermodal network development: half of the port's container volume is transported inland via an extensive domestic rail network of container shuttles. The rail network also extends to Norway.

Some of the ports in this region are gearing up to welcome more direct calls of mainline vessels. This is particularly felt in ports like Gothenburg and Aarhus which are already acting as regular ports of call on quite a few intercontinental liner services. While these ports have a good position to act as turntables for the Baltic on many trade routes, the insertion of these ports as regular ports of call on the Europe-Far East trade remains uncertain. The large vessel sizes deployed on this route, the associated reduction in the number of ports of call and the additional diversion distance make regular direct calls to the multi-port gateway region Kattegat/The Sound less viable compared to other trade routes. The P3 Network, the alliance between Maersk Line, MSC and CMA CGM, plans to include Gdansk and Aarhus in its rotation for the BALTIC service with ships of 14,000 TEU while Gothenburg will act as a port of call in the SKAW service with ships of 13,000 TEU.

3.5 Discussion and Conclusion

Container port competition is becoming ever more complex and intense, not only between ports of the same range or multi-port gateway region, but also between ranges and multi-port gateway regions. The current logistics and economic environment makes European container ports increasingly compete not as individual places that handle ships but within transport chains or supply chains. The logistics chain has become more than ever the relevant scope for analyzing port competitiveness. If a seaport wants to attract or retain some of the megacarriers (be it shipping lines, logistics service providers or shippers) it has to position itself as an efficient intermodal hub and logistics service center acting within extensive transport and communications networks.

European integration has created a single European market with economic centres in East and Central Europe, the Nordic triangle and the Iberian Peninsula increasingly making their mark on the European economic scene, thereby giving rise to new load centres and inland transport corridors. While a large part of the throughput of European container ports remains locally generated and stimulated by the ports' centrality with respect to a strong regional hinterland, ports can no longer expect to attract cargo simply because they are natural gateways to rich hinterlands. Seaports are competing fiercely to extend their hinterlands across frontiers (see De Langen 2007, for the case of Austria). The increased competition decreases 'natural' gateways and captive hinterlands. This tendency is further enhanced by the development of intermodal corridors and inland terminals in Europe. By developing strong functional links with particular inland terminals a port might intrude in the natural hinterland of competing ports. 'Islands' in the distant hinterland are created in which the load centre achieves a comparative cost and service advantage vis-à-vis rival seaports (Notteboom and Rodrigue 2005).

At the maritime side, new liner service network configurations and larger ships force ports into head-on-head competition. In the past 20 years the center of gravity of economic development moved to Asia. The Europe-Far East trade gradually overtook the Atlantic to become the most important intercontinental most trade. This geographical shift gave Med ports a new impulse to play a very active part in the deep sea container business. The ever larger vessels push ports to stretch their nautical accessibility profile, their infrastructure and intermodal offer.

The consolidation process in the container handling industry also has a large impact on individual ports. Large terminal operators are becoming more footloose as the network approach loosens their former strong ties with one particular seaport. At the same time, competition has shifted from port authorities to private terminal operators who are establishing terminal networks. The global stevedoring companies have acquired a very strategic position in a port's future. In the present port competition model, ports are frequently pushed into making investment decisions of a speculative nature. Many European container ports make significant investments without any degree of assurance that traffic will increase and shipping lines will retain their loyalty. Their only belief is that a lack of investments will certainly not increase traffic.

In the analysis of port competition, a supply chain perspective is required. Especially important is the spatial distribution of supply chains. These have a huge impact on container volumes in ports. As a 'rule of the thumb' a large distribution center of about 50,000 m^2 'floor space' may generate 3000–5000 TEU per year. So inland nodes with large volumes of distribution centers (Duisburg, Venlo and Meerhout, to mention arguably the three largest ones in Germany, Belgium and The Netherlands) account for large container volumes (Venlo and Meerhout both $>$ 300,000 TEU, Duisburg $>$ 600,000 TEU). In this sense, the spatial competition for distribution centers is central for understanding port competition (see Ferrari et al 2006).

All these changes in the port environment have an impact on the port hierarchy in the European container port system. First, seaports located far away from each other are now to some extent competing. Ports in the Hamburg-Le Havre range are competing with UK ports, especially for UK-bound transshipment traffic (see Ng and Yu 2006). Competition is also growing between the Mediterranean port system and the Hamburg-Le Havre range, as these two different port systems are in a good position to reach the economic and industrial heartland of Europe.

Secondly, the position of some large load centers is challenged by medium-sized ports and new hub terminals. In Southern Europe new hubs have emerged since the mid-1990s (e.g. Algeciras, Marsaxlokk, Gioia Tauro, Taranto and Cagliari). The term 'west med hubs' in Fig. 3.6 refers to these 'pure' transshipment hubs in the Mediterranean. They all have a transshipment incidence of above 80 %. The success of these ports is partly the result of the fact that a call involves a minimal diversion for a mainline vessel transiting the Mediterranean between the Suez Canal and the Straits of Gibraltar. A lot of carriers are using these Med hubs to shift boxes between linehaul services in order to serve more markets with fewer vessels. However, ports whose competitive strategy is completely based on their intermediacy may find themselves in an unstable and highly fragile position, as this kind of traffic flow is more volatile and footloose and depends solely upon the strategy of shipping lines with respect to their service networks. In Northern Europe, successful upstream ports such as Antwerp and Hamburg as well as existing large coastal ports such as Rotterdam and Bremerhaven are facing competition from new terminal initiatives. Good examples are the JadeWeserPort project in Wilhelmshaven, the port of Gdansk, London Gateway and the development of a container terminal in Amsterdam—that so far failed to attract customers. The new terminal facilities might give shipping lines and alliances more opportunities to use their bargaining power to play off one port against another. Still, as it takes more than cranes and quay walls to become successful, some doubt whether the new terminals will be able to become effective competitors of the existing large ports which are also investing on a continuous basis to strengthen their position even further. However, the fact that some entrants in the container business have not been successful (e.g. the terminal in Amsterdam), does not imply that all new terminal projects in non-hub ports have few chances of becoming successful.

Competition between European container ports focuses mainly on their capacity to attract the maximum container volume in order to justify direct calls. Some of the key factors to success include the proximity to a strong cargo generating and cargo receiving hinterland, a favourable location both nautically as vis-à-vis the hinterland, a strong sea and land connectivity (both in infrastructures but also frequency, quality and price of transport services), a high port and terminal efficiency, the right pricing and a supply chain approach.

Given the fact that the European port throughput in 2013 is likely to remain below the 2008 levels, whereas new capacity has been added (through new ports as Londen Gateway and JadeWeserPort as well as capacity expansion in established ports, as most visible expansion project the Second Maasvlakte in Rotterdam, but investments in expansion in various other large ports) and productivity has increased (partly due to increasing ship sizes and investments in new cranes and other equipment) competition between ports as well as terminals has intensified. There is no longer a market in which all players can record growth—the gain of one port (terminal) is increasingly the loss of another port (terminal).

Especially for port authorities this leads to increasing pressure and ongoing debates about the need to reform port governance structures, as port authorities in Europe are government owned except for the UK, see De Langen (2004) for a theoretical discussion on port governance, Brooks and Cullinane (2006) for a large number of descriptive cases of port governance and Baird (2013) for a description of the changing ownership structures of the private UK ports.

Furthermore, the increasingly complex playing field for port authorities has led to debates about the need for port authorities to move beyond their traditional 'government owned landlord' role and develop new commercial capabilities as well as capabilities to effectively develop and execute investment programs with benefits for the port community at large, such as port community systems, inter-terminal transport, coordination and planning of hinterland services and regulated road access to the port, see Verhoeven (2010) and Van der Lugt et al (2013).

References

Baird, A. J. (2013). Acquisition of UK ports by private equity funds. *Research in Transportation Business & Management, 8,* 158–165.

Brooks, M. R., & Cullinane, K. (Eds.). (2006). *Devolution, port governance and port performance. Research in Transportation Economics*, (Vol. 17).

De Borger, B., Proost, S., & Van Dender, K. (2008). Private port pricing and public investment in port and hinterland capacity. *Journal of Transport Economics and Policy (JTEP), 42*(3), 527–561.

De Langen, P. (2004). Governance in seaport clusters. *Maritime Economics & Logistics, 6*(2), 141–156.

De Langen, P. W. (2007). Port competition and selection in contestable hinterlands; the case of Austria. *European Journal of Transport and Infrastructure Research, 7*(1), 1–14.

De Langen, P. W., & Pallis, A. A. (2006). Analysis of the benefits intra-port competition. *International Journal of Transport Economics, 33*(1), 69.

Feng, L., & Notteboom, T. (2013). Peripheral challenge by Small and Medium Sized Ports (SMPs) in multi-port gateway regions: The case study of northeast of China. *Polish Maritime Research, 20,* 55–66.

Ferrari, C., Parola, F., & Morchio, E. (2006). Southern European ports and the spatial distribution of EDCs. *Maritime Economics & Logistics, 8*(1), 60–81.

Heaver, T. D. (1995). The implications of increased competition among ports for port policy and management. *Maritime Policy and Management, 22*(2), 125–133.

Kaselimi, E. N., Notteboom, T. E., Pallis, A. A., & Farrell, S. (2011). Minimum Efficient Scale (MES) and preferred scale of container terminals. *Research in Transportation Economics, 32*(1), 71–80.

Liu, L., Wang, K. Y., & Yip, T. L. (2013). Development of a container port system in pearl river delta: Path to multi-gateway ports. *Journal of Transport Geography, 28*, 30–38.

Luo, M., Liu, L., & Gao, F. (2012). Post-entry container port capacity expansion. *Transportation Research Part B: Methodological, 46*(1), 120–138.

Ng, A., & Yu, K. (2006). Assessing the attractiveness of ports in the North European container transhipment market: An agenda for future research in port competition. *Maritime Economics & Logistics, 8*(3), 234–250.

Notteboom, T. (2009). *Economic analysis of the European seaport system. Report serving as input for the discussion on the TEN-T policy*. Report prepared for European Sea Ports Organisation (ESPO). Antwerp: University of Antwerp.

Notteboom, T. E. (2010). Concentration and the formation of multi-port gateway regions in the European container port system: An update. *Journal of Transport Geography, 18*(4), 567–583.

Notteboom, T. E., & Rodrigue, J. P. (2005). Port regionalization: Towards a new phase in port development. *Maritime Policy & Management, 32*(3), 297–313.

Notteboom, T., & Rodrigue, J. P. (2012). The corporate geography of global container terminal operators. *Maritime Policy & Management, 39*(3), 249–279.

Notteboom, T., Ducruet, C., & de Langen, P. W. (Eds.). (2009). *Ports in proximity: Competition and coordination among adjacent seaports*. Ashgate, Surrey.

Rodrigue, J. P. & Notteboom, T. (2009). The terminalization of supply chains: Reassessing the role of terminals in port/hinterland logistical relationships. *Maritime Policy & Management, 36*(2), 165–183.

Van den Berg, R., & De Langen, P. W. (2011). Hinterland strategies of port authorities: A case study of the port of Barcelona. *Research in Transportation Economics, 33*(1), 6–14.

van der Lugt, L., Dooms, M., & Parola, F. (2013). Strategy making by hybrid organizations: The case of the port authority. *Research in Transportation Business & Management, 8,* 103–113.

Veenstra, A., Zuidwijk, R., & van Asperen, E. (2012). The extended gate concept for container terminals: Expanding the notion of dry ports. *Maritime Economics & Logistics, 14*(1), 14–32.

Verhoeff, J. M. (1981). Seaport competition: some fundamental and political aspects. *Maritime Policy & Management, 8*(1), 49–60.

Verhoeven, P. (2010). A review of port authority functions: Towards a renaissance? *Maritime Policy & Management, 37*(3), 247–270.

World Bank. (1992). *Port marketing and the challenge of the third generation port*. Geneva: World Bank.

Chapter 4
Container Port Competition and Competitiveness Analysis: Asian Major Ports

Paul Tae-Woo Lee and Jasmine Siu Lee Lam

Abstract This chapter evaluates the competitive edge of major Asian container ports, i.e. Busan, Hong Kong, Shanghai, and Singapore ports, referring to the customer-centric community ports, so-called the Fifth Generation Ports, of which the concept has been modified for evaluation. The literature of competitiveness of container ports applied a number of methods, among others, including time series analysis, DEA and SFA methods, service quality analysis with importance-performance analysis and Kano model, multi-criteria evaluation, survey of container ship operators and logistics managers, shift-share analysis and diversification indexes such as Herfindahl–Hirschmann, marginal cost pricing approach, and game theory. The previous literature measured the relative competitiveness of ports in Asia and Europe. Such analysis results are useful to evaluate the competitiveness of container ports at the given evaluation angle at the given time. However, they do not consider port competitiveness in relation to port devolution according to a globalized economy with changes of production and distribution channels, technology, city-port interface, government policy, port users' behavior, pricing, environmental issue, as well as security and safety. Having considered the above limitation in the literature, this chapter argues that a novel approach is required to evaluate inter-port competition in a comprehensive way to reflect cross-sectional, longitudinal and horizontal aspects of the port evolution. In this regard, it can be said that this novel approach, i.e. the Fifth Generation Ports, with empirical test by descriptive and quantitative methods contributes to port competition studies in the literature.

P. T.-W. Lee (✉)
Department of Business Administration, Soochow University, Taipei, Taiwan
e-mail: paultwlee@scu.edu.tw

J. S. L. Lam
Division of Infrastructure Systems and Maritime Studies, School of Civil
and Environmental Engineering, Nanyang Technological University, Singapore, Singapore

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_4

4.1 Introduction

A container port has become a key element of global supply chain management (SCM) in tandem with international trade development. Table 4.1 enables us to grasp striking structural and geographical changes in the number of container hub ports by region in the world over the last four decades. Numerous European and American container ports have left the top echelons of the world container port rankings. Among the global top 25 ports in terms of container cargo handling throughput, the number of the top ports in Asia increased dramatically from 2 in 1970 to 17 in 2012, while in North America and Europe the incidence decreased from 7 to 3 and from 14 to 4 over the same period, respectively. Out of 312.7 million TEUs of the total container handling throughput in 2012, Asia's share amounts to 78.1 %, totalling 244.3 million TEUs.

Claiming that the underlying principles of port development policy drivers i.e., the Anglo-Saxon and the Continental (European) Doctrines (Bennathan and Walters 1979), are not sufficient to explain the remarkable growth of container port developments over the last four decades, Lee and Flynn (2011) proposed the Asian (Port) Doctrine. Concisely speaking, the Anglo-Saxon Doctrine considers that the port should be financially self-sufficient and make a profit (or at least should not make a loss), while the Continental (European) Doctrine views the port as part of the social infrastructure and hence assesses its value in terms of its contribution to the development of the region and not necessarily primarily in terms of profitability. The Asian Port Doctrine discusses how the 'Asian Doctrine' policy has fostered the construction of or led to the construction of container terminals in Singapore, Korea, China, and Taiwan. In addition, it illustrates the central role of public ownership (with later supportive involvement of private capital), the centrality of the economic policy maker as an agent in port development, pricing approaches specific to the new doctrine, plus providing examples of how Asian governments are more extensive in their investment in waterside and landside infrastructure. In the public enterprise approach, port authorities under the central governments in Korea and China discounted part of the total construction based on the impact of a port on the national and regional economy, and then they set port charges in an administered fashion. On the contrary, the Anglo-Saxon, European and Asian doctrines sit on a continuum of less to more government involvement in ports related to ownership, operation, pricing, financing, and provision of landside and waterside infrastructure. Unlike port developments in the UK and most of those in Europe, Asian port developments have been driven by multi-dimensional roles of central governments as port designer, developer, operator, port pricing maker, mediator, and investor.

Generally speaking, it seems that Asian container ports such as Busan port and most Chinese ports have paid attention to their capacity expansion to meet their own international trade at the initial development stage responding to export-oriented economic policies, while the Singapore container port is more concerned with transshipment cargoes. Port capacity expansion policy requires more aggressive port

Table 4.1 Distribution of the top 25 container ports in the world by region in terms of container throughput. (Source: Compiled by authors from Containerisation International (several issues))

Year	North America	Europe	Asia	Others	Total
1970	7 (1 156 975)	14 (1 974 566)	2 (144 299)	2 (276 112)	25 (3 551 952)
1980	7 (5 789 727)	6 (5 071 076)	8 (6 734 319)	4 (1 886 731)	25 (19 481 853)
1990	6 (9 290 276)	6 (10 830 567)	11 (26 749 331)	2 (2 297 767)	25 (49 167 941)
2000	3 (12 530 252)	6 (20 215 448)	14 (96 523 374)	2 (5 711 587)	25 (134 980 661)
2010	3 (18 966 142)	5 (32 385 576)	15 (211 033 271)	2 (15 806 937)	25 (278 191 926)
2011	3 (19 505 095)	4 (35 470 795)	17 (233 986 586)	1 (13 000 000)	25 (301 962 476)
2012	3 (19 653 285)	4 (35 480 192)	17 (244 331 628)	1 (13 270 000)	25 (312 735 105)

Numbers in parentheses indicate container throughput in twenty-foot equivalent unit (TEU)

pricing and incentive policy with well-established port infrastructure to capture transshipment cargoes. We can see such a crystal case of port pricing policy in the paper of Lee and Lee (2012). They designed an efficient and competitive lease charging system to provide incentives and competitive edge for terminal operators to increase transshipment cargo throughput at the Port of Busan, with factors and rationale underpinning a reasonable lease calculation model. Today's competing ports are facing multi-facet challenges arising from a new logistics paradigm as an integral element of the supply chain in the context of a globalized economic system with free trade agreements, global warming issue, security, and a resilient system caused by natural disasters. In addition, container ports have been devolved in line with development of the globalized economy, accordingly transformation of supply chain systems, and increasing interaction with the cities where the ports are located (Lam and Song 2013). Therefore, port competition may not be improved without tacking the above key issues timely and properly. In other words, ports need multi-dimensional instruments to improve port competitive edge in the context of a socio-economic system, reflecting high service quality, well organized governance system, efficient port performance, and high level of port infra-structure and supra-structure.

Numerous studies have attempted to classify ports into typologies and analyze their roles and functions because port competitiveness should be considered in the socio-economic system (UNCTAD 1994 and 2011a; Ishii et al. 2013; Hayuth 1982; Hoyle and Hilling 1984; Hoyle 1998; Gilman 2003; Flor and Defilippi 2003; Paxio and Marlow 2003; Beresford et al. 2004; Bichou and Gray 2005). Flynn et al. (2011) focus on UNCTAD (2011a); Beresford et al. (2004), and Bichou and Gray (2005) to elucidate port generations in terms of functional and spatial development of port and then suggest the Fifth Generation Ports, so-called 'the port ladder—customer-centric community-focused port'. Flynn et al. (2011) argue that 'Fourth Generation Ports' (4GP) described in UNCTAD (2011a) reflected the focus on internal efficiency, while their survey of current trends sees a strong and intensifying focus on the port functions required by community and needs of port users. Attempted to fill this analytical gap, Flynn et al. (2011) develop a conceptual framework for the "Fifth Generation Ports"

(5GP) with world-class customer-centric and community port representing the next evolutionary step on the port ladder. That would be a new port typology to improve the competitive edge of container ports.

This chapter aims to examine the status of major Asian container ports on whether they have reached the next evolutionary stage of 5GP beyond the existing 4GP, referring to the characteristics on eight items of 4GP and 5GP, namely service quality, information technology (IT), community environmental impact, port cluster, maritime cluster, logistics hub, inland connection, and waterside connection. However, we modified the concept of the 5GP proposed by Flynn et al. (2011) partly because the description of the factors is to some extent vague and does not have a sharp demarcation line among some items for the sake of comparison and evaluation of port generation. The other reason is because some items, e.g., the port and maritime clusters and the logistics hub and inland are not independent functions to achieve the core of the customer-centric community ports as indicated in Table 4.2. This chapter applied the modified concept of 5GP to evaluate inter-port competition of the four major container ports in Asia in a comprehensive way to reflect the above from cross-sectional, longitudinal and horizontal aspects of the port evolution. In this regard, this chapter employs a novel approach with empirical test which combines a descriptive method and a quantitative approach to study port competition and competitiveness. The chapter selects four representative container ports which are among the busiest ports in the world, i.e., Busan, Hong Kong, Shanghai, and Singapore for conducting case studies.

4.2 Customer-Centric Community Ports: the Fifth Generation Ports[1]

Ports are by nature dynamic entities interfacing with various types of users. As the world is being transformed into a globalised economy, ports have been incorporated into a huge, changing and competitive system, providing commercially-oriented as well as customer centric services. The concept of generation type ports first introduced in 1994 by UNCTAD with the third generation port has been followed by several well-founded proposals for the fourth generation ports. Arguing that a four generation framework is insufficient to reflect the port functions required by community stakeholders' requirements and the needs of port users in the rapidly evolving globalised economic system, Flynn et al. (2011) propose the 5th Generation Ports, so-called 'the port ladder—customer-centric community-focused port' based on key differentiating features of 5GP. They list up eight differences between 4GP and 5GP

---

[1] This section is based on Flynn, M., Lee, P.T.W. and Notteboom, T. 2011. The next step on the port generations ladder: customer-centric and community ports, in Notteboom, T. (ed.), 2011. Current Issues in Shipping, Ports and Logistics, Brussels: University Press Antwerp, 497–510.

Table 4.2 Comparison of key features of the fourth and the fifth generation ports

Items	The Fourth Generation Ports by UNCTAD	The Fifth Generation Ports modified by the Authors
Service quality	Meeting regulations and general levels of standards	Finding dynamic incentives to perform beyond basic standards and to meet customers' satisfaction
Information Technology (IT)	Cargo clearance & tracking	IT focuses on one stop service and security to improve port performance and users' satisfaction. IT is not only based on tracking and tracing of both cargoes and information via a 'single window' system but also on performance measurement including gas emission information
Community environmental impact	Regulatory compliance with environmental impact and planning statutes	Active outreach to community stakeholders in port-city interface, planning and decision making process, in particular waterfront development. Active green port policy with rewarding system is envisaged
Port cluster	Handled through land-use planning	Port services provision integral to mission and vision. Port leaders have role as "port cluster managers" in tandem with maritime cluster contributing to generating value-added in in the context of logistics hub
Maritime cluster	Treated as separate from port function	Still functional independent of the port cluster, but subject to clustering, functional interrelated with creative financial incentives to attract shipowner and cargo by creating jobs and value-added
Logistics hub	Logistics developed as a back of port function; and Physical Free Trade Zones and Logistics Parks	Logistics seen as part of a maritime logistics chain; Airport interface for high-value added flexibility; and Advanced Free Trade Zone and Logistic Park functions. This logistics function is interrelated to the feature of 'inland' to maximise its synergy effect

Table 4.2 (continued)

Items	The Fourth Generation Ports by UNCTAD	The Fifth Generation Ports modified by the Authors
Inland	Inland connections develop through natural evolution	Ports develop hinterland strategies through pricing & incentive policies ensuring that evolution does advantage interest of cargo owners and generates efficiency of intermodal system with possible reduction of total transportation costs
Waterside	Port marketing as two dimensional price and quantity approach	Ports developing foreland strategies to capture transshipment cargoes in tandem with SCM through pricing and other incentive policies

The authors modified key features of the Fifth Generation Ports in Table 4.2 taken from Flynn et al. (2011, p. 503)

drawn by their survey among port experts, port service users, and port system managers as shown in Table 4.2, highlighting what factors are important for the ports to characterize as ports moving from 4GP to 5GP.

However, the description of the factors is to some extent vague and does not have a sharp demarcation line among some items for the sake of comparison and evaluation of port generation. For example, Flynn et al. (2011) states that maritime cluster covers functional independent port cluster and maritime cluster of the port cluster. In fact, the two clusters are not only interrelated to each other to achieve synergy effect of 5GP, but also are implemented in a free trade zone and a logistic park. Therefore, the evaluation of a port generation status should be made in a comprehensive way by considering the three factors. In addition, Flynn et al. (2011) describe the feature of service quality as finding dynamic incentives to perform beyond basic standards. We argue that it has to harbor another incentive to meet customers' satisfaction to embody the core of 'Customer-centric Community Ports'. Moreover, the feature of information technology (IT) in the original concept of 5GP is too broad to measure the 5GP status. IT should cover single window system having one-stop service because it is closely related to the feature of a logistics hub. Current ports often receive pressure from their community stakeholders' request to develop waterfront area (Lee et al. 2013) with minimisation of gas emissions from port operation. Therefore, the item should reflect such points into its features as shown in Table 4.2. Furthermore, the function of logistics hub can be strengthened by supporting of the item of 'inland' that requires hinterland strategies through pricing and incentive policies. As for waterside, the purpose of foreland strategies should be added to its features: to improve inter-port competition by capturing transshipment cargoes. Referring to the above critiques on the features of 5GP given by Flynn et al. (2011), we indicated the modified features of 5GP in Table 4.2 to evaluate port generation status of the four ports.

While 4GP is concerned with total container throughput, port efficiency and productivity, this new concept of 5GP is more concerned with attracting and keeping clients and also serving community stakeholders. This implies that the scope and contents of port competition should be expanded on top of traditional factors affecting port competitive edge reflected in 4GP. Therefore, the concept of 5GP contributes to illuminating a new angle and key factors of contributing to improve port competitive edge. Flynn et al. (2011) explains how leading ports meet their customers' and port stakeholders' needs and requirements to make them 5GP, i.e. customer-centric community port. Flynn and Lee (2010) propose a new port classification into five levels: Cargo ports, Logistics ports, SCM ports (bilateral E-ports), Globalised E-ports, and Customer-centric ports.[2] The 5GP is required to raise the bar of the levels up to the customer-centric level by utilising market mechanism, incentives, government policy for port users and port operators. This example can be found in green reward/passport incentive system for shipping lines with performance of low greenhouse gas emission.

4.3 Case Studies

Hub port development in shipping and aviation sectors has become a common trend in the globalized world in association with SCM. Table 4.3 shows the top 25 container ports in terms of annual container throughput in the years 2000, 2010, and 2011–2012. As for 2013, the container handling throughput of the top 10 ports can be gathered at the time of writing and is shown in Table 4.4. Out of the top ten container ports, nine ports are located in Asia, except Dubai ranked on 9th. Since 2010, Shanghai port has become the top container port in the world, handling 29.1 million TEU. In the following years, the port has kept top ranking, handling 31.7 million TEU in 2011, 32.5 million TEU in 2012, and 33.6 million TEU, respectively. Out of the top nine ports in 2013, mainland China has six container ports, Shanghai, Shenzhen, Ningbo, Guangzhou, Qingdao, and Tianjin in ranking order, while in 2006, only Shanghai was ranked on 6th in the world. China has recorded a remarkable growth of container ports within only one decade. In this chapter, Hong Kong is categorised under a Special Administrative Region (SAR), China.

With structural changes of production line and worldwide distribution channels accelerated by the free trade agreements (FTA), container ports have played a critical role in the international logistics and SCM, facing fierce inter-port competition. Consequently, port developers and terminal operators are more concerned with service

---

[2] On the diagramme of the port ladder, see Flynn and Lee (2010), which is quoted by Flynn et al. (2011, p. 501). The customer-centric port concept is influenced by a customer centric strategy described by Charan (2007).

Table 4.3 Top 25 container ports in the world in years 2000, 2010, and 2011 (Source: Compiled by authors from Containerisation International (several issues))

Rank in 2012	Port name	Country	2000		2010		2011		2012	
			TEU (Rank)		TEU (Rank)		TEU (Rank)		TEU (Rank)	
1	Shanghai	China	5,613,000	(6)	29,069,000	(1)	31,739,000	(1)	32,529,000	(1)
2	Singapore	Singapore	17,086,900	(2)	28,430,800	(2)	29,937,700	(2)	31,649,400	(2)
3	Hong Kong	China	18,098,000	(1)	23,532,000	(3)	24,384,000	(3)	23,117,000	(3)
4	Shenzhen	China	3,993,714	(11)	22,509,700	(4)	22,570,800	(4)	22,940,130	(4)
5	Busan	S. Korea	7,540,387	(3)	14,157,291	(5)	16,184,706	(5)	17,046,177	(5)
6	Ningbo	China	902,000	(69)	13,144,000	(6)	14,510,000	(6)	15,670,000	(6)
7	Guangzhou	China	1,429,900	(38)	12,550,000	(7)	14,260,040	(7)	14,743,600	(7)
8	Qingdao	China	2,120,000	(24)	12,012,000	(8)	13,020,010	(8)	14,503,000	(8)
9	Dubai	UAE	3,058,886	(13)	11,600,000	(9)	13,000,000	(9)	13,270,000	(9)
10	Tianjin	China	1,708,423	(32)	10,080,000	(11)	11,580,760	(11)	12,300,000	(10)
11	Rotterdam	Netherlands	6,280,000	(5)	11,145,804	(10)	11,876,900	(10)	11,865,916	(11)
12	Port Klang	Malaysia	3,206,753	(12)	8,870,000	(13)	9,603,926	(13)	10,000,000	(12)
13	Kaohsiung	Taiwan	7,425,832	(4)	8,871,745	(12)	9,636,289	(12)	9,781,221	(13)
14	Hamburg	Germany	4,248,247	(9)	7,900,000	(15)	9,014,165	(14)	8,863,896	(15)
15	Antwerp	Belgium	4,082,334	(10)	8,468,475	(14)	8,664,243	(15)	8,635,169	(14)
16	Los Angeles	US	4,879,429	(7)	7,831,902	(16)	7,940,510	(16)	8,077,714	(16)
17	Dalian	China	1,011,000	(61)	5,242,000	(21)	6,400,030	(19)	8,060,400	(17)

Table 4.3 (continued)

Rank in 2012	Port name	Country	2000		2010		2011		2012	
			TEU (Rank)		TEU (Rank)		TEU (Rank)		TEU (Rank)	
18	Tanjung Pelepas	Malaysia	418,218	(115)	6,530,000	(17)	7,520,000	(17)	7,700,000	(18)
19	Xiamen	China	1,084,700	(51)	5,820,000	(19)	6,450,500	(18)	7,201,700	(19)
20	Tanjung Priok	Indonesia	2,476,152	(19)	4,714,857	(24)	5,649,119	(23)	6,200,000	(20)
21	Bremerhaven	Germany	2,751,793	(17)	4,871,297	(23)	5,915,487	(21)	6,115,211	(21)
22	Long Beach	US	4,600,787	(8)	6,263,499	(18)	6,061,099	(20)	6,045,662	(22)
23	Leam Chabang	Thailand	2,104,950	(25)	5,068,076	(22)	5,660,000	(22)	5,830,000	(23)
24	N.Y./NJ	US	3,050,036	(14)	5,292,020	(20)	5,503,486	(24)	5,529,909	(24)
25	Ho Chi Minh	Vietnam	824,708	(72)	4,367,900	(28)	4,814,000	(26)	5,060,000	(25)

Table 4.4 Top 10 container ports in the world in year 2013. (Source: Compiled by authors from ports' official websites and JIFFA (2014))

Rank	Port name	Country	TEU
1	Shanghai	China	33,611,000
2	Singapore	Singapore	32,579,000
3	Shenzhen	China	23,278,000
4	Hong Kong	China	22,288,000
5	Busan	S. Korea	17,675,000
6	Ningbo-Zhoushan	China	17,327,000
7	Qingdao	China	15,520,000
8	Guangzhou	China	15,309,000
9	Dubai	UAE	13,640,000
10	Tianjin	China	13,000,000

qualities, minimization of logistics and transportation costs, environmental issue, security, contribution to regional economies with increase of value-added, mitigation of conflicts between port and city. Reviewing port devolution from the 1st Generation Port to the 4th Generation Port (4GP), Flynn et al. (2011) proposed the "Fifth Generation Port" (5GP).

With reference to the conceptual framework of the 5GP with world-class customer-centric and community ports, this chapter examines the status of major Asian container ports on whether they have reached the next evolutionary stage beyond the existing 4GP, referring to the characteristics on the eight items of the port generation as shown in Table 4.2. This chapter selects a representative port from each economy having container ports listed on top ten ports, i.e., Busan, Hong Kong, Shanghai, and Singapore. The analysis to the port case study is based on desk research, the authors' field trips to the ports, interviews, and questionnaires. Descriptive approach is also applied to ports' capability and orientation to find their future by increasingly looking at market opportunities through the eyes of their customers and also by adapting to meet the ever higher expectations of their host communities. A more thorough analysis by performing a comparison of the service qualities of the four ports together will be discussed in a separate section, referring to Hu and Lee (2010, 2011) and Lee and Hu (2011).

4.3.1 Case of Busan Port

Busan container port handled 17.67 million TEUs, being ranked on the 5th in the world. (see Table 4.4.) Container terminals at the port were constructed and managed by the Korea Container Terminal Authority (KCTA), a non-profit corporation authorized by the Ministry of Maritime Affairs and Fisheries, from the late 1970s until

2004 when the Busan Port Authority (BPA) was established. During the period, there were four container terminals in the Port of Busan: Busan Container Terminal Operation Corporation (BCTOC), Pusan East Container Terminal (PECT), Uam Terminal Company and Gamman Container Terminal (GCT). All of them were operated by domestic operators. To resolve congestion and prepare for the demand increase due to the expected economic development in China, Korean governments worked with private companies to develop Busan New Port (BNP) with an eventual total of 30 berths with each capacity for the world's largest super post-Panamax vessels. As of November 2013, it has a capacity of 9.4 million TEUs and a hinterland for business logistics and value added services. In the meantime, according to the decentralization policy of the M. H. Roh's government, the Port Authority Act was enacted in 2003. The following year, BPA took over the role of construction, financing, management and operation of container terminals in Busan Port from KCTA. As far as the appointments of and structures of board of directors and president of BPA and their decision making process are concerned, BPA has more autonomous power than KCTA. In particular, this development has paved a way for Busan City Government to some extent to involve the port development and planning process. Although in 2012, the appointment of president of BPA was transferred from President of Korea to Ministry of Oceans and Fisheries (MOF)[3], but the governance and autonomous power of BPA is still weaker than those of port authorities in Europe and North America, such as LA Port Authority. It can be said that BPA is not only theoretically owned in terms of shares but also practically governed to some extent by the central government[4].

Ports receive social pressures from adjacent cities regarding waterfront port development (WPD) owing to a decline in the port functions, changes in port-city interactions, and alterations of a port governance system. The urban planning and port development literature cited numerous cases of WPD (e.g., Hayuth 1982; Robinson 1985; Hoyle 1988, 1989, 2000; Charlier 1992; Gordon 1993, 2004; Craig-Smith 1995; Brown 2008; Lee et al. 2013). Hoyle (2000) proposed a historical outline of the evolving port-city interface with a six-stage model (i.e., "primitive port/city, expanding port/city, modern industrial port/city, retreat of the city from the waterfront, redevelopment of the waterfront, and renewal of port/city links" (Hoyle 2000, p. 405). Grobar (2008) pointed out conflicts arising between stakeholders and policymakers who want to preserve the benefits of port users, to maintain port functions and local communities that transform the port area into an amenities function, generating a friendlier port-city environment with added value.

---

[3] The Ministry of Oceans and Fisheries (MOF) established in 2013 under President Park is a former organization of the Ministry of Maritime Affairs and Fisheries (MOMAF), which had been merged into the Ministry of Land, Transportation, and Maritime Affairs in 2008 under Lee's Government.

[4] The concept of BPA and its governance system are quite different from those of port privatization of European and American cases addressed in port studies literature. Therefore, a careful approach and interpretation is required to deal with port authorities under Korean port authorities under the Port Act, such as Busan Port Authority (BPA), Incheon Port Authority, Gwangyang Port Authority and so on, although they have been named 'port authority'.

Since 2010, Busan Port has faced a similar case because the opening of Busan New Port (BNP) and decline demand for general cargoes in central piers of the Busan Northern Port have triggered to debate over the WPD in the port. BPA has played a central role in accommodate and reflect views of Busan city government and citizen groups, holding public hearings and conferences to draw the WPD plan. As a result, the Korean central government, BPA and the city government have concluded a master plan of WPD. This process is a good example of one of key features of 5GP to serve community stakeholders. In this regard, Busan Port is moving forward to 5GP in terms of 'community environmental impact'.

A cruise terminal with one 80,000 gross tons berth started its operation in April, 2007. The terminal recorded 335 ships and 356,593 passengers in the period of 2007–2012. Although the performance of the cruise terminal is very small in terms of connectivity and frequencies and cruise ship size at Singapore and Hong Kong, the cruise terminal of Busan has been gaining power of positive influence on the sector of waterside development and increase of value added in the sector of port cluster.

Recognising that Korea is well located to link the entire area of Northeast Asia with the world utilizing Incheon International Airport, Busan seaports, and the surrounding free trade zones, the Korean government has launched 'National Logistics Master Plan, 2011–2020'. In light of this comprehensive policy, the Korean government has created Port Management Information System (Port-MIS) 2.0 on 19 April 2010, of which previous version Port-MIS was run by EDI method since early 1990s (Lee et al. 2000).

'Yes! U-Port' (integrated management brand for shipping & port-logistics) is a very-first collaborative business project planned by the government and private enterprises to upgrade the maritime industry in Korea. It is a representative port-logistics system built by integration of domestic Shipping and Port-logistics Information System (SP-IDC) (see Fig. 4.1). It provides total logistics information infra to realize safer, faster and easier information technology (IT) port network that supports harmonious flow of port-logistics business, regardless of time or place. Korean integral port systems include ATOMS (Advanced Terminal Operation & Management System), GCTS (Global Container Tracking System), and GICOMS (General Information Center on Maritime Safety & Security). The overall system of IT in the Busan Port has moved from 4GP to 5GP.

IT in port is interrelated among stakeholders involving logistics and maritime cluster to develop maritime logistics chains, to generate value-added and to activate free trade zone adjacent to ports. Single window system as IT in port is a facility that allows parties involved in trade and transport to lodge standardized information and documents with a single entry point to fulfill all import, export, and transit-related regulatory requirements. It simplifies processes through integration with government agencies and port authorities' system as well as port users' individual systems. In this regard, it is meaningful to make a comparison of the overall IT system among the four port cases (see Table 4.4). Seven aspects are taken into consideration for evaluating the overall IT system, namely, single window, integrated function, government and private enterprises collaboration, cargo tracking system, maritime safety and security,

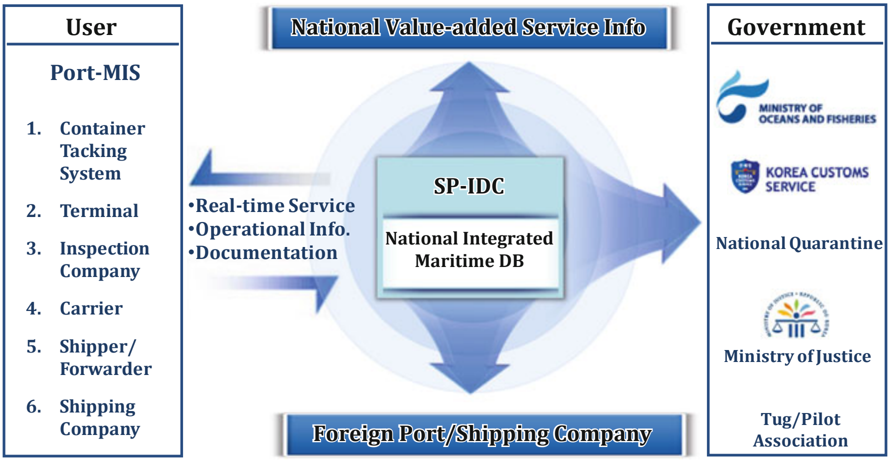


Fig. 4.1 Concept of shipping and port-logistics information system (SP-IDC). (Source: KL Net http://www.klnet.co.kr/english/product/product1_2.html, Accessed 30 Oct 2013)

radio frequency identification (RFID), and mobile service by using smart phone. As far as the single window system is concerned, Shanghai container port is quite behind the three ports, i.e., Hong Kong, Singapore, and Busan.

Now we turn to the discussion of port cluster and maritime cluster development in Busan Port.

According to the World Bank's Logistics Performance Index (LPI) in 2012, Korea scored 3.7 ranking 21 in the world. In terms of the performance in international shipment, Korea ranked 12 among 155 economies (World Bank 2012). Busan has an international seaport having a high record of connectivity and service frequency with well-organized maritime auxiliary services. However, Busan is not a free port with flags of convenience registration. No tax incentives are provided for foreign shipowners and maritime related businesses, unlike Hong Kong and Singapore. As discussed above, Busan City Government has neither independent power to implement logistics and tax incentive policy nor authority to use port land. These conditions are unfavourable conditions to promote port cluster and maritime cluster. The BNP has the Busan-Jinhae Free Economic Zone (BJFEZ) and the Kimhae International Airport nearby the seaport. The free economic zone has arranged the master plans for five areas: (1) logistics, distribution and maritime affairs, (2) high-tech, air logistics, high-tech & manufacturing, (3) high-tech, R&D, Busan science and industrial park, (4) residential complex, mechatronics, R&D professional education, and (5) tour and leisure, logistics and distribution. Therefore, it is a good opportunity to activate clustering. As of December 2012, 30 logistics consortiums from home and abroad operate in the distripark at BNP, creating more than 611,000 TEU cargo and 1300 of new jobs. However they have not shown formidable outcomes owing to several reasons. Among others, the free economic zone is governed by BJFEZ Authority under the Ministry of Trade, Industry and Energy so that they cannot have speedy and efficient decision making system among the port authority, two city governments (Busan and Jinhae), and the BJFEZ authority. In addition, each stakeholder has no full autonomous decision making power so that they should get permission from their ministerial level. It implies that foreign investors and service users in the zone may not expect one stop services like Singapore.

This is also interrelated to a negative impact on the item of the 'maritime cluster' because such an inefficient system is not easy to draw creative financial incentives to attract ship owners and cargo by creating jobs and value-added referring to the free zone. Having considered the above argument, Busan is on the stage of 4GP in the system of 'logistics hub' and 'maritime cluster'.

Although Korea is a peninsular, South Korea is divided by demilitarised zone so that it is like an island. Therefore, Busan port has a limited boundary hinterland with well-connected by railways and highways within South Korea. As discussed above in this section, Korea has well established IT systems for cargo tracking and single window system that enables port users and logistics providers to enjoy smooth flow of cargo and information. Korea has conducted feasibility studies of Trans-Korean Railway (TKR)/Trans-Siberian Railway (TSR) plans for container cargoes bounding for Europe. (KMI 1992; Lee 1993) Having considered current geo-political situation, opening hinterland connection of TKR and TSR is expected to be a long way to go. Therefore, Busan port has limitation to capture gateway cargoes for inland connectivity to China, Mongolia, and Russia that should pass through North Korea. The inland connectivity competitiveness of Busan port would not be improved unless such TKR and TSR are developed. There is another potential market of inland connectivity for Busan Port. It is the Kaesong Industrial District (KID) in Kaesong City, North Korea, which started to produce mainly manufacturing goods in 2005 to achieve future of national co-prosperity as a way for new history of inter-Korean reconciliation and cooperation. It produced approximately 2 billion US$ in the period of 2005–2012 (KIC 2013). The first joint inter-Korean industrial complex, KID, employed labors from two Koreas working in the same companies for the first time in 60 years of the division of Korea, aiming to keep peace and co-existence and to be the internationally competitive industrial complex. However it was shut down owing to North Korea's attack to South Korea in 2012. Although it has been reopened in September 2012 after full shut-down for a couple of months, cross-border hinterland transportation between the two Koreas is still volatile owing to their political conflicts. Some container cargoes generated from the industrial complex can be shipped at Incheon Port or Pyeongtaek Port located in the west coast of Korea, having the advantage of hauling distances compared to Busan port as well as better location to be connected to Chinese ports. Conclusively, Busan Port is still in the 4GP stage in terms of inland connectivity.

Next, let's move on to the item of 'waterside' in the key features of 5GP. According to this item of 4GP, port marketing relies on a two dimensional price and quantity approach. In order to raise its level up to 5GP, a port is required to develop foreland strategies to capture transshipment cargoes and improve inter-port competitive edge through pricing and incentive policies. Busan is located on the main trunk of the route of west coast of North America-East Asia-Indian Ocean-Mediterranean-Europe, serving by some 61 international shipping lines with over 358 container liner services per week (BPA 2013). The transhipment cargo portion over the last five years at the port ranges from 43.2 to 47.8 % of total container throughput handled at the port. The sharp rises of Chinese ports such as Shanghai, Ningbo, Tianjin, and Qingdao, which are relatively in a competitive positions with Busan to capture transhipments cargoes generated from northeast China and Japan, has caused Busan port to redesign port pricing and incentives. The Korean financial crisis was a turning point to change port pricing and incentive policies at Busan Port (Lee and Lee 2012). Thanks to the introduction of a new port pricing into the port, the transhipment cargoes at the port increased from 848,000 TEUs in 1999 to about 4.3 million TEUs in 2003 including empty containers, despite the fact that the Korean economy suffered from the financial crisis during that period. Reviewing the overall proactive port pricing and incentives, it can be said that Busan Port is moving to the stage of 5GP with proactive pricing policy.

4.3.2 Case of Hong Kong Port

The role of Hong Kong as a relay hub for China has been established when Hong Kong was the only gateway linking the semi-open Chinese economy with the rest of the world before the implementation of the open door policy in 1978. The port of Hong Kong has developed to a world-class standard while the economy of Hong Kong enjoyed a fast economic growth rate. However, China's foreign trade has experienced a remarkable development since its economic reform to open trade. Corresponding to its prosperous foreign trade, China's maritime transport and port industry have grown rapidly since the 1990s. The port of Hong Kong faces competitive challenges from the ports of the Chinese mainland due to the northward shift of most of Hong Kong's export-oriented manufacturing activities into the Guangdong Province, particularly in the Pearl River Delta (PRD) region in South China. In particular, the major neighbouring container port of Shenzhen has developed as a key competitor to Hong Kong (Lam 2011). Since 2013, Shenzhen became the third busiest container port in the world and ranked higher than Hong Kong in terms of container handling throughput (JIFFA 2014).

It can be seen that the port of Hong Kong has experienced dramatic changes in the economic environment over the past decades. Correspondingly, the port has also evolved from a cargo-centric orientation to a more logistics- and information-focused status. The first item for assessing the extent to which the port has evolved is service quality. Hong Kong is one of the busiest ports in the world in the three categories of shipping movements, cargo handled, and passengers carried. The major cargo handled is container traffic, which stood at 22.29 million TEUs, making the port the fourth busiest in the world in 2013 (JIFFA 2014). Capable of handling such a huge volume, Hong Kong is one of the world's most efficient container ports (Cullinane and Wang 2007). The port also provides comprehensive services of world class quality to visiting ships, including more sophisticated services like vessel traffic services, Differential Global Positioning System broadcasting and waste reception services (MD 2006). Therefore, having performed beyond basic standards, we consider the port a 5GP in terms of service quality.

With regards to IT, Hong Kong Port has a long history in utilising the world wide web technology in disseminating information and providing services to customers and users. The Hong Kong Marine Department maintains a number of IT systems for port management and facilitation such as the Vessel Traffic Management System, the Dangerous Goods Information System and the Marine Department Electronic Business System (MD eBS) (MD 2006). For example, MD eBS provides a one-stop electronic submission solution to all port formality procedures in Hong Kong via the internet. This has significantly increased port formality efficiency. The port is operated by several international professional terminal operators with respective IT portals. OnePort Limited is an initiative by Hongkong International Terminals Limited (HIT), Modern Terminals Limited (MTL) and COSCO-HIT Terminals to offer an integrative IT platform. In particular, through the establishment of its Electronic Data Interchange centre, all the container terminals, a vast majority of shipping lines and a sizable number of freight forwarders are connected (OnePort 2013). Hence the IT standard in the port is customer- and stakeholder-focused, making Hong Kong a 5GP in the IT aspect.

In terms of community environmental impact, Hong Kong Port is quite active and has done beyond regulatory compliance. We will discuss the port's commitment at the governmental level and the industry level. At the governmental level, the Marine Department under the Hong Kong Special Administrative Region (HKSAR) Government functions as the maritime authority and is responsible for the administration of the Port of Hong Kong (MD 2013). The Marine Department is responsible for implementing policies that have been endorsed by the bureaucratic level with regard to the protection of the marine environment from ships, in particular air pollution and water quality. These can be accomplished by controlling pollutions from ships in the waters of Hong Kong through vessel traffic management, harbor patrol, and enforcing international maritime regulations as stipulated by the International Maritime organization (IMO) (Zhu and Agarwal 2011). Under the Transport Branch in Transport and Housing Bureau, there are three councils, namely Hong Kong Maritime Industry Council, Hong Kong Port Development Council, and Hong Kong Logistics Development Council, which act as advisory bodies to advise the HKSAR Government on Hong Kong's development as an international leading maritime centre as well as a logistics services centre. The councils play an active role in offering suggestions in green port policy issues. Regardless of these efforts, the HKSAR Government's outreach to the community in terms of green port initiatives is negligible. The Government adopts a laissez-faire policy which means that transactions between private entities are free from government interventions (Schiffer 1991).

Hong Kong Port's community outreach in the environmental aspect is largely led by the maritime industry. For example, in addition to government schemes, shipowners under Hong Kong Shipowners Association supported to adopt a lower global sulphur cap of 0.5 %. Craft operators also take initiatives to reduce vessels' fuel consumption. Hong Kong's largest tugboat operator the HUD Group commits to become the world's first tug company being completely carbon neutral (Galbraith et al. 2008) which is a major step taken in the tug industry. These green initiatives have created positive community environmental impact on the port of Hong Kong. This shows that the port is proactive in working towards self-motivating ecological solutions. Such deportment is more sustainable than merely regulatory compliance. The port would be able to achieve more if the government can play a more active role in community outreach, such as creating discussion forums without direct intervention on private entities. As a whole, Hong Kong is considered moving towards a 5GP in the community environmental impact aspect.

Now we turn to the discussion of Hong Kong's port cluster and maritime cluster development. Taking an overall review throughout major international maritime clusters, it can be observed that many maritime clusters, including Hong Kong, developed from port operations since the early stage (Zhang and Lam 2013). Thus the port cluster and the maritime cluster are interlinked with regards to their development. In addition to port roles and functions, institutional structuring varies significantly from generation to generation (UNCTAD 1992). Although Hong Kong's port service quality has achieved a Fifth Generation Port's standard as discussed above, its institutional structuring has not been transformed for port leaders to take the role as "port cluster managers". In terms of the HKSAR Government, it plays an essential role in port cluster land-use planning given the fact that space constraint is a major challenge to the port of Hong Kong (Yap and Lam 2013). Hong Kong Port Development Council provides advices to the HKSAR Government on port development strategies and port planning. However, port cluster development is to a large extent market driven also due to the laissez-faire policy adopted by the Government. This also means that not being an active port cluster manager is the Government's choice. Therefore, we consider that Hong Kong's port cluster remains in the Fourth Generation status.

Nevertheless, as the port of Hong Kong faces fiercer competition from the mainland Chinese counterparts, the port has evolved to offer more high value-added maritime services in addition to cargo handling. Hong Kong possesses its unique competitive advantages versus Shenzhen as a maritime cluster. Hong Kong has an independent and transparent legal system. It is also a well-established financial centre. These advantages are crucial for maritime services such as marine insurance, ship brokering and chartering, shipping finance, and maritime legal services. These advantages are also difficult for other competitors to replicate. For example, Hong Kong is successful in attracting more shipowners to register their ships under its flag over the years. As at end June 2013, the Hong Kong Shipping Register has a fleet of more than 83 million gross registered tonnes, making Hong Kong the fourth largest shipping register in the world following Panama, Liberia, and Marshall Island (HKTDC 2013). As noted by Zhang and Lam (2013), Hong Kong is an advanced maritime cluster which is regarded as the supply chain hub in global/regional economic and trade market, enjoying massive economies of density and scope by the effect of a hub-and-spoke system. The important characteristic is integrated resources allocation. Such advanced maritime clusters integrate not only products but capital, information and technology as well. Hong Kong's maritime cluster has reached a 5GP status.

As for the logistics hub status, Hong Kong is in the process of transforming from a freight transport hub city to a knowledge-based global supply chain management centre (Wang and Cheng 2010). According to the World Bank's Logistics Performance Index (LPI) in 2012, Hong Kong scored 4.12, ranking second in the world after Singapore. In terms of the performance in international shipment, Hong Kong ranked first among 155 economies (World Bank 2012). The seaport and transport network offer world class logistics infrastructure. It is evidential that Hong Kong has achieved a very high international standard in logistics services. Other than industry efforts, the HKSAR Government plays an important role by setting the whole of Hong Kong as a free port that does not levy any customs tariff and has limited excise duties. This creates a low cost and open environment for logistics activities. Furthermore, Hong Kong Logistics Development Council provides a forum for public and private sector stakeholders to discuss and co-ordinate matters of concern to, or affecting, the logistics industry. Overall, Hong Kong is regarded as a 5GP in the logistics aspect.

Even Hong Kong Port has achieved superior logistics performance, its inland connectivity is considered a major weakness that undermines the port's competitiveness. Hong Kong is still in the Fourth Generation stage in terms of inland connectivity. Cross-border hinterland transport between Hong Kong and Mainland China is often congested and inefficient. With increasing number of factories being relocated to inland provinces in China, to enhance the PRD ports' performance and in particular for Hong Kong to continue growing, there is an urgent need for better inland transport infrastructure, including roads and railway. The Hong Kong–Zhuhai–Macau Bridge is an ongoing construction project due for completion in 2015 for enhancing hinterland connection (HZMB 2013). A key reason for manufacturing companies to shift towards inland China is due to the connectivity to Shanghai and Ningbo ports (Levesque 2011), which are the major gateways of export to the global market. Therefore, improving hinterland connection enhances not only the competitiveness of Hong Kong, but also the competitiveness of the PRD ports as a whole.

Lastly, Hong Kong has good seaward connectivity in terms of shipping services deployed. Hong Kong Port is served by some 80 international shipping lines providing over 380 container liner services per week connecting to over 550 destinations worldwide (HKPDC 2013). Thus the port is a vital node for numerous links with substantial volumes connected to various forelands. Waterside connectivity is a key performance indicator to analyse ports (as nodes) and routes and shipping lines (as links) that are embedded within the maritime supply chain (Lam 2011). However, liner networks are ephemeral and dynamic since container shipping lines periodically restructure their networks to adjust to the demands from the market. Building on its high service quality and efficiency, Hong Kong's waterside port marketing strategy focuses more on growing transhipment traffic. Over the past decade, Hong Kong has handled increasing proportion of transhipment cargo generated by hub-and-spoke and interlining shipping operations carried out at the port. The port has enhanced its role as a transhipment hub serving larger number of supply chains with various origins and destinations in East Asia (Lam and Yap 2011a). Therefore, Hong Kong has done well as a 4GP in waterside connection, but currently has not evolved further to a more advanced stage.

4.3.3 Case of Shanghai Port

The Shanghai Port handled 5.6 million TEUs in 2000 as the 6th busiest container port in the world. In one decade, the port has been ranked as the top container port in the world, handling 29 million TEUs in 2010. (see Table 4.4) Since then, the port has kept remarkable record in the container cargoes volume. The port's development attributes to a typical causal relationship between the central government role and globalized economic development. Shanghai and the ports in the PRD and some Chinese coastal ports, such as Ningbo, Qingdao and Tianjin are gateway ports of vast hinterlands in China. Although Shanghai Port played a leading role in handling container cargo volumes, it could not meet the fast-growing demand for cargo volumes in terms of deep-water berths, connecting road for hinterlands, container handling facilities, and modern logistics system until after six years of feasibility studies of the Yangshan project. China completed the first phase of Shanghai Yangshan Port on the islands of Xiao Yangshan and Da Yangshan in Hangzhou Bay. The location is 25.7 km away from Shanghai's southern coast and under the jurisdiction of Zhejiang Province. This successful project was a turning point to make Shanghai to put it on the world's top container port since 2010. As briefly mentioned in the section of Introduction, the port development was driven by the Asian Doctrine under the central government which has contributed to solving conflicts between different provinces and Shanghai Municipal Government, integrating several shipping, port, and terminal agencies, and providing financial support for the project including port infrastructure and social infrastructure outside the port.

Shanghai port does not have a single window system to share cargo information among the Chinese ports. Table 4.5 shows a comparison of information technology system in the four container ports under the case study in this chapter. Shanghai has no collaborative interaction with service by using mobile smart phone as an integrated function in the port IT system. As can be seen in the Korean case, IT service development is required to develop close collaboration between government agencies and private enterprises to share cargo information and trade information among stakeholders and to integrate them into IT. We consider Shanghai Port a 4GP in terms of IT.

In terms of community environmental impact, the Chinese government and Shanghai authorities acknowledge the importance of controlling pollution and preserving the environment. The port of Shanghai has implemented various green policies. Pricing policy is the most common one which imposes fines on marine oil spills. The pricing initiatives in Shanghai are mostly penalty schemes. By fining the wrong doers stated in laws and regulations, Shanghai adopts the principle of letting the polluters pay for pollution (Lam and Notteboom 2014). Marine Environment Protection Law

Table 4.5 Comparison of IT systems of container ports in Asia. (Source: Authors)

Service and Function Items	No specific system (Shanghai)	Portnet (Singapore)	OnePort (Hong Kong)	Yes! U-port (Busan)
Single Window	X	√	√	√
Integrated Function	Limited	√	√	√
Government and Private Enterprises Collaboration	Limited	√	√	√
Cargo Tracking System	√	√	√	√
Maritime Safety & Security	√	√	√	√
RFID	√	√	√	√
Mobile Service By Using Smart Phone	X	√	Developing	√

states clearly under what circumstances should polluter pay for what amount covering both port industrial activities and port expansion (PRC 2000). In the meantime, regulatory control in Shanghai conforms to the International Maritime Organization convention which is MARPOL Annex VI on air pollution (IMO 2012). The content of regulatory control also covers a prohibition where certain activities causing damage to the marine environment are prohibited by the Chinese law on the Prevention and Control of Marine Pollutions from ships. In addition, it is stipulated in law to monitor the air quality and water quality to keep track of the port's environmental performance (PRC 2000). For example, monitoring the port's carbon footprint is highlighted in the green port guide information. However, there are not much specific measures implemented yet mainly due to the relatively recent commencement of these green initiatives in 2012 (China ACC 2011). Hence, Shanghai Port performs well as a 4GP in regulatory compliance with environmental impact and planning statutes. In order to progress to 5GP in the environmental aspect, it has to conduct more active outreach to the community in planning and decision making process. While the central government implemented the project, Luchaogang New Harbour City was designed as a maritime cluster. As one of the 11 satellite towns of Shanghai's urban planning, it is located in the southeast corner of Nanhui District, about 30 km away from Yangshan Island and 30 km away from Pudong International Airport. The city has functions of container distribution and storage, offshore processing, shipping market, logistics centre, residence, financial and commercial service, and tourism (Song 2007).

Shanghai is strategically located next to one of the world's largest manufacturing regions, the Yangtze River Delta (YRD). With the manufacturing cluster as backyard, Shanghai is now establishing a presence as an international shipping and financial centre. According to the central government of China, it plans to create a world-class international maritime centre in Shanghai by year 2020 (The State Council of China 2009). While cargo handling activities have expanded rapidly for a long period since the 1990s, more attention on developing expertise on maritime services was given in recent years. Shanghai's Shipping Service Centre is an integral component to cultivate a world-class maritime cluster. Among other objectives, the Shipping Service Centre especially aims at establishing a multifunctional business district to serve the maritime sector. One of its core functions is to develop legal expertise. International commercial arbitration centres and specialized maritime arbitration services have been set up. As the experiences of Hong Kong and Singapore show, a developed maritime cluster counts maritime arbitration as an important component. Shanghai has also spent great efforts in the aspect of research and development. For instance, in 2010, Zurich Financial Services in collaboration with the Pudong New Area government created the Zurich Research and Development Centre which focuses on developing new solutions for the shipping and finance industries (McKinnon 2011). Though Shanghai draws much of its strength from the size of its port and the burgeoning YRD, it takes plenty of resources and long time to establish advanced maritime services. Shanghai is still not considered "Singapore" or "Hong Kong" like maritime cluster because its complementary service sector has a long way towards maturity. It can be regarded as a 4GP maritime cluster. However, with central government's strong support, Shanghai's maritime cluster is expected to grow rapidly.

Given the fact that Shanghai Port serves a huge manufacturing region in the YRD, connectivity to the vast hinterland and the associated logistics services are integral to the port's performance. According to the World Bank's Logistics Performance Index (LPI) in 2012, China scored 3.52, ranking 26th in the world. China has the same score as Ireland, and has outperformed a number of developed countries, such as Portugal and New Zealand (World Bank 2012). As a developing country, China is considered highly ranked in the logistics indicators. Separate scores for Shanghai are unavailable from World Bank's assessment. However, since Shanghai is located at among the most established region in China, the LPI for China can be taken as a reference point that Shanghai would be at least reaching the standard for the overall China. Shanghai has ample inland distribution facilities and infrastructure which is regarded as a major advantage. Nevertheless, as revealed by the relatively lower ranking of customs performance (ranked 30th), Shanghai's customs has room for improving its charges, flexibility and efficiency (Li 2011).

Thanks to the leading role of the central government, Shanghai Port has developed to world-class standard while the Chinese enjoyed the highest economic growth rate over the last two decades. China's foreign trade has experienced a remarkable development since its economic reform to open trade. At the same time, the port faces competitive challenges from neighbor ports, especially Ningbo as well as Pearl River Delta ports such as Hong Kong and Shenzhen. Shanghai has to some extent common hinterlands with them except Busan. In addition, Shanghai has faced fierce competition with Busan to capture transshipment cargoes. Along with the BNP development plan, Chinese ports, including Shanghai, Ningbo, Tianjin, and Qingdao ports, have also invested heavily in port development, to capture transshipment cargoes generated in Northeast Asia as well as gateway cargoes for the world trade. In addition, according to Chinese foreign and economic policy with Africa and South America, COSCO in collaboration with their strategic alliance shipping lines has started to deploy fleet on the routes to the two continents. Shanghai port is one of the Chinese ports to handle cargoes bounding for the region. Having considered the above developments, Shanghai is doing well for establishing port infrastructure to develop a port cluster and logistics hub. The Shanghai port works towards achieving an important goal of 5GP having smooth and dynamic inland transportation linked to the port in developing SCM with hinterland strategies via pricing and incentive policies for port users. As for pricing, the central government's influence is still strong and has low flexibility responding to market situation and port users' needs. Therefore, Shanghai remains at the stage of 4GP in terms of port cluster and logistics hub shown in Table 4.2.

Related to established logistics infrastructure, Shanghai Port is also the biggest transportation hub in China. The port is connected to an extensive transport network, including waterway, railway, road and air transport penetrating the vast hinterland in the YRD and central part of China. In particular, the domestic road network makes transport flexible (Li 2011). Shanghai International Port Group (SIPG) announced its direction to embark on three major strategies to anchor Shanghai's position as the premier container hub port in East Asia. The first strategy is called the Yangtze River Strategy. It aims to secure a commanding presence in all the major ports of the Yangtze River, namely Chongqing, Wuhan and Nanjing, in order to secure cargoes from Central China. The second strategy, namely Northeast Asia Strategy, aims to develop a coastal feeder network which is targeted at securing transshipment cargoes from other coastal ports in China (SIPG 2013). Shanghai is increasingly a favourable port of call serving growing number of regions and supply chains. The performance of the port was most impressive in Europe–Far East trade where liner shipping capacity calling at the port grew unprecedentedly and surpassed Busan since 2003 (Lam and Yap 2011b). Therefore, Shanghai Port has developed hinterland strategies through incentive policies and we can consider the inland transport network aspect towards a 5GP. The waterside connectivity has enhanced over the years, supporting import and export activities but also directing the port to focus more on transshipment traffic. This aspect would be classified as 4GP.

4.3.4 Case of Singapore Port

Singapore is one of the world's busiest ports with a leading position in terms of annual vessel arrival by shipping tonnage, transshipment container throughput, and bunker sales. It is home to more than 120 international shipping groups. In the year of 2012, annual vessel arrival tonnage hit a new record high of 2.25 billion gross tons (GT). Container throughput also registered a record high of 31.6 million TEUs in the same year while total cargo tonnage handled reached 538.0 million t. In terms of bunker sales, a total 42.7 million t were sold, enabling Singapore to remain as the world's top bunkering port in 2012. The Singapore Registry of Ships also ranked amongst the top 10 ship registries in the world with a total tonnage of 65 million GT flying the Singapore flag (MPA 2012). In 2013, container throughput handled by the port continued to grow, reaching a volume of 32.58 million TEUs (see Table 4.4).

However as a result from Tanjung Pelepas, a major port located in the Malaysian state of Johor, emerging as a credible alternative for transshipment operations, Singapore port also faces competitive challenges.

To face the challenges and competition posed by emerging transshipment hubs that are gunning for the Europe-Asia and intra-Asia container traffic, Singapore is in the process transferring itself into a 4GP. As a 5GP, service quality is a very important criterion, including in the area of port operations. Similar to Hong Kong, Singapore is one of the world's most efficient container ports (Cullinane and Wang 2007). Being the world's busiest container transshipment hub, Singapore is a world-class seaport that is able to deliver a seamless stream of activities related to the whole spectrum of container shipping operations, such as advanced container shipping infrastructure and facilitation of ancillary services including logistics and freight forwarding. Singapore was also named the Best Seaport in Asia while PSA Singapore Terminals was conferred Best Container Terminal in Asia for the 24th time at the 27th Asian Freight and Supply Chain Awards 2013 (Transport Weekly 2013). In addition, to further support the increasing push in the port industry, the Workforce Development Agency (WDA) worked closely with the port industry by conducting competency based training for port employees and professionals. Together, they have designed and developed the Port Services Workforce Skills Qualifications (WSQ) competency framework which can be used to facilitate skills development and career progressions (WDA 2012). Therefore, we can consider Singapore to be a 5GP attributed to its outstanding service quality.

In terms of IT, PSA as Singapore's container terminal operator has achieved a lot of recognitions globally especially with its flagship product PORTNET, which is the world's first nation-wide business to business (B2B) port community solution system. PORTNET helps shipping lines, hauliers, freight forwarders, and government agencies in terms of managing information and synchronizing complex operational processes. Most recently, their new product PORTNET® Mobile can allow users to access the website via any mobile devices at anytime and anywhere with information such as container status and berthing enquiry (PSA 2013). In the meantime, the Maritime and Port Authority of Singapore (MPA) also provides a community-based system named Marinet which is an E-service that helps to achieve faster clearance of port and shipping documents as well as disseminating critical ship arrival and departure information (MPA 2013a). In addition, there is also the Tradenet system which is used for the purpose of facilitating trade through simplifying and harmonizing formalities and procedures, and doing away with paperwork for the trade community. Hence, the port of Singapore provides a customer and stakeholder focused system which qualifies Singapore to be a 5GP in the IT aspect.

Regarding community environmental impact, the port of Singapore is increasingly active and has come up with three different programmes to reduce the environmental impact of shipping. These are namely the Green Ship Programme, the Green Port Programme, and the Green Technology Programme. In 2011, MPA Singapore pledged S$100 million for funding these programmes. In 2013, several enhancements were announced by the Minister of Transport to further encourage companies to adopt environmentally-friendly shipping practices (Maritime Singapore Green Initiative

2013). For example in year 2011 under the Green Ship Programme, Singapore-flagged ships which adopt energy efficient ship designs exceeding IMP's Energy Efficiency Design Index would enjoy 50 % reduction of Initial Registration Fees and 20 % rebate on Annual Tonnage Tax. In 2013, it was announced that if the ship adopts both energy efficient ship designs and approved SOx scrubber technology exceeding IMO's requirements, the ship will enjoy 75 % reduction of Initial Registration Fees and 50 % rebate on Annual Tonnage Tax (Enhancements to Maritime Singapore Green Initiative Annex A (2013). As Flynn et al. (2011) suggested a case to raise the bar of the five levels for 5GP in the environmental sector, Singapore Port implements a marketing tool with incentives to create a win-win situation between the shipping companies and the port by gas emissions. This has proved that Singapore is moving towards a 5GP in the community environmental impact aspect.

Now we continue to discuss Singapore's status in terms of port cluster and maritime cluster development. Similar to Hong Kong and Shanghai, Singapore's maritime cluster developed from port operations since the early stage (Zhang and Lam 2013). Thus the port cluster and maritime cluster are intimately connected with regards to their development. With huge support from the government and MPA, the institutional structuring of Singapore Port has transformed from being a leading port in terms of cargo throughput to becoming a leading promoter and facilitator of port and maritime cluster development. The Singapore port and maritime clusters are anchored on three key areas: the port, the shipping community, and ancillary services that are required to support the first two areas. Singapore is well known as a global transshipment hub and hub for vessel traffic. The port city also hosts a large community of shipping lines and shipping agencies. According to UNCTAD's Review of Maritime Transport 2012 report, Singapore has the 11th largest fleet in terms of ship ownership, being 38.6 million deadweight tons or 2.8 % of the world's total (UNCTAD 2012). As for ancillary services, the port city has made great strides in the past decade, attracting several international companies involved in activities such as shipbroking, ship finance and classification societies among others to establish offices in Singapore. For example in the area of ship broking, the top twelve companies in this field have presence in Singapore, second only to London.

The maritime cluster employs more than 170,000 people and contributes around 7 % of Singapore's GDP. MPA has introduced the Maritime Cluster Fund (MCF) to facilitate the growth of Singapore's maritime cluster by supporting the industry's manpower and business development efforts as well as its drive for productivity improvements. There are three key components under MCF which are the MCF-Manpower, the MCF-Business and the MCF-Productivity. Recently on 1st June 2013, MPA introduced MCF-Productivity which is to support initiatives by the maritime industry that will lead to productivity gains (MCF 2013). In addition, Singapore is reputed for its quality Ship Registry, which is administered by MPA. In 2012, it ranked the fifth largest registries in the world with more than 4000 registered vessels, totaling over 65 million gross tons (MPA 2013b). The maritime cluster in Singapore also consists of a host of shipping and port-related activities such as ship chandling, ship survey, ship repair, marine insurance, ship chartering, and maritime legal services among others. From the rapid and strong development of the port and concrete support from MPA, we can see that Singapore is very proactive in developing its port and maritime clusters. That is why we consider that Singapore's port cluster and maritime cluster have reached the 5GP status.

As for the logistics hub status, as stated in the Logistics Performance Index 2012 published by the World Bank, Singapore ranked first among 155 economies. Singapore also ranked first for efficiency in terms of customs clearance and timeliness of shipments delivered (2012). In addition, Singapore is also ranked first for ease of trading across borders in its Doing Business report for six consecutive years from 2008 to 2013 (Channel news Asia 2013). As a favorite destination for multinational corporations (MNCs) to establish regional distribution centers, Singapore's strategic location, world-class infrastructure and excellent connectivity have made it a competitive global logistics hub and supply chain management centre. Singapore has remained consistently competitive in terms of the efficiency of customs clearance process, quality of trade- and transport-related infrastructure, ease of arranging competitively priced shipments, quality of logistics services, ability to track and trace consignments, and frequency with which shipments reach the consignee within the scheduled time. With all these accomplishments and achievements, Singapore has established its status as a 5GP in the logistics aspect.

In the area of inland connectivity, the local hinterland of the port of Singapore effectively consists of three regions. The first is the Johor-Melaka region lying to the north of Singapore. Hinterland transportation is organized primarily by road haulage despite the presence of a railway line running north-south from Kuala Lumpur into Singapore. Most hauliers prefer to use the North-South Expressway that runs parallel to the railway line. The second region is the Riau islands which lie to the south of Singapore. Hinterland transportation is organized almost entirely by barging. The third region is within the country of Singapore. As with the Johor region, hinterland cargo is transported by road haulage. Although Singapore is largely a transshipment hub, there is a significant amount of containers handled in Singapore connected to its local hinterland, totalling over 4.5 million TEUs. The handling of hinterland cargo traffic is facilitated by an effective inland transportation system that consists largely of road networks. However, the port's biggest constraint is land scarcity (Yap and Lam 2013). As a small island state without extensive freight corridors to neighbour countries, there is capacity limitation for Singapore's growth in inland connectivity. So far inland connections are developed mainly through natural evolution. Hence, Singapore can be considered as a 4GP in terms of inland connectivity.

Lastly, Singapore has good seaward connectivity in terms of shipping services deployed. Waterside connectivity is a key performance indicator for ports and shipping lines in global supply chains (Lam 2011). The port of Singapore is a focal point for some 200 shipping lines with links to about 600 ports in over 120 countries worldwide (MPA 2013c). Thus the port is a vital node for numerous links with substantial volumes connected to various forelands. The port has continuously pursued a strategy of common user terminals since containers were first handled in the 1970s. However, PSA has diverged from this long standing position by offering dedicated handling facilities through joint venture terminals with MSC, COSCO, and PIL. Hence, the port is able to anchor major shipping lines without compromising on its ability to continue to bring about strong connectivity to other regions in the world. Overall, Singapore is likely to maintain its strategy as a major transshipment hub. As such, we consider that Singapore has done well in waterside connection, and is somewhere between 4GP and 5GP.

4.3.5 Comparison of Service Quality of the Four Ports [1]

This section aims to evaluate service quality of the four ports to justify their generation port status, referring to Hu and Lee (2010, 2011), and Lee and Hu (2012). Major Asian container ports face fierce inter-port competition to capture transshipment cargoes due to development of short sea shipping networks as well as gateway cargoes arriving from the hinterland. Therefore, port authorities and terminal operators have to pay more attention to regularly evaluate their port service quality and to compare it with the service quality of competitors to satisfy the needs of shippers and ship operators with multi-dimensional constructs of port-service quality.

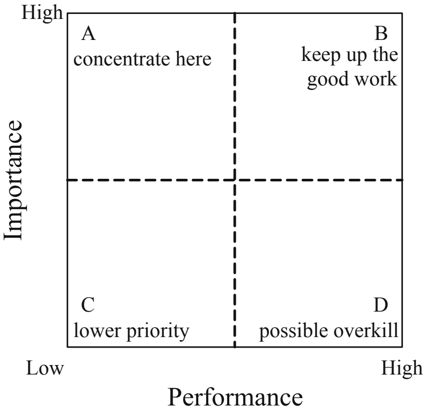
Hu and Lee (2010) evaluated port service quality of major container ports in Asia, i.e. Singapore, Hong Kong, Busan, Shanghai, and Kaohsiung, applying the Kano model. According to the scope of the case study in this chapter, Kaohsiung test results were excluded from this section. However, Kano's model has no function to prioritise service quality attributes even though it has been widely used to identify categories of service quality attributes and can offer useful information about developing improvement strategies of service quality (Löfgren and Witell 2008). Therefore, Hu and Lee's (2010) findings are limited to provide terminal operators with more concrete managerial and operational alternatives and to help them develop strategies of port service quality improvement. To solve such drawbacks, Hu and Lee (2011) conducted the importance-performance analysis (IPA), which was developed by Martilla and James (1977) and later refined by Jemmasi et al. (1994), recognising it as a useful tool for prioritizing attributes. The IPA technique can prescribe prioritisation of attributes for improvement and provide guidance for strategy formulations by exploring the competitive situation that port operators face, confirming the opportunity of improvement, and indicating the strategic direction in a four-quadrant matrix (Hawes and Rao 1985; Myers 2001; Sampson and Showalter 1999).

IPA uses the mean of importance and performance as coordinates for plotting each attribute on a two-dimensional matrix (see Fig. 4.2). A manager should focus on the position of each attribute in the four quadrant boundaries of IPA matrix that show the relative urgency of improvement. The top-left quadrant A means that attribute has high importance and low performance, thus this attribute has a higher priority to be improved and the manager needs to "concentrate here". The bottom-left quadrant C means the attribute has low importance and low performance, and then the attribute has the "lower priority" of improvement. The top-right quadrant B indicates that

---

[1] This section is based on Hu and Lee (2010), Hu and Lee (2011), and Lee and Hu (2012).

Fig. 4.2 Importance-performance analysis matrix



the attribute has high importance and high performance, thus the attribute has an excellent performance relative to its importance and the manager needs to "keep up the good work". The bottom-right quadrant D indicates that the attribute has low importance and high performance, thus this attribute may be a "possible overkill". From the viewpoint of efficiency of resource allocation, resources should be shifted from quadrant D to quadrant A to improve overall performance of service attributes.

The 19 attributes in Table 4.6 are classified into tangibles, reliability, responsiveness, assurance and empathy based on a comprehensive literature review done by Hu and Lee (2010). It aims to compare satisfaction level of service quality attributes among the four major ports in Asia and to draw operational and managerial strategy to improve quality service at each port.

The results of satisfaction of these four ports are listed in Table 4.7. Singapore port is the most satisfied port that rated by container shipping lines. Shanghai port has lower satisfaction in terms of responsiveness and assurance than the other ports. As for service reliability, responsiveness, and service assurance, Shanghai is the worst port among the four ports. On the contrary, the port has good satisfaction outcome in terms of tangibles compared to Busan. Total satisfaction of Shanghai and Busan port, both at 3.84, is lower than its average value (3.93).

Further close analysis on Table 4.8 enables each port operator to identify its weaknesses and strengths compared to the competing ports via IPA analysis. One of advantages of the IPA matrixes can help port manager to identify the position of each attribute in the four quadrant boundaries of IPA matrix and help them develop strategies to upgrade from 4GP to 5GP, referring to relative urgency of improvement among the attributes under considerations. According to the results of IPA, there is no big difference among the four ports. Most attributes are located at the right side in IPA matrixes, i.e. in quadrants B and D. Table 4.8 shows attributes of each port to be considered by port managers for moving up from 4GP to 5GP.

Most of the ports had port location (2), security procedures (7), communication system (8) and settlement of accident claims (10) in quadrant B (keep up the good work), except for Shanghai, where accident claims (10) were missing in quadrant

Table 4.6 Port service quality attributes. (Source: Quoted from Hu and Lee (2010))

Attribute No	Attribute
1	Cargo handling facilities at this port are well maintained and properly working
2	This port is well located to capture container cargo with feeder shipping networking
3	Port congestion is not from time to time faced at this port
4	On-dock operation of containers at this port is efficient
5	Ready information of port-related activities (e.g. port MIS, EDI) contributes to saving ship's time costs and increasing a port user's utility
6	This port has well harmonized and integrated system among Customs Office, Port Authority and Quarantine Office to avoid waste of time of a port user
7	Container cargo security procedures requested by international organization and USA is efficient enough not to deteriorate port productivity
8	This port has good communication system with port users on the port service and operation
9	This port takes and replies to a port user's opinions and requirements promptly
10	The settlement of accident claims in this port is amicably made without lingering process
11	This port has transparent system in pricing negotiation and administrative process
12	This port conducts survey and questionnaire to see a port user's satisfaction on the port services
13	This port lives up to its service promises at the time
14	The port personnel have required skills and knowledge for better port services
15	This port is concerned with increasing value-added of a port user
16	This port doesn't have bureaucratic aspect in dealing with cargo claims and a port user's needs
17	The employees in this port provide a port user with trustworthy and reliable services
18	This port has a port user's oriented policy and reviews it timely responding to market changes
19	The employees at this port can understand a port user's specific needs and properly respond to it

B. Busan also had port congestion (3) in this quadrant implying that these attributes performed well in proportion to their importance. Therefore, managers just have to 'keep up the good work' regarding those attributes. The transparent system in pricing negotiation and administrative process (11) and the bureaucratic aspect in dealing with cargo claims (16) were located in quadrant C (lower priority), except for Singapore, where attribute 11 was missing. This implies that shipping lines calling the four ports do not perceive these attributes as very important contradicting complaints on that issue from most international business companies in developing countries and thus requires further study through face-to-face interviews with managers. All other attributes were located in the bottom-right quadrant D (possible

Table 4.7 Results of service satisfaction at four major ports in Asia. (Source: Table 4.2 in Lee and Hu (2012, p. 205) and modified by the authors)

Dimension	Attribute No	Singapore	Shanghai	Hong Kong	Busan	Average
Tangibles	1	4.76	4.44	4.32	4.20	4.43
	2	4.92	3.84	4.44	3.60	4.20
Reliability	3	3.24	3.28	3.28	3.92	3.43
	4	4.44	4.16	4.12	3.88	4.15
	5	4.64	3.92	4.60	4.12	4.32
Responsiveness	6	4.60	3.96	4.52	4.20	4.32
	7	4.52	4.04	4.32	4.24	4.28
	8	4.56	4.00	4.40	4.04	4.25
	9	4.32	3.60	4.04	3.76	3.93
	10	4.00	3.28	3.84	3.60	3.68
	11	3.80	3.32	3.36	3.40	3.47
	12	3.36	2.92	3.40	3.16	3.21
Assurance	13	4.36	3.88	4.16	3.76	4.04
	14	4.40	3.72	4.32	4.00	4.11
	15	4.24	3.64	4.04	3.88	3.95
	16	2.88	2.60	2.64	2.52	2.66
	17	4.56	3.84	4.60	4.04	4.26
Empathy	18	4.36	3.72	4.04	3.80	3.98
	19	4.52	3.88	4.32	3.76	4.12
Total Satisfaction		4.64	3.84	4.28	3.84	3.93

Full score is 5

Table 4.8 Attributes to be considered by port managers for moving up to fifth generation port. (Source: Lee and Hu (2012, p. 208))

	Quadrant A: Concentrate here	Quadrant B: Keep up the good work	Quadrant C: Lower priority	Quadrant D: Possible overkill
Singapore	3, 12	2, 7, 8, 10	16	1, 4, 5, 6, 9, 11, 13, 14, 15, 17, 18, 19
Shanghai	3, 10, 12	2, 7, 8	11, 16	1, 4, 5, 6, 9, 13, 14, 15, 17, 18, 19
Hong Kong	3, 12	2, 7, 8, 10	11, 16	1, 4, 5, 6, 9, 13, 14, 15, 17, 18, 19
Busan	12	2, 3, 7, 8, 10	11, 16	1, 4, 5, 6, 9, 13, 14, 15, 17, 18, 19

overkill). For example, attributes 17 ('the employees in this port provide a port user with trustworthy and reliable services'), 9 ('prompt replies to container port user's opinions and requirements'), and 14 ('port personnel's skill and knowledge for better port service') had lower importance levels and lower service quality (satisfaction) compared to the attributes in quadrant C. This indicates that the shipping lines currently rate these attributes with relatively low importance but the four ports perform very well by meeting customers' satisfaction or even exceeding their requirements. This would be attributed to the world-class standard in general achieved by the four ports. In view of fierce competition, the ports are expected to continuously improve their performance. For efficient allocation of resources for improving the more important attributes, the ports may not have to over-emphasize quadrant D but would be appropriate to maintain at least an average rate of customer satisfaction.

This section has evaluated port service quality at the four container ports by testing service satisfaction of shipping liners calling at the four ports and applying IPA to judge generation port level with some meaningful significance. Singapore and Hong Kong Port provide the highest satisfaction to shipping liners. First, this study results enable us to recognise importance and satisfaction levels of each attribute from perspectives of the container shipping lines. Second, the results of IPA can help port managers draw managerial and operational implications for moving up from 4GP to 5GP by improving their service quality as well as meeting port users' needs. Having considered the satisfaction results of the four ports and attributes in IPA analysis, Singapore and Hong Kong provide satisfactory service to shipping liners, while Shanghai and Busan offer relatively lower satisfaction to them.

4.4 Cooperation Among Asian Ports?

Based on the above case studies, it can be seen that these Asian container ports are confronted with keen competition from other ports, especially those neighbouring ports. This is because neighbouring ports are competitive in nature, especially those that share a common hinterland or transhipment market and aim to attract higher cargo throughput. Would there be any rooms for cooperation among Asian ports? Cooperation is an agreement in which two or more ports jointly participate in mutually beneficial operations. Compared to port competition, the research area of port cooperation has been less investigated in the literature. However, ports engaging in intense competition with rival ports at the local and regional levels, coupled with the increasing bargaining power of customers, i.e. shipping companies, actually face a squeeze on their market power (Song 2002). Shipping lines as container ports' major customers are involved in mergers and strategic alliances to gain economies of scale, yet ports intensely compete against each other and are fragmented in their commercial actions such as pricing and terminal expansion projects. This has led to a weaker collective market power for the container port industry against its customers.

To counter such a trend from the micro level, port cooperation or port complementarity as a strategic option should be considered by ports in the current globalised market (Yap and Lam 2004; Wang et al. 2012). The cooperation approach embodies the sharing of resources among ports and creates a form of connection for the participating ports to operate in a consistent and coherent manner which assists them to achieve their main objective, that is to create a sustainable competitive advantage. For example, Song (2002) conducts a study on the case of Hong Kong and South Chinese ports, illustrating the collective benefits that can be reaped for the ports in the South China region, through cross investments made in the ports of Hong Kong and Shenzhen and by the various terminals operators in the region, such as Hutchison Port Holdings (HPH) Group and MTL. Equity joint ventures allow firms to gain market power without losing their flexibility in core operations to compete globally (Parola et al. 2013), hence would be an appropriate form of collaborative arrangement for global terminal operators. From the point of view of ship calls, the decision by liner services to call at particular ports can be influenced by joint competitive offers of a group of ports in the PRD instead of one single entity (Lam and Yap 2011a). Hence ports and terminal operators could adopt a long term view of achieving a win-win situation through cooperative agreements.

From the macro level, port authorities and governments can cooperate at a national or regional scale for ports' mutual benefits. Cooperative agreements may be drawn between two or more ports in which parties agree to cooperate in matters related to maritime trade and port developments (Lam et al. 2014). For instance, through exchanging of information in port infrastructure developments, staff training, and industry best practices in port operations and management, ports can benefit from each other's experiences and improve their own competitive advantage. Possible trade promotion between parties may also lead to higher cargo throughputs for these collaborating ports. In recent years, there are some favourable developments generating higher port complementarity in the South China region. The Framework Agreement on Hong Kong/Guangdong Co-operation (2010) was signed between the Hong Kong Special Administrative Region and Guangdong Provincial Governments with the aim to outline concrete measures for long-term development positions on Hong Kong and Guangdong cooperation. Thus there should be higher integration within the PRD ports towards establishing a port cluster in the PRD, with its port functions complementary to each other. This is a positive movement for the port of Hong Kong in terms of port cluster development in the port generation status. In other words, more cooperative efforts with the PRD are helpful for enhancing Hong Kong's status as a 5GP.

Other than economic cooperation at the macro level, port authorities and governments can also work together for non-commercial initiatives such as enhancing environmental performance and maritime security. For example, Singapore partnered with neighbouring countries including Malaysia and Indonesia in the Joint Oil Spill Exercise 2006, which was organized by the Maritime and Port Authority of Singapore (MPA 2006). Joint forces can smoothen the implementation of contingency plans if an extensive oil spill caused by ship collision happens leading to water contamination across neighbouring countries. Port operators will be more prepared and can effectively exercise reactive measures to mitigate the impact of water pollution. Such kind of collaboration will facilitate the stabilisation of the operations conditions and environmental quality of the ports, thus contribute to several aspects of 5GP including service quality, community environmental impact, port cluster performance and others.

Nevertheless, inter-port relationships are often complex and dynamic. There are various factors and conditions affecting cooperation and competition among regional ports (Wang et al. 2012). First, ports cooperate for mutual interests. The benefits of cooperation should be sufficient to motivate the involved ports' decision makers. Benefits should outweigh the overall costs of starting, executing, and maintaining the cooperation. This also means that a cost and benefit analysis should be performed in order to understand if a cooperation is justifiable. Second, port cooperation involves multiple stakeholders who have diverse interests. There is a need to engage in well-structured negotiation and discussion for formulating a cost and benefit allocation mechanism acceptable to the institutional environments in the ports of concern. Third, some cooperative agreements may involve ports in different countries. Hence political issues will inevitably determine, or at least influence, the feasibility and format of port cooperation. The evolving and multi-dimensional relationships among ports make port cooperation a challenge. However, port cooperation is in need in the era of 5GP which embraces the concept of clustering and community impact. Therefore, more efforts are encouraged to investigate this emerging and relatively under-researched area in the future.

4.5 Discussions and Concluding Remark

The literature of competitiveness of container ports[2] applied a number of methods, among others, including time series analysis, DEA and SFA methods, service quality analysis with IPA and Kano model, multi-criteria evaluation, survey of container ship operators and logistics managers, shift-share analysis and diversification indexes such as Herfindahl–Hirschmann, marginal cost pricing approach, and game theory (e.g., Lombaerde and Verbeke 1989; Haralambides 2002; Ha 2003; Cullinane et al. 2004; Yap and Lam 2006; Cullinane et al. 2004; Zan 1999; Song 2002; Anderson et al. 2008; Hu and Lee 2010; Hu and Lee 2011; Lee and Hu 2012; Ishii et al. 2013). The above studies aim to measure the relative competitiveness of the ports in Asia and Europe. Such analysis results are useful to evaluate major ports at the given evaluation angle at the given time. However, these methods do not consider competitiveness analyses of container port in relation to its devolution according to a globalised economy with the change of production and distribution channels, technology, city-port interface, government policy, port users' behavior, port stakeholders' needs, port pricing, IT development, environmental issue, security and safety. This chapter attempted to evaluate port competitive edge of major Asian container ports, i.e. Busan, Hong Kong, Shanghai, and Singapore, referring to the customer-centric

---

[2] On the literature information of port competition studies, see Chang and Lee (2007) and Pallis et al (2010).

community ports, so-called the Fifth Generation Ports (5GP) proposed by Flynn et al. (2011). However, we modified the concept of the 5GP partly because the description of the factors is to some extent vague and does not have a sharp demarcation line among some items for the sake of comparison and evaluation of port generation. Furthermore, some items, e.g., the port and maritime clusters and the logistics hub and inland are not independent function to achieve the core of the customer-centric community ports as indicated in Table 4.2. This chapter applied the modified concept of 5GP to evaluate inter-port competitiveness of the four major container ports in Asia in a comprehensive way to reflect the above from cross-sectional, longitudinal and horizontal aspects of the port evolution. In this regard, it can be said that this novel approach with empirical test by a descriptive method plus partly quantitative approach contributes to port competition studies in the literature.

Table 4.9 summarises the evaluation results of the port generation status of the four ports, referring to the eight items of 4GP and 5GP as shown in Table 4.2 in section 2 above: service quality, information technology (IT), community environmental impact, port cluster, maritime cluster, logistics hub, inland connection, and waterside connection.

The case studies compared the situations of port generations of the four major container ports in Asia. Our research has revealed the similarities and differences among the four ports' status and approaches with regards to customer- and community-centric port development. In general, the ports are no longer entirely 4GP. In other words, leading container ports are evolving and progressing towards a higher hierarchy in the port ladder. In a volatile business environment, the dynamics between a port and its client base become much more complex. To innovate and improve indeed helps ports stay competitive. Among the eight items for analyzing port generation status of the four ports, IT has reached the most advanced stage with three ports being classified as 5GP, except for Shanghai. In the era of globalization and supply chain management, high-end IT solutions meeting customer and community's sophisticated and diversified demands are indispensable. The common feature among the 5GP is the development of integrated IT solutions. Korea and Singapore are among the leaders in seaport electronic information system. Both systems of the two nations are good at the 'Single Window', 'Integrated Function', and 'Cargo Tracking System' so that at anytime and anyplace, port users can obtain data such as vessel arrival/departure and overall information. These attributes would be good reference for Shanghai and other ports to further develop their IT solutions.

In view of the above case analysis, we conclude that Singapore Port has transformed into a 5GP in many of the aspects and hence can be treated as a pioneer in the 5GP model. The only aspect out of the eight assessment items that is considered staying at the 4GP status is inland connectivity. As compared with major gateway ports such as Shanghai, massive intermodal connection is not really applicable to Singapore due to the small size of the city state and the port's focus on transshipment traffic. On the contrary, inland connectivity is a major strength of Shanghai Port. This is the only item that Shanghai is rated as towards a 5GP. The findings draw managerial insights that the ports have evolved at different pace for various items largely to

Table 4.9 Evaluation status and comparison of the ports generation of the four ports. (Source: The authors)

Items	Busan	Hong Kong	Shanghai	Singapore
Service quality	Forwarding to Fifth Generation most efficient container port, providing comprehensive services to ships but some service attributes requiring improvements	Fifth Generation, one of the world's best ports satisfied by shipping liners, providing comprehensive services to ships	Fourth Generation, gaining lowest service satisfaction from the port users	Fifth Generation, one of the world's best ports satisfied by shipping liners, providing comprehensive services to ships
Information Technology (IT)	Fifth Generation, integrated IT solutions	Fifth Generation, integrated IT solutions	Fourth Generation, no single window system connected with Chinese container ports	Fifth Generation, integrated IT solutions
Community environmental impact	Fourth Generation, Busan Port Authority (BPA) having limited authority to implement decisions on the community impact	Towards Fifth Generation, largely led by the maritime industry	Fourth Generation, focusing on regulatory compliance	Towards Fifth Generation, largely led by Maritime and Port Authority
Port cluster	Fourth Generation, Busan Port Authority with limited autonomy to use land	Fourth Generation, government as land use Fourth Generation, government as land use planner	Fourth Generation, government as land use planner	Fifth Generation, Maritime and Port Authority and PSA Corporation as port cluster developers
Maritime cluster	Fourth Generation, offering neither FOC nor tax incentives for maritime business	Fifth Generation, offering advanced maritime services, and the fourth largest shipping register in the world	Fourth Generation, developing maritime services expertise	Fifth Generation, offering advanced maritime services, and the fifth largest shipping register in the world
Logistics hub	Fourth Generation, Korea's Logistics Performance Index LPI is relatively low at 21st rank in the world but moving forward to Fifth Generation under the Government's Master Plan	Fifth Generation, the whole of Hong Kong is a free port, ranked second in World Bank's LPI	Fourth Generation, China's LPI is relatively low at 26th, but highly ranked as a developing country	Fifth Generation, ranked first in World Bank's LPI

Table 4.9 (continued)

Items	Busan	Hong Kong	Shanghai	Singapore
Inland	Fourth Generation hindered by North Korea	Fourth Generation, hindered by congested and inefficient cross-border hinterland transport	Towards Fifth Generation, extensive transport network penetrating the vast hinterland in the Yangtze River Delta and central China	Fourth Generation, small island state without extensive freight corridors to neighbour countries
Waterside	Moving forward to Fifth Generation, focusing on transshipment cargoes with proactive port pricing	Fourth Generation, extensive connectivity to forelands with growing transshipment traffic	Fourth Generation, developing a coastal feeder network in its Northeast Asia Strategy	Somewhere Fourth to Fifth Generation, focusing on common user strategy as a major transshipment hub

achieve strategic fit of the port’s major clientele. This is virtually a customer-centric approach in enhancing the port’s competitive advantage.

Referring to Table 4.9, Hong Kong has achieved or is towards the 5GP status in five items. As discussed before, a major weakness is inland connectivity with mainland China. This is hindered by congested and inefficient cross-border hinterland transport which has existed for long time. The Hong Kong SAR government realised the issue and the Hong Kong–Zhuhai–Macau Bridge is being constructed to improve the situation. Therefore, even in principle the Government adopts a laissez-faire policy, it does step up to facilitate strategic directives for strengthening the long term competitiveness of the port and maritime industries.

As for Busan, the port is more of a 4GP than a 5GP based on its current conditions. Information technology is a 5GP item, whilst service quality and waterside strategy are moving towards a 5GP status. Among the five items which are classified as 4GP, the relatively more essential items for enhancement would be the development of port and maritime clusters, as well as improving Busan’s logistics performance. However, active port-city interface among port stakeholders is found in the course of waterfront development in Busan Port. It is a good signal to raise the bar to 5GP in the terms of ‘Community environmental impact’ shown in Table 4.2.

The concept of 5GP is required to explicitly consider green port priorities to assist with gas emissions reductions and waterfront development to harmonize with urban planning, although the Port Ladder indicates the environmental issue (Flynn et al. 2011). The above two issues are critical for further developing port generations. In addition, the literature of port generations points out that the concept of 4GP is subjective. It implies that the concept of 5GP drawn from the features of 4GP would be subjective. Having said that, we need to further expand the concept of the 5GP model to make it more sophisticated and quantifiable, responding to port evolution. This chapter is an original attempt for analysing port generation status by combining a descriptive approach and a quantitative method. Therefore, future research has first to develop a more sophisticated concept of 5GP and then to evaluate generation status of container ports in an objective and quantifiable way taking more cases of container hub port in Asia and Europe. Future research can also further investigate the topic of port cooperation, for example, in terms of its interplay with port competition and competitiveness in the 5GP model.

Acknowledgement The study is partly supported by Singapore MPLP project, Nanyang Technological University ref. M4061473.

References

Anderson, C. M., Park, Y-A, Chang, Y-T, Yang, C-H, Lee, T-W., & Luo, M. (2008). A game-theoretic analysis of competition among container port hubs: the case of Busan and Shanghai. *Maritime Policy & Management, 35*(1), 5–26.

Bennathan, E., & Walters, A. A. (1979). *Port Pricing and Investment Policy for Developing Countries*. London: Oxford University Press.

BPA, Busan Port Authority. (2013). 2012 Port of Busan Container Statistics, Busan: Busan Port Authority.

Brown, P. H. (2008). *America's waterfront revival: Port authorities and urban redevelopment*. University of Pennsylvania Press.

Chang, Y. T., & Lee, P. T-W. (2007). Overview of interport competition: Issues and methods. *Journal of International Logistics and Trade, 5*(1), 99–121.

Channel news Asia. (2013). Singapore highly regarded as logistics hub. http://www.channelnewsasia.com/news/singapore/singapore-highly-regarded/849440.html. Accessed 31 Oct 2013.

Charan, R. (2007). *What the customer wants you to know how everybody needs to think differently about sales*. New York: Penguin Books.

Charlier, J. (1992). The regeneration of old port areas for new port uses. In B. S. Hoyle & D. A. Pinder (Eds.), *European port cities in transition* (pp. 137–154). London: Belhaven Press.

China ACC. (2011). http://www.chinaacc.com/new/184_900_201112/20zh880717040.shtml. Accessed 3 Nov 2013

Craig-Smith, S. J., & Fagence, M. (1995). *Recreation and tourism as a catalyst for urban waterfront redevelopment: An international survey. Praeger Publishers*.

Cullinane, K., & Song, D. W. (Ed). (2007). Asian *Container Ports: Development, Competition and Co-operation*. New York: Palgrave Macmillan.

Cullinane, K., & Wang, T. F. (2007). Data envelopment analysis (DEA) and improving container port efficiency, *Research in Transportation Economics, 17,* 517–566.

Cullinane, K., Song D. W., Ji, P., & Wang, T. F. (2004). An Application of DEA Windows Analysis to Container Port Production Efficiency*, Review of Network Economics*, *3*(2), 1–23.

Enhancements to Maritime Singapore Green Initiative, Annex A 2013. http://www.mpa.gov.sg/sites/pdf/annex%20a%20-%20Enhancements%20to%20Maritime%20Singapore%20Green%20Initiative.pdf. Accessed 31 Oct 2013.

Flor, L., & Defilippi, E. (2003). Port infrastructure: An access model for the essential facility. *Maritime Economics & Logistics, 5*(2), 116–132.

Flynn, M., & Lee, P. T-W. (2010). Customer-Centric and Community Ports as the Next Step on the Port Ladder: The Fifth Generation Ports, The 8th Asia Pacific Transportation Development Conference, May 27–30, 2010, Tainan, Taiwan.

Flynn, M., Lee, P.T-W., & Notteboom, T. (2011). The next step on the port generations ladder: customer-centric and community ports. In T. Notteboom (Ed.), *Current Issues in Shipping, Ports and Logistics*, (pp. 497–510) Brussels: University Press Antwerp.

Galbraith, V., Curry, L., & Loh, C. 2008. Green Habours: Hong Kong & Shenzhen, Civic Exchange, Hong Kong.

Gilmans, S. (2003). Sustainability and national policy in UK port development. *Maritime Policy & Management, 30*(4), 275–291.

Gordon, D. L. A. (1993). Architecture: How not to build a city - implementation in Battery Park City. *Landscape and Urban Planning, 26*, 35–54.

Gordon, D. L. A. (2004). *Implementing urban waterfront redevelopment, in urban land institute. Remaking the urban waterfront* (pp. 80–89). Washington DC: Urban Land Institute.

Grobar, L. M. (2008). The economic status of areas surrounding major US container ports: Evidence and policy issues. *Growth and Change*, 39(3), 497–516.

Ha, M. S. (2003). A comparison of service quality at major container ports: implications for Korean ports. *Journal of Transport Geography, 11*(2), 131–137.

Haralambides, H. (2002). Competition, excess capacity, and the pricing of port infrastructure. *International Journal of Maritime Economics, 4,* 323–347.

Hawes, J. M., & Rao, C. P. (1985). Using Importance-performance Analysis to Develop Health Care Marketing Strategies. *Journal of Health Care Marketing, 5*(4), 19–25.

Hayuth, Y. (1982). The port-urban interface: An area in transition. *Area, 14*(3), 219–224.

Hong Kong Port Development Council (HKPDC). 2013. www.pdc.gov.hk. Accessed 16 Oct 2013.

Hong Kong Trade Development Council (HKTDC). 2013. Sea Transport Industry in Hong Kong. http://hong-kong-economy-research.hktdc.com/business-news/article/Hong-Kong-Industry-Profiles/Sea-Transport-Industry-in-Hong-Kong/hkip/en/1/1X000000/1X0018WG.htm. Accessed 15 Aug 2013.

Hong Kong—Zhuhai—Macao Bridge (HZMB). (2013). http://www.hzmb.hk/eng/about_overview_01.html. Accessed 2 Oct 2013.

Hoyle, B. S. (1988). Development dynamics at the port-city interface. In B. S. Hoyle, D. A. Pinder, & M. S. Husain (Eds.), *Revitalising the waterfront: International dimensions of dockland redevelopment* (pp. 3–19). London: Belhaven Press.

Hoyle, B. S. (1989). The port—City interface: Trends, problems and examples. *Geoforum, 20*(4), 429–435.

Hoyle, B.S. (1998). Development dynamics at the port-city interface. In B. S. Hoyle, D. A. Pinder, & M. S. Hussain (Eds.), *Revitalising the waterfront* (pp. 3–19). London: Belhaven Press.

Hoyle, B. (2000). Global and local change on the port-city waterfront. *Geographical Review, 90*(3), 395–417.

Hoyle, B. S., & Hilling, D. (Eds.). (1984). *Seaport systems and spatial change, technology, industry and development strategies*. Chichester: Willey.

Hu, K. C., & Lee, P. T-W. (2010). The Service Quality Attributes of Major Container Ports in Asia, Incheon Port Policy Forum & 2010 the International Conference, 169–185.

Hu, K-C., & Lee, P. T-W. (2011). Developing a New Technique for Evaluating Service Quality of Container Ports, (pp. 337–356). In T. Notteboom (Ed.), *Current Issues in Shipping, Ports and Logistics*, Brussels: University Press Antwerp.

Ishii, M., Lee, P. T. W., Tezuka, K., & Chang, Y. T. (2013). A game theoretical analysis of port competition. *Transportation Research Part E: Logistics and Transportation Review, 49*(1), 92–106.

IMO. 2012. Status of conventions as at 29 February 2012. http://www.imo.org/About/Conventions/StatusOfConventions/Pages/Default.aspx.

Jemmasi, M., Strong, K. C., & Taylor, S. A. (1994). Measuring service quality for strategies planning and analysis in service firms. *Journal of Applied Business Research, 10*(4), 24–34.

JIFFA, Japan International Freight Forwarders Association Inc. (2014). Shanghai Retains Its Top Position in World's Box Port Ranking. http://www.jiffa.or.jp/en/news/entry-2691.html. Accessed 24 Mar 2014.

Kano, N., Seraku, N., Takahashi, F., & Tsuji, S. (1984). Attractive Quality and Must-be Quality. *The Journal of the Japanese Society for Quality control, 14*(2), 39–48.

KIC (Kaesong Industrial District Management Committee). bhttp://www.kidmac.com/. Accessed 1 Nov 2013.

KMI, Korea Maritime Institute (Seoul) and Soyuzmorniiproekt (Moscow, Russia). (1992). Sea Transportation between Russia and Korea, Prospects for the Development of Trans-Siberian Container Service, Policy Research Report 056. Korea Maritime Institute: Seoul.

Lam, J. S. L. (2011). Patterns of maritime supply chains: slot capacity analysis. *Journal of Transport Geography, 19*(2), 366–374.

Lam, J. S. L., & Notteboom, T. (2014). The greening of ports: a comparison of port management tools used by leading ports in Asia and Europe. *Transport Reviews,* DOI:10.1080/01441647.2014.891162.

Lam, J. S. L., & Song, D. W. (2013). Seaport network performance measurement in the context of global freight supply chains, *Polish Maritime Research, 20,* 47–54.

Lam, J. S. L., & Yap, W. Y. (2011a). Container port competition and complementarity in supply chain systems: evidence from the Pearl River Delta. *Maritime Economics and Logistics, 13*(2), 102–120.

Lam, J. S. L., & Yap, W. Y. (2011b). Dynamics of liner shipping network and port connectivity in supply chain systems: analysis on East Asia. *Journal of Transport Geography, 19*(6), 1272–1281.

Lam, J. S. L., Ng, A. K. Y., & Fu, X. (2013). Stakeholder management for establishing sustainable regional port governance. *Research in Transportation Business & Management, 8,* 30–38.

Lee, T-W. (1993). A Study on Development of Linkages of Ports and Internal Transport between Russia, China and Korea towards the 21st Century, *Journal of Korean Association of Shipping Studies, 16,* 81–131.

Lee, P. T-W., & Flynn, M. (2011). Charting A New Paradigm of Container Port Development Policy: The Asian Doctrine, *Transport Reviews, 31*(6), 791–806.

Lee, P. T-W., & Hu, K-C. (2012). Evaluation of the service quality of container ports by importance-performance analysis. *International J of Shipping and Transport Logistics, 4*(3), 197–211. doi:10.1504/IJSTL.2012.047479.

Lee, P. T-W., & Lee, T-C. (2012). A New Lease Charging System for Busan Container Terminals: A Historical Case Study. *Maritime Policy & Management, 39*(1), 91–105.

Lee, T-W, Park, N-K, Joint, J. F., & Kim, W-K. (2000). A new efficient EDI system for container cargo logistics, *Maritime Policy & Management, 27*(2), 133–144.

Lee, P. T-W, Wu, J-Z, Suthiwartnarueput, K., Hu, K-C, & Rodjanapradied, R. (2013). A Comparative Study of Key Critical Factors of Waterfront Port Development: Case Studies of Incheon Inner Port and Bangkok Port, Proceedings of 2013 Asian Logistics Round Table Conference, 2–4 April, Chulalongkorn University, Bangkok, Thailand. (Forthcoming in an international journal Growth and Change).

Levesque, P. (2011). Maritime Industry: Today and Tomorrow. Asian Logistics and Maritime Conference [pdf]. http://www.asianlogisticsconference.com/download/speeches/peter_levesque_speech.pdf. Accessed 10 Aug 2013.

Li, Yuehua. (2011). Current logistics situation of container transport in Shanghai Port, http://www.fhnw.ch/wirtschaft/dienstleistung/studierendenprojekte/olten/bisherige-projekte/bachelor-thesis-diplomarbeiten-2011/pdf/port. Accessed 3 Nov 2013.

Löfgren, M., & Witell, L. (2008). Two decades of using kano's theory of attractive quality: a literature review. *The Quality Management Journal, 15*(1), 59–75.

Lombaerde, P. D., & Verbeke, A. (1989). Assessing international seaport competition: A tool for strategic decision making. *International Journal of Transport Economics, 16*(2), 175–192.

Marine Department (MD). (December 2006). *Port Benchmarking for Assessing Hong Kong's Maritime Services and Associated Costs with other Major International Ports*, Hong Kong Marine Department.

Marine Department (MD). (2013). http://www.mardep.gov.hk/en/home.html. Hong Kong Marine Department. Accessed 18 Sept 2013.

Maritime Cluster Fund. (2013). http://www.mpa.gov.sg/mcf. Accessed 27 Oct 2013.

Maritime Singapore Green Initiative. (2013). http://www.mpa.gov.sg/sites/maritime_singapore//msgi/maritime-singapore-green-initiative.page. Accessed 31 Oct 2013.

Martilla, J. A., & James, J. C. (1977). Importance-Performance Analysis*, Journal of Marketing, 41*(1), 77–79.

McKinnon, A. (2011). HONG KONG AND SHANGHAI PORTS: Challenges, Opportunities and Global Competitiveness. http://www.cityu.edu.hk/slw/HKCMT/Doc/Working_Paper_4_-_Shanghai_-_Final_(v4).pdf. Accessed 3 Nov 2013.

MPA. (2006). Joint Oil Spill Exercise to Test Response Strategy. http://www.mpa.gov.sg/sites/global_navigation/news_center/mpa_news/mpa_news_detail.page?filename † =nr060929.xml Accessed 16 April 2014.

MPA. (2012). Singapore's 2012 Maritime Performance. http://www.mpa.gov.sg/sites/global_navigation/news_center/mpa_news/mpa_news_detail.page?filename†=nr130110.xml. Accessed 31 Oct 2013.

MPA. (2013a). E-services. http://www.mpa.gov.sg/sites/global_navigation/e-services/e_services.page. Accessed 31 Oct 2013.

MPA. (2013b). Maritime Sector. http://www.mpa.gov.sg/sites/global_navigation/news_center/mpa_news/mpa_news_detail.page?filename†=nr120928.xml. Accessed on 31 Oct 2013.

MPA. (2013c). http://www.mpa.gov.sg/sites/maritime_singapore/what_is_maritime_singapore/premier_hub_port.page. Accessed 27 Oct 2013.

Myers, J. (2001). Measuring Customer Satisfaction: Hot Buttons and Other Measurement Issues. Chicago: American Marketing Association.

OnePort. (2013). http://www.oneport.com/eng/overview.htm. Accessed 10 Sept 2013.

Paixão, A. C., & Marlow, P. B. (2003). Fourth generation ports–a question of agility? *International Journal of Physical Distribution & Logistics Management, 33*(4), 355–376.

Pallis, A. A., Vitsounis, T. K., & De Langen, P.W. (2010). Port Economics, Policy and Management: Review of an Emerging Research Field, *Transport Reviews, 30*(1), 115–161.

Parola F. Satta G., & Caschili S. 2013. Unveiling co-operative networks and 'hidden families' in the container port industry, *Maritime Policy & Management*. DOI:10.1080/03088839.2013.782442.

People's Republic of China (PRC). (2000). Marine Environment Protection Law. http://law.shmsa.gov.cn/smls/WebPages/Law/LawDetail.aspx?CatalogId † =15692& &Keyword = %E6%B5%B7%E6%B4%8B%E6%B1%A1%E6%9F%93 & & TreePath=100000020020 & & typelang = LangType. Accessed 3 Nov 2013.

PSA. (2013). PSA Products. http://www.portnet.com/WWWPublic/products.html. Accessed 31 Oct 2013.

Robinson, R. (1985). Industrial strategies and port development in developing countries: The Asian Case. *Tijdschrift voor Economische en Sociale Geografie, 76*, 133–143.

Sampson, S. E., & Showalter, M. J. (1999). The Performance-Importance Response Function: Observations and Implications. *The Service Industries Journal, 19*(3), 1–25.

Schiffer, J. R. (1991). State policy and economic growth: A note on the Hong Kong model. *International Journal of Urban and Regional Research, 15*(2), 180–196.

Shanghai International Port Group (SIPG). (2013). http://www.portshanghai.com.cn/en/. Accessed 2 Nov 2013.

Song, D. W. (2002). Regional container port competition and co-operation: the case of Hong Kong and South China. *Journal of Transport Geography, 10*(2), 99–110.

Song, B. L. 2007. Shanghai: a Newly & rapidly Developing Hub Port. In K. Cullinane, & D. W. Song. (Eds), *Asian Container Ports: Development, Competition and Co-operation* (pp. 11–33). New York: Palgrave Macmillan.

Teng, J. Y., Huang, W. C., & Huang, M.J. (2004). Multicriteria evaluation for port competitiveness of eight East Asia container ports. *Journal of Marine Science and Technology, 12*(4), 256–264.

The Hong Kong SAR Government and Guangdong Provincial Government. (2010). Framework Agreement on Hong Kong/Guangdong Co-operation, www.info.gov.hk. Accessed 10 April 2014.

The State Council of China. (2009). http://www.gov.cn/zwgk/2009–04/29/content_1299428.htm. Accessed on 1 Nov 2013.

Transport and Housing Bureau (THB). (September 2013). *Summary Statistics on Port Traffic of Hong Kong*. Hong Kong: Transport and Housing Bureau.

Transport Weekly. (2013). http://www.transportweekly.com/pages/en/news/articles/100889/. Accessed 31 Oct 2013.

UNCTAD. (1992). Port marketing and the challenge of the third generation port. Trade and development board committee on shipping ad hoc intergovernmental group of port experts, 13–23

UNCTAD. (1999). The fourth generation port. *UNCTAD Ports Newsletter, 19,* 9–12.

UNCTAD. (2012). Review of maritime transport. http://unctad.org/en/PublicationChapters/Chapter%202.pdf. Accessed 31 Oct 2013.

Wang, J. J., & Cheng, M. C. (2010). From a hub port city to a global supply chain management center: a case study of Hong Kong, *Journal of Transport Geography, 18*(1), 104–115.

Wang, K., Ng, A. K. Y., Lam, J. S. L., & Fu, X. (2012). Cooperation or competition? Factors and conditions affecting regional port governance in South China, *Maritime Economics and Logistics, 14*(3), 386–408.

Workforce Development Agency. (2012). WDA Port Services. http://www.wda.gov.sg/content/wdawebsite/L207 - AboutWSQ/L301 - WSQIndustryFramework-PortServices/L401–013PortServices.html. Accessed 31 Oct 2013.

World Bank. (2012). Logistics Performance Index. http://lpisurvey.worldbank.org/. Accessed 9 Sept 2013.

Yap, W. Y., & Lam, J. S. L. (2004). An Interpretation of Inter-Container Port Relationships from the Demand Perspective. *Maritime Policy & Management, 31*(4), 337–355.

Yap, W. Y., & Lam, J. S. L. (2006). Competition dynamics between container ports in East Asia. *Transportation Research Part A, 40*(1), 35–51.

Yap, W. Y., & Lam, J. S. L. (2013). 80 Million-TEU Container Port? Sustainability Issues in Port and Coastal Development. *Ocean and Coastal Management, 71,* 13–25.

Zan, Y. (1999). Analysis of container shipping port policy by the reaction of an equilibrium shipping market. *Maritime Policy & Management, 26*(4), 369–381.

Zhang, W., & Lam, J. S. L. (2013). Maritime cluster evolution based on symbiosis theory and Lotka-Volterra model. *Maritime Policy & Management, 40*(2), 161–176.

Zhu, L., & Agarwal, S. K. (2011). A Review of the Legal and Policy Framework for Vessel Source Pollution in Hong Kong. *Ocean Development & International Law, 42*(3), 264–279.

Chapter 5
Choosing a Business Model of Container Terminal Operations

Chin-Shan Lu and Pei-Hsuan Chang

Abstract This chapter empirically identifies the crucial criteria for choosing a business model of container terminal operations from the viewpoint of the Taiwan International Ports Corporation (TIPC)—the former port authority of Kaohsiung, Keelung, Taichung, and Hualien in Taiwan. An analytic hierarchy process approach is employed to assess the relative importance of these criteria and business models. Results indicate that benefit and operational capability are the two most important criteria for selecting business models of container terminal operations. The optimal business model in the short to medium term for TIPC is "wholly-owned, operate-outsourced", whereas "wholly-owned, operate-owned" is the optimal business model in the long term. Implications of the research findings for port corporations and global terminal operators are discussed.

5.1 The Role of Container Terminals in the Global Supply Chain

With the increased containerization of international trade and the development of transshipment, world container port throughput rose by a significant 3.8 % to 601.8 million TEUs in 2012 (UNCTAD 2013). A container terminal plays an important role in the global supply chain and provides an interface between sea and land transports (Fransoo and Lee 2013; Zhang et al. 2002). The major functions of a container terminal include loading and unloading containers on the quayside, and the temporary storage of containers dropped off by or to be delivered to inland shippers (Jin et al. 2014). The container terminal business is dominated by a few global container terminal operators. In terms of throughput, the top 10 global terminal operators accounted for 36 % of world throughput in 2012 (Drewry Maritime Research 2014). The five largest global container terminal operators are the Port of Singapore Authority (PSA),

---

C.-S. Lu (✉) · P.-H. Chang
Department of Logistics and Maritime Studies, C.Y. Tung International Centre for Maritime Studies, The Hong Kong Polytechnic University, Hung Hom, Hong Kong
e-mail: chin-shan.lu@polyu.edu.hk

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics*,
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_5

Hutchison Port Holdings (HPH), APM Terminals, Dubai Ports World (DP World), and COSCO Group. In particular, the efficiency of terminal operations is vital for the development of seaports. Governments and global terminal operators have shown strong interest in efficient business models of container terminal operations (Cheon et al. 2010).

With the globalization and institutional reforms in the port sector, public-private partnerships in port development have become a common business model in the last 25 years. The participation of the private sector (e.g. HPH, DP World, and APM Terminals) brings much needed capital and know-how, and is expected to lead to efficiency gains from combining construction, maintenance and operations arrangements (UNCTAD 2013). In addition, the creation of port corporations has brought many opportunities for operators based in one country to get involved in the container terminal operations in other countries. Notably, PSA remains one of the few terminal (corporations/operators?) still owned by its national government. Similarly, the Taiwan International Ports Corporation (TIPC), which was established on 1 March 2012 as a result of the port organization reform in Taiwan, is looking to expand its container terminal operations to international markets. TIPC is a government-owned organization and can adopt a wide range of policies that aim at increasing the operational efficiency of their port network. This includes decisions regarding the legal and institutional framework, the selection of an ownership model or the allocation of funds for infrastructure investment (UNCTAD 2013). In this chapter, we attempt to find answers to the following questions: What are the important criteria for choosing a business model for container terminals? What kinds of business models based on private sector participation, ownership, and corporatization are suitable for TIPC? What kinds of business models are optimal in the short to medium term, and the long term?" We will do so by examining the relative importance of the various criteria from the perspectives of port senior managers and academics. The findings will contribute to research on the theory and managerial practice of container terminal operations.

The majority of extant studies on terminal operations have focused on quayside operations (Chao and Lin 2011; Chen et al. 2012; Goodchild and Daganzo 2007), crane deployment and operations (Cheung et al. 2002; Kim and Bae 2004; Meisel and Bierwirth 2013; Vis and Carlo 2010; Zhang et al. 2002), as well as yard capacity and management (Jin et al. 2014; Kim and Kim 2002; Petering 2011; Taleb-Ibrahimi 1993; Zhang et al. 2003). Relatively little previous research has examined the criteria for choosing a business model of container terminal operations based on ownership, organizational management structure, and the period of implementation. Thus, the objective of this chapter is to investigate the criteria for choosing a business model of container terminal operations and examine the optimal business model in the short to medium term as well as the long term from a port corporation's perspective.

The remainder of this chapter is organized as follows. Section 2 comprehensively reviews previous studies and briefly delineates the major business models of container terminal operations in terms of ownership and concession of operation, as well as the criteria for choosing a business model. Section 3 describes the methodology including samples, methods used to assess the important criteria for choosing a business model

Table 5.1 The port ownership model. (Source: Alderton 2008)

Assets/operations	Public port	Tool port	Landlord port	Private port
Infrastructure (including port land)	Publicly owned	Publicly owned	Publicly owned	Privately owned
Superstructure	Publicly owned	Publicly owned	Privately owned	Privately owned
Stevedoring operations	Publicly operated	Privately operated	Privately operated	Privately operated
Other cargo handling operations	Mainly publicly operated	A mixture of publicly and privately operated	A mixture of publicly and privately operated	Mainly privately operated

of container terminal operations, and justifications for why some business models are feasible while others are not. Section 4 presents the results of analyses. Finally, Sect. 5 concludes with policy implications of the major business models of container terminal operations in the short to medium term, and the long term.

5.2 Literature Review

5.2.1 Business Models of Container Terminal Operations

Business models of container terminal operations can be categorized by such characteristics as public, private, or mixed provision of service, ownership of infrastructure, superstructure, equipment, and stevedoring operations. As can be seen in Table 5.1, a public port (e.g. the Port of Singapore) is one that is controlled by the government and the port authority owns, maintains, and operates assets, as well as conducts cargo handling activities. On the contrary, a private port (e.g. the Port of Felixstowe and Dublin Port) is one that is owned, operated and managed by private operators. In the tool port model, the port authority owns, develops, and maintains the infrastructure, superstructure, and cargo handling equipment, whereas stevedoring operations on the quay are usually carried out by private operators contracted by shipping carriers and licensed by the port authority. Finally, the landlord port is characterized by its mixed public-private services. The port authority leases infrastructure and superstructure out to private operators. The private operators are responsible for the provision of equipment and the loading/unloading operations. Examples of landlord ports include the Port of Kaohsiung, the Port of Rotterdam, and the Port of Antwerp (Alderton 2008).

Heaver et al. (2000) described the three basic roles of a port authority: conservator, facilitator, and entrepreneur. Verhoeven (2010) identified the four basic functions—namely landlord, regulator, operator and community manager—of a port authority based on a renaissance matrix which can be applied at the local, regional and global

Table 5.2 Classification of terminals

Terminals operated by public and for public use	General public terminals
Terminals operated by private companies and for dedicated use	Public terminals for priority berthing
	Terminals dedicated and rented out to shipping companies
	Terminals operated jointly by shipping companies
	Terminals constructed by the government and subsequently leased out to terminal operators
Terminals operated by private companies and for public use	Terminals constructed and operated by terminal operators
	Terminals constructed and operated by shipping companies
	Terminals constructed and operated through joint ventures between terminal operators and shipping companies

levels. A review of the relevant literature on the business models of container terminal operations (Baird 1995; Tongzon and Heng 2005; Chen 2009), the types of operators and the purpose of use of terminals shows that the terminals can basically be classified into "publicly operated and for public use", "privately operated and dedicated" and "privately operated and for public use." As shown in Table 5.2, two types of terminals are classified as "publicly operated and for public use" terminals. They are public terminals and public terminals for priority berthing. Three types of terminals are classified as "privately operated and dedicated" terminals. They are terminals dedicated to shipping companies, terminals jointly operated by shipping companies and terminals constructed by the government and leased out to terminal operators. There are also three types of "privately operated and for public use" terminals. They are terminals constructed and operated by terminal operators, terminals constructed and operated by shipping companies, and terminals constructed through joint ventures between terminal operators and shipping companies.

In general, public container terminals (have adopted/are adopting?) the traditional approach of queuing up for berths. A vessel arriving late at a public container terminal will be required to queue up again for a berth. This makes it more difficult for shipping companies to manage their sailing schedules. Therefore, ports that wish to attract large container carriers with vessels sailing on a regular schedule but have relatively few terminals tend to adopt the business model of container terminal for priority berthing (e.g. the Port of Keelung in Taiwan). Under this business model, the port authority and the shipping companies enter into an agreement which gives the latter priority access to the container terminals. The business model of leasing out container terminals for dedicated use is currently adopted in the Port of Kaohsiung. Under this business model, the port authority leases the container terminals and the affiliated facilities to container shipping companies through long-term tenancy agreements.

Table 5.3 Business models of container terminal operations

Operations Ownership		Operate-owned	Outsource-operated	Jointly-operated
Subsidiary company	Wholly-owned	Wholly-owned, Operate-owned	Wholly-owned, Outsource-operated	Wholly-owned, Jointly-operated
	Joint venture	Joint venture, Operate-owned	Joint venture, Outsource-operated	Joint venture, Jointly-operated

In the case of public container terminals operated by private companies, the port authority receives a rental income and management fees merely as the landlord of the terminals while the terminals are operated by private companies. Under this business model, the use of the terminals operated by the private companies is not limited to berthing by (privately/publicly?) owned vessels. Moreover, the terminals can be operated like enterprises, thus increasing berth operational efficiency.

The terminals can be further classified according to their ownership structure into wholly-owned terminals and joint venture terminals, and according to the operation of the terminals into operate-owned terminals, outsourced terminals and jointly-operated terminals. Upon consideration of ownership structure and operations, there are six types of business models of container terminal operations as shown in Table 5.3.

With a trend towards privatization and liberalization, the business models of container terminals have been changing over the past years (Pallis et al. 2010). The involvement of privately-owned business units in port management can enhance the operational efficiency of ports (Tongzon and Heng 2005). As a result, it has become increasingly among for container terminals to be operated jointly by the government and private companies (Wiegmans et al. 2002). In practice, most of the terminal operations such as unloading/unloading, towing, warehousing, etc. are outsourced to stevedoring companies. For example, the current operators of the container terminals at the Port of Kaohsiung tend to keep ownership and concession separate. Although the terminals are owned by the major investors, most of the various operations of the container terminals are outsourced to individual specialist operators.

5.2.2 *The Criteria for Choosing a Business Model of Container Terminal Operations*

This chapter refers to the results of the relevant past studies and interviews with practitioners on container terminal operations and describes the various criteria for selecting a business model of container terminal operations. There are four main factors to consider: the benefit factor, the concession factor, the environmental factor, and operational capabilities.

1. The benefit factor
    - Sources of funding

The requisite infrastructure for container terminal operations (rivers, wharf and container yards), upper facilities (cranes, straddle carriers and tractors) and other related assets involve massive capital investment (Musso et al. 2006). Apart from this, the proportion of fixed assets required by terminal operations is far greater than that required by other types of businesses while the depreciation period is longer. The fixed assets are also limited to terminal operation use. Augmented with the risk and uncertainty embedded in the shipping industry, estimating when the investment will break even becomes even more difficult. As a result, the source of funding is one of the important criteria for selecting a business model of container terminal operations (Musso et al. 2006). Under the current port policy, port construction and related facilities are usually financed through the introduction of private funds.

- Financial benefit

An enterprise will take into account operating costs when making decisions on investments. As container terminal operations involve tremendous capital investment, investors will attempt to estimate the financial benefits, particularly when they are private companies. Private companies are usually more willing than government departments to inject more capital into the facilities that can benefit operations. At the same time, private investors will pay relatively more attention to the rate of financial returns from operations (Wiegmans et al. 2002). Musso et al. (2006) pointed out that when making decisions on investments, potential investors evaluate two key factors: whether the net present value (NPV) obtained from an analysis of the cost effectiveness of a container terminal is positive, and whether the internal rate of return (IRR) of the investment in the container terminal is greater than the interest rate in the market. An analysis of the cost effectiveness, which is one aspect of the financial benefits of operating a container terminal, will help investors estimate whether the operations can generate profits. This evaluation dimension is vital to the operation and management of container terminals.

- Economic effectiveness assessment

The infrastructure of container terminals is a type of public asset. Research conducted by Wiegmans et al. (2002) found that the objective of container terminal operations is not merely to make profits, but also to bring benefits to society. Container terminals with sound operation and well developed ports will attract investment from enterprises, creating job opportunities indirectly and in turn economic benefits (Musso et al. 2006). When selecting a business model of container terminal operations, one should take into account the contribution the model can bring to the industry and the economy and not focus solely on the financial benefits (Lu et al. 2012)

- Timeline for implementing the business model

Since a lot of planning is involved in the construction of a container terminal before implementation can even begin, a tremendous time cost is incurred (Musso et al. 2006). When the shipping industry is robust and the market is in good shape, it will be easier to raise capital as more investors will be interested. This will bring a greater profit margin within a shorter period of time. On the contrary, when the shipping industry is stagnant or when supply is greater than demand, competition will be keener and it will be difficult to boost the volume of required operation. In this regard, the timeline for implementing the business model will be a very important criterion. Notably, the decision making process, the recruiting system and the financial management system under the current port operation system are still subject to the statutory orders and the restrictions set out by the administrative authorities, rendering them less autonomous. As a result, their operational efficiency and the effectiveness of their implementation timeline cannot be compared to those of private enterprises. When selecting a particular business model of container terminal operation, one should try to determine which model will reach its implementation stage quicker.

2. The concession factor
   - Ownership of the container terminal

The involvement of private companies in container terminal operations has become a trend (Farrell 2012). There are two major types of ownership by private companies: wholly-owned and joint venture. A joint venture in container terminal operations is when two or more companies jointly establish and operate a container terminal. A terminal is said to be wholly-owned when it is operated and owned solely by a port corporation. Lu et al. (2012) revealed that the Port of Kaohsiung, at which the container terminals are wholly-owned, is subject to the restrictions relevant to state-owned companies which impose more constraints on the operation of the container terminals. On the contrary, the container terminal operations are more flexible in a joint venture. However, under the legal restrictions in some districts, at least one party in the joint venture must be a public unit (Farrell 2012).

- Concession of container terminal operations

The government always plays a pivotal role in port operation. In recent years, many governments have attempted to consolidate their status in the market by engaging themselves in port operation (Musso et al. 2006). These governments usually hold the majority of shares which gives them strong control over the operation of the container terminals, e.g. those at the Port of Shanghai and the Port of Ningbo. While traditional economic theories generally opine that government intervention in the market will bring adverse effects, Wiegmans et al. (2002) argued that since perfect competition does not exist anyway, the government's holding of a certain proportion of the ownership of the container terminals can promote the healthy development of the industry as it allows them to participate in and monitor the operation of the terminals. Container terminal operations include shipside loading/unloading operations, tractor operations and other operations in container yards. Concession of container terminal operations exist in different forms under different business models of container terminal operations. It can thus be seen that the role played by government departments and private companies is interrelated with the decision making in investment and the selection of financial strategies (Dekker and Verhaeghe 2012).

3. The environmental factor
   - Shipping market and cargo volume

One of the important objectives in port business is to generate commercial activities at the port and meet stakeholders' requirements. In other words, investment in container terminals is aimed at increasing container volumes and enhancing the efficiency of the loading/unloading operations (Musso et al. 2006). Tsai (1991) believed that in selecting the investment strategy, overseas enterprises will evaluate the following aspects: extent of economic development in the region, the scale of the market, and whether there is adequate demand in the shipping market to cover the costs. These aspects should be considered in the selection of a business model of container terminal operations. In particular, the development of the shipping market is usually a factor when deciding on the level of investment and the business model of container terminal operations. Nevertheless, when deciding whether to invest in a specific container terminal, one should not consider merely the current situation of the shipping market, i.e. whether the industry is robust or not. As a result, any decisions on investment should be based on the short-term, medium-term and long-term analyses of the potential shipping market development.

- The willingness of potential investors

The most important factors when deciding whether to form a joint venture is who the business partner is and whether or not the partner (e.g. global terminal operator) is willing to invest. In the past few years, the trend has been for private companies instead of the public to own and operate the container terminals. In the current container terminal operation, the partners or investors are mostly container shipping companies and specialized container terminal operators like the Port of Singapore Authority (PSA), Hutchison Port Holdings Ltd (HPH), and Dubai Ports World (DPW). When selecting the business model of container terminal operations, one should factor in the willingness of the international container terminal operators to invest.

- The impact of existing terminal operators

Tsai (1991) argued that the investment behavior of overseas investors hinges greatly on the competition among existing container terminal operators. Musso et al. (2006) pointed out in a research report on port investment that the port is usually deemed public infrastructure and is expected to bring benefits to the national economy. If a port corporation and other operators in the industry decide to engage themselves in the operation of container terminals in the form of joint venture and if the port corporation does not need to pay land rates while enjoying greater cost advantage than other players in the industry, then the port corporation may be accused of competing with private companies for profits. Also, the existing operators in the port may request a port corporation to reduce the land rates of the container terminals and the port-related costs. This will affect the existing market mechanism (Lu et al. 2012).

- Restrictions imposed by the laws

The imposing of too many restrictions on the functions of human resources management, budgeting and purchasing of a port operation system will result in less flexibility on its operation. Farrell (2012) demonstrated that for a joint venture of container terminal operation in which at least one of the partners is a public organization, its operation will more likely be subject to legal restrictions. Wiegmans et al. (2002) suggested that due to the complexity of the legal process, it may take a long time for the container terminal operation strategy to reach its implementation stage and this may in turn reduce the investors' willingness. Kolstad and Villanger (2008) contended that since investment projects of infrastructures involve close contact with public departments, the investors should possess a good understanding of the local legal system. It was found that since the Taiwan International Ports Corporation is a state-owned company, its acquisition of equipment and services is subject to the law of tender for state-owned industries. Therefore, legal restrictions should be factored in to the selection of a business model of container terminal operations.

4. Operational capabilities
    – The specialized skills required for container terminal operations

The trend towards operation by private companies means that investors in container terminals will include many existing container terminal operators in the port and international container terminal companies. The main reason is that those operators have a better understanding of the market or possess the specialized skills required for container terminal operations which will enable their business partners to become familiar with the market situation more quickly and this will bring more advantages to the operation (Farrell 2012). In addition, the facilities related to the operation of infrastructure such as container terminal are always unique. The functions of different kinds of machinery and equipment are different. Such difference will have an impact on the estimation of cost. The participation of investors possessing specialized skills in the operation will not only enhance the operational efficiency of the container terminals, but will also reduce the running cost as a whole (Wiegmans et al. 2002). If TIPC desires to engage itself in the operation of container terminals, then forming a joint venture with an industry player possessing specialized techniques for container terminal operations will be a good option since its own staff lack the practical experience in container terminal operations.

– The effectiveness of acquisition of equipment

Wiegmans et al. (2002) noted that the acquisition of machinery and equipment takes a long time as it has to go through many legal processes. Making matters worse is that no income is generated during this acquisition period while extra costs such as interest will be incurred. Musso et al. (2006) commented that it is quite difficult to acquire the machinery and equipment and estimate the associated costs. The major reason is that the period of usage of the machinery and equipment is comparatively longer (20 years on average) than a private terminal operator while the amount of capital required may differ substantially for different kinds of operation. It may take 15–30 years to break even. In particular, the operational efficiency of the machinery and equipment is one important dimension for shipping companies when deciding whether their vessels should berth at a specific container terminal (Wiegmans et al. 2002). As the acquisition of machinery and equipment for a stated-owned company is subject to the restrictions under the law of tendering for stated-owned business, the process will take a long time. The time required for the acquisition should therefore be taken into account when selecting a business model of container terminal operations.

- Sales and marketing capabilities

Container terminals are operating in an increasingly competitive environment and the importance of sales and marketing capabilities in terminal business and development is gradually being recognized by the terminal operators. Ng and Yu (2006) studied the choice of a port and container terminal for berthing, and concluded that besides the port charges, customers usually also take into account the marketing capabilities and reputation of the port. As a result, it is important to acquire the ability to promote services and attract customers, regardless of the model of container terminal operations. It is suggested that a new container terminal operator should form a joint venture to acquire the ability to market its services and attract customers.

5.3 Methodology

The purpose of this chapter is to develop an evaluation framework for the selection of a business model of container terminal operations from the viewpoint of the Taiwan International Ports Corporation (TIPC) by assessing data collected from a survey questionaire completed by experts and practitioners in the port sector. The primarily survey was revised and adjusted according to the experts' recommendations. The data collected was analysed through the analytic hierarchy process approach to decide on a business model of container terminal operations for TIPC according to the weights of the evaluation criteria.

5.3.1 Sample

In general, the business model of container terminal operations is decided by a few senior managers or above. Therefore, the survey questionnaire was given to senior managers or above at TIPC and academics who were conducting research on container terminal operations. Seven copies of the survey were dispatched to TIPC and three to academics at maritime-related departments in Taiwanese universities on 25 February 2013. Five completed surveys were received on 11 March 2013. Any incomplete surveys were deemed invalid. The effective response rate was 40 %.

5.3.2 Analytic Hierarchy Process

The analytic hierarchy process is a multiple criteria decision-making method (Dolan 2008; Vaidya and Kumar 2006) developed by Saaty (1980). It has been used successfully by managers to make better decisions in complicated environments (Dolan

Table 5.4 A preliminary proposal on the business models of container terminal operations

Operation Ownership	Operate-owned	Operate-outsourced	Jointly-operated
Wholly-owned	Option 1 (WOOW): wholly-owned, operate-owned	Option 2 (WOOU): wholly-owned, operate-outsourced	Option 3 (WOJO): wholly-owned, jointly operated
Joint venture	Option 4 (JVOW): joint venture, operate-owned	Option 5 (JVOU): joint venture, operate-outsourced	Option 6 (JVJO): joint venture, jointly operated

2008). The process frames a decision as a hierarchy. Pairwise comparisons based on the scale ranging from 1/9 for 'less important than', to 1 for 'equal', and to 9 for 'absolutely more important than' are easy to understand for practitioners and managers (Dolan 2008; Vaidya and Kunar 2006). The key steps of this approach consists of defining the problem and objective, identifying the criteria that influence the choice of business models of container terminal operations, structuring the problem in a hierarchy of various levels of sub-criteria and alternatives or options, comparing the weight of each criterion based on the scales, performing a consistency test involving the consistency index and the consistency ratio, and developing an overall priority decision ranking (Ho 2008; Vaidya and Kunar 2006).

5.3.3 *A Preliminary Proposal on the Business Model of Container Terminal Operations*

This research is conducted from the perspective of TIPC. Based on the literature review in the preceding paragraphs, container terminals can be classified according to the form of ownership (either wholly-owned or a joint venture) and the type of investment (namely, operate-owned, operate-outsourced or jointly-operated). These two dimensions combine to give six categories of container terminals, namely "wholly-owned, operate-owned", "wholly-owned, operate-outsourced", "wholly-owned, jointly operated", "joint venture, operate-owned", "joint venture, operate-outsourced" and "joint venture, jointly operated" as shown in Table 5.4.

1. Justifications for the wholly-owned business models (options 1–3)

In the three wholly-owned models of container terminal operations, the port corporation concerned has made full investment in the container terminals. The advantage to these three models is that the port corporation has full control of the terminal operations, since it holds 100 % ownership of the terminal. As a result, wholly-owned container terminals enjoy greater flexibility in its operations which helps to attract shipping carriers.

A "wholly-owned, operate-owned" container terminal is one that is wholly invested by a single company with the shares wholly held by that company (David and Stewart 2010). This model of setting up a specialized department or a subsidiary company under the whole ownership of a port corporation for the purpose of operating container terminals is similar to the business model adopted by Hutchison Port Holdings Ltd (HPH). HPH, which was established in 1994, is a subsidiary company of Hutchison Whampoa Limited (HWL). HPH is responsible mainly for the management of the port business and related global services of the Whampoa Group (Notteboom and Rodrigue 2010). PSA is another example of this kind of business model. This company was established as a subsidiary company under the whole ownership of PSA and is specialized in the provision of all-rounded quality services for shipping carriers (Notteboom and Rodrigue 2010). HPH and PSA are both examples of success. In the meantime, the operation of the two companies is no longer limited to their countries. Rather, the two companies have been participating actively in the investment, operation, and development of container terminals in different countries around the world.

With regard to the business model in which the concession of the operation is outsourced, Notteboom and Winkelmans (2001) revealed that if an enterprise is not competitive enough in the global environment, then that enterprise should outsource some of its business activities. In addition, under the trend of globalization, the business model of outsourcing will enable shipping companies, freight forwarders and container terminal operators to develp a new business. Under the business model of "wholly-owned, operate-outsourced", the container terminal is wholly invested by the port corporation while the associated business activities of the container terminal are outsourced to different operators, including stevedoring companies and other container terminal operators. The port corporation is not involved in the shipside operations or operations in the container yards. Having outsourced other associated business to companies with specialized techniques, the port corporation can then focus on the development of its core business. This may be one of the best business models (Quinn and Hilmer 1994). Nevertheless, the port corporation, which lacks practical experience in container terminal operations, will be deprived of the opportunity to learn and acquire specialized skills and capabilities required for container terminal operations since all the specialized skills are provided by outside companies (Quinn and Hilmer 1994).

Under the business model of "wholly-owned, jointly operated", the port corporation owns and maintains container terminals while the operations are taken up jointly by the port corporation and stevedoring companies or container terminal operators. The advantage of this business model is that the party investing in the container terminals can also participate in the operation of the terminals. This model can help enhance the operational capabilities of the staff of the investing company, i.e. the port corporation (Lu et al. 2012). Nevertheless, the port corporation is required to make tremendous capital investment in the machinery and equipment at the initial stage as it is involved in some of the container terminal operations. In addition, it is likely that disputes will arise between the partners of the joint operation over how the profits should be split. The disputes will adversely affect the partnership.

2. Justifications for the joint venture business models (options 4–6)

Under the business model of joint venture, there are three options, namely “joint venture, operate-owned”, “joint venture, operate-outsourced” and “joint venture, jointly operated”. David and Stewart (2010) pointed out that a joint venture is formed when two or more companies have co-ownership of the concerned company. This kind of partnership is usually formed when the local government has imposed some legal restrictions on foreign investors. In some cases, the container terminal operators form partnerships with local container terminal operators such as a port corporation to achieve greater profits as doing so can reduce the hindrance and setback imposed by the local government. The container terminal can thus be operated more smoothly. While the capital investment in container terminal facilities is tremendous, forming joint ventures with partners can, to a certain extent, help solve the problem of inadequate capital for a single container terminal operator. In addition, in the case where the partner of a joint venture is a specialized container terminal operator, the investing company can acquire the specialized techniques and experience required for container terminal operations.

The business model of “joint venture, operate-owned” refers to the scenario where a port corporation co-invests in the container terminal operations with other investors. A partner of the joint venture owns some shares according to the proportion of its investment. The partner with the largest investment usually holds the greatest proportion of ownership. Under this business model, the port corporation could enjoy high flexibility in the operation since it is solely in charge of all the related activities of the container terminal. Nevertheless, since the port corporation lacks the experience required for container terminal operations, a higher operational cost may be incurred. This will in turn reduce the profit margin of the investment and reduce its attractiveness to the existing business partners who may decide to discontinue their investment. In view of this situation, there is a recent trend for global container terminal operators that have set up container terminal companies through joint ventures to enter into profit-sharing agreements with shipping companies with which they have been co-operating for a long time (Notteboom and Rodrigu 2010). The profit-sharing terms set forth in the agreement can help the terminal operators to minimize their operational risk.

Container terminal operators need to provide a number of services to shipping companies. These services include planning of the vessels’ berthing time, sufficient machinery and equipment, warehousing facilities, comprehensive tractor services inside the yard and cargo tracking facilities, all of which are related to the operation of container terminals (Rao and Young 1994). In the case where a company needs to provide a diversity of services but does not possess the specialized skills required by the different departments, then this company may choose to outsource (Rao and Young 1994). Smith et al. (1998) explained that outside companies usually possess the advantage of economies of scale. This will render greater efficiency in controlling the operational cost and in turn generate a greater profit margin. Mcfarlan and Nolan (1995) also revealed that outsourcing different business activities to different companies will save the container terminal operators much time and resources needed to coordinate the different departments. Gupta and Gupta (1992) also pointed out that outsourcing will not affect the operational efficiency of the organization in its original areas of operation and will not reduce the ability of the organization to react to the market. In summary, the greatest advantage of the business model of operate-outsourced is the reduction in operational cost and the provision of professional services for customers. In addition, no special risk will arise from this business model of operation (Rao and Young 1994).

The business model of "joint venture and operate-outsourced" refers to the scenario where the port corporation and other operators co-invest in the container terminals but none of the parties will be involved in the operation of the container terminals. Rather, they will delegate the work to stevedoring companies and other container terminal operators specialized in the relevant service fields (Lu et al. 2012). This business model combines the advantages of joint ventures and outsourcing. Nevertheless, it has some disadvantages. For example, the various parties investing in the container terminal operations may be deprived of the opportunity to nurture the interdisciplinary ability of integrating the different functions because none of them would be involved in the operation of the container terminal (Quinn and Hilmer 1994).

The business model of "joint venture, jointly operated" refers to the scenario where the port corporation and other operators co-invest in the container terminal operations. The proportion of ownership of each party is usually related to the the amount of money invested by each. Nevertheless, there are cases where a container terminal operator is subject to local legal restrictions and therefore invests the same amount of money as a local container terminal operator or the local government, or where the foreign investor's proportion of equity shares is less than that of the local container terminal operator or the local government (Rossignol 2007). This business model possesses the merits of joint venture and joint operation. A port corporation that operates a container terminal jointly with stevedoring companies or other container terminal operators will hold the ownership of the container terminal and at the same time has the opportunity to acquire experience in operating a container terminal. However, since the concerned port corporation may be required to co-operate with a large number of operators from different fields, the operation will become more complicated and less manageable for the port corporation.

The business models of "wholly-owned, jointly-operated", "joint venture, jointly-operated", and "joint venture, operate-owned" are (seldom/never?) adopted by TIPC. The greatest problem with the business model of "jointly-operated" is the time constraint of the operation. Container terminal operations have been moving towards greater segregation and differentiation of the functions. In addition, licences have to be acquired for many of the operations of the container terminal. Since the port corporation is a unit under the state-owned sector, it will be required to go through many complicated and lengthy application processes to acquire the licences. This may reduce the willingness of industrial players to form joint ventures with the port corporation. In addition, although the port corporation may be able to acquire experience in container terminal operations during the course of co-operation with the container terminal operators, whether this can be achieved depends on the willingness of the container terminal operators to transfer the specialized knowledge to the staff of the port corporation. If this can be achieved, it will help to enhance the port corporation's operational capabilities. Furthermore, as the profits generated from the joint venture will be shared between the port corporation and its partner(s), the profits that the port corporation will receive from the joint venture may ultimately be less than that obtained if the terminal is operate-owned or outsource-operated. In view of this, this research will exclude the "wholly-owned, jointly-operated" and "joint venture, jointly-operated" business models under the main category of "jointly operated" from evaluation (Lu et al. 2012).

The business model of "joint venture, operate-owned" can reduce the financial burden of the port corporation. Nevertheless, since the port corporation may not possess any practical experience in container terminal operations, its operational capabilities and its ability to generate profits may not be satisfactory. This may reduce the willingness of container terminal operators to form joint ventures with them. Further, in the case where the port corporation has relied too much on foreign investment while owning the concession of the operation, when the two parties hold opposing views about the operation leading to a withdrawal of capital or unwillingness to invest, there is the fear that the container terminal operations will be adversely affected. The willingness of the partners to invest is reduced if the container terminal is operated entirely by the port corporation.

Since the port corporation is in control of the operation of the container terminal, it is more likely that disputes will arise and result in operational risks. In view of these drawbacks, this research will also not evaluate the "joint venture, operate-owned" business model. That leaves three business models of container terminal operation to be analyzed, namely "wholly-owned, operate-owned", "wholly-owned, operate-outsourced" and "joint venture, operate-outsourced". The analytic hierarchy process is adopted and the evaluation framework for the container terminal operation for a port corporation is shown in Fig. 5.1.

5.4 Results of Empirical Analyses

5.4.1 Choosing a Business Model of Container Terminal Operation in the Short To Medium Term

- Importance of the main criteria for choosing a business model of container terminal operations in the short to medium term as perceived by respondents

The questionnaires received were tested for consistency of measures on the main criteria and sub-criterion using a software called Expert Choice 2000. Subsequently, the value of the primitive weight and the relative weight under the original dimensions are ranked. According to Satty (1990), a consistency ratio < 0.1 means that the pairwise matrix is consistent. The short to medium term is defined as a period of less than 5 years in this research. As shown in Table 5.5, the dimension of "capability" received the highest weight (0.336), followed by "benefit" (0.312), "environment"

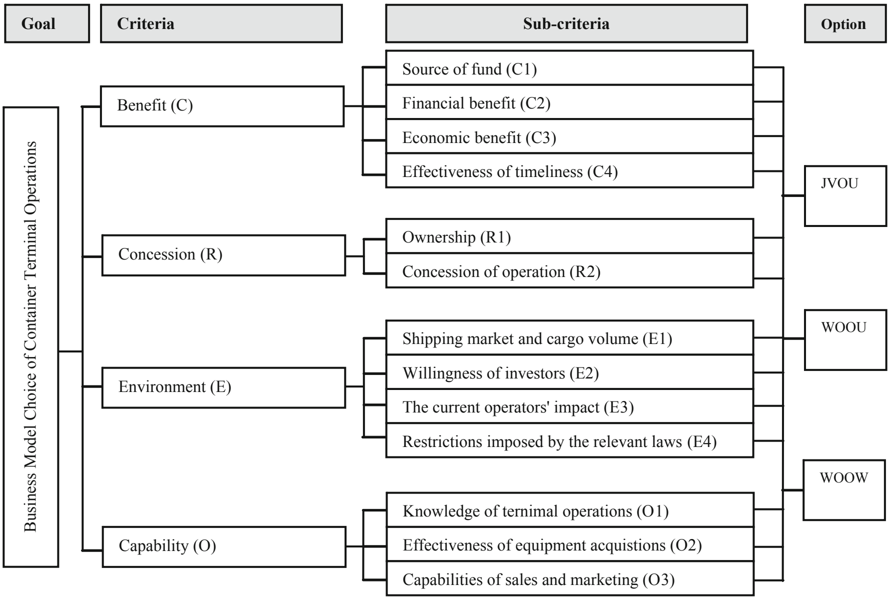


Fig. 5.1 Criteria for choosing business models of container terminal operations based on the analytic hierarchy process

Table 5.5 Importance of the main criteria for choosing a business model of container terminal operations in the short to medium term

Main criterion	Weight	Overall ranking
Benefit	0.312	2
Concession	0.153	4
Environment	0.199	3
Capability	0.336	1

(0.199) and “concession” (0.153). The result suggest that at present the port corporation does not possess sufficient capabilities to operate a container terminal by itself. Thus, this dimension was perceived as the most important dimension influencing the selection of a business model of container terminal operations.

- Importance of the sub-criteria for choosing a business model of container terminal operations in the short to medium term as perceived by respondents

From the ranking of the sub-criteria by weight as shown in Table 5.6, it is seen that the five most important sub-criteria are “knowledge of container terminal operations” (0.1515), “sales and maketing capabilities” (0.1391), “financial benefits” (0.1198), “effectiveness of the implementation timeline” (0.1017) and “ownership” (0.087975). Contrarily, “sources of funding” (0.0293) is the least important sub-criterion. The result highlights the relative importance of the knowledge of container

Table 5.6 Importance of the sub-criteria for choosing a business model of container terminal operations in the short to medium term

Main criterion	Sub-criterion	Original weight	Relative weight	Overall ranking
Benefit	Sources of funding	0.094	0.0293	13
	Financial benefits	0.384	0.1198	3
	Economic benefits	0.197	0.0615	7
	Effectiveness of the implementation timeline	0.326	0.1017	4
Concession	Ownership	0.575	0.0880	5
	Concession of operation	0.425	0.0650	6
Environment	Shipping market and cargo volume	0.176	0.0350	12
	Investors' willingness	0.257	0.0511	10
	The current operators' impact	0.293	0.0583	8
	Restrictions imposed by the relevant laws	0.274	0.0545	9
Capability	Knowledge of container terminal operations	0.451	0.1515	1
	Effectiveness of equipment acquisitions	0.134	0.0450	11
	Sales and marketing capabilities	0.414	0.1391	2

terminal operations and sales and marketing capabilities to attract customers in the selection of business models. It also confirms that TIPC has sufficient sources of funding to operate container terminals in the short to medium term.

- Results of analysis on which business model of container terminal operations is the best for TIPC in the short to medium term

Table 5.7 shows that the "wholly-owned, operate-outsourced" model (0.524) ranks the highest, followed by the "joint venture, operate-outsourced" model (0.319) and the "wholly-owned, operate-owned" model (0.157). The "capability" is clearly the most important major criterion in the short to medium term (within 5 years) which is why the port corporation should adopt the "operate-outsourced" business model. Moreover, as the TIPC will not encounter a lack of funding in the short to medium term while participants having placed importance on the dimension of "ownership

Table 5.7 Business models of container terminal operations for TIPC in the short to medium term

Business model	Weight	Overall ranking
Joint venture, operate-outsourced	0.319	2
Wholly-owned, operate-outsourced	0.524	1
Wholly-owned, operate-owned	0.157	3

Table 5.8 Importance of the main criteria for choosing a business model of container terminal operations in the long term

Main criterion	Weight	Overall ranking
Benefit	0.333	1
Concession	0.138	4
Environment	0.239	3
Capability	0.289	2

of container terminal" (ranking fifth in the list of sub-criteria), it can be said that the "wholly-owned, operate-outsourced" model will be more feasible than the "joint venture, operate-outsourced" model.

5.4.2 *Choosing a Business Model of Container Terminal Operations in the Long Term*

- Importance of the main criteria for choosing a business model of container terminal operations in the long term as perceived by respondents

Results shown in Table 5.8 indicate that "benefit" (0.333) is the most important criterion for choosing a business model of container terminal operations in the long term, followed by "capabilities" (0.289), "environment" (0.239) and "concession" (0.138). These results suggest that, in the long term, as the port corporation has sufficient time to nurture its capabilities of operating the container terminal, operational capabilities will not be the most important dimension. On the contrary, more importance will be placed on the ability to generate profits for the port corporation. As a result, the dimension "benefit" was ranked the highest in terms of importance.

- Importance of the sub-criteria for choosing a business model of container terminal operations in the long term as perceived by respondents

Table 5.9 shows the results of the five most important sub-criteria for choosing a business model of container terminal operations in the long term including "sales and marketing capabilities" (0.1731), "financial benefits" (0.1252), "effectiveness of the implementation timeline" (0.0896), "knowledge of container terminal operation"

Table 5.9 Importance of the sub-criteria for choosing a business model of container terminal operations in the long term

Main criterion	Sub-criterion	Original weight	Relative weight	Overall ranking
Benefit	Sources of funding	0.116	0.0386	12
	Financial benefits	0.376	0.1252	2
	Economic benefits	0.239	0.0796	5
	Effectiveness of the implementation timeline	0.269	0.0896	3
Concession	Ownership	0.534	0.0737	6
	Concession of operation	0.466	0.0643	7
Environment	Shipping market and cargo volume	0.256	0.0612	10
	Investors' willingness	0.263	0.0629	9
	The current operators' impact	0.213	0.0509	11
	Restrictions imposed by the relevant laws	0.268	0.0641	8
Capability	Knowledge of container terminal operations	0.296	0.0855	4
	Effectiveness ofequipment acquisitions	0.105	0.0303	13
	Sales and marketing capabilities	0.599	0.1731	1

(0.0855) and "economic benefits" (0.0796). Contrarily, the sub-criteria of "effectiveness of equipment acquistions" (0.0303), "sources of funding" (0.0386) and "the current port operators' impact" (0.0509) were the least important sub-criteria. The port corporation will not encounter problems in acquiring capital even in the long term. It has already acquired the necessary machinery and equipment and the impact of the existing port operators has gradually subsided. In this case, the port corporation will place more importance on its marketing capability and its ability to attract customers and increase cargo volume. The port corporation will also place an emphasis on financial benefits. Thus the economic benefits to the nation and society may also be included as an important sub-criterion for choosing a business model of container terminal operations.

Table 5.10 Business models of container terminal operations for TIPC in the long term

Business model	Weight	Overall ranking
Joint venture, operate-outsourced	0.224	3
Wholly-owned, operate-outsourced	0.230	2
Wholly-owned, operate-owned	0.546	1

- Results of analysis on which business model of container terminal operations is the best for TIPC in the long term

As can be seen in Table 5.10, results indicate that the "wholly-owned, operate-owned" model (0.546) was ranked the highest in terms of importance, followed by the "wholly-owned, operate-outsourced" model (0.230) and the "joint venture, operate-outsourced" model (0.224). The raising of funds will not be a problem for TIPC in the long term and the company does not need any partners to share the operation risk. Also, it has sufficient time to nurture its capabilities of operating container terminals. Therefore, the "operate-outsourced" model will not be considered. Rather, the "self operation" model will more likely be adopted. Apart from this, the sub-criterion of "economic benefits", in addition to "financial benefits", will be taken into consideration in the long term. Under the "wholly-owned, operate-owned" business model, the port corporation will find it easier to make a contribution to the national economy than if either of the other two business models is adopted. As a result, the "wholly-owned, operate-owned" business model is more feasible in the long term.

5.5 Discussion and Conclusion

Container terminals facilitate the loading/discharging of containers, and provides storage and infrastructure for ship operations which play an important role in global supply chains. Terminal operations, including crane deployment and operations (Cheung et al. 2002; Kim and Bae 2004; Meisel and Bierwirth 2013; Vis and Carlo 2010; Zhang et al. 2002), yard capacity and management (Jin et al. 2014; Kim and Kim 2002; Petering 2011; Taleb-Ibrahimi 1993; Zhang et al. 2003), and quayside operations (Chao and Lin 2011; Chen et al. 2012; Goodchild and Daganzo 2007) have been widely examined. But few studies have examined the choice of a business model of container terminal operations based on the ownership and operational management structure. Theoretically, this chapter highlights the importance of ownership and operational capabilities in explaining the choice of a business model for a new entrant or investor. We answer several important questions with regard to the choice of a business model of container terminal operations. First, which ownership and operational management structure are the major business models based on, and what do investors perceive as the criteria for choosing a business model? Second, what is the relative importance of the main criteria and sub-criteria based on the analytic hierarchy process in the short to medium term and the long term? To the best of our knowledge, this is the first study to provide empirical evidence on the importance of a criterion or choosing a business model of container terminal operations. More specifically, this research fills the gap in the literature by explaining business models from the perspective of ownership and operational management structure in container terminal operations.

5.5.1 Implications

The research findings bear several implications. First, benefit and operational capability are the most important criteria influencing the choice of a business model of container terminal operations for TIPC and investors not only in the short to medium term but also in the long term. Survey respondents considered knowledge and sales and marketing capabilities to be of the greatest importance in the short to medium term. This suggests that the port corporation should focus on nurturing the specialized knowledge and capabilities of operating the container terminal. Then the port corporation will no longer need to rely on other parties for operation. This approach will give the port corporation fuller ownership and concession of the port corporation. In addition, the nurturing of operational capabilities will facilitate the port corporation's expansion of business overseas.

Second, this research indicates that whether it is the short to medium term or long-term development of the container terminal operation, a key parameter influencing the selection of a business model of container terminal operations is the marketing strategy and incentives to investment. The port corporation should therefore make reference to the marketing strategies adopted by the largest container terminals in the world such as the ports of Singapore, Shanghai, Hong Kong, Busan, etc. and tailor marketing strategies to shipping companies. To do so, the port corporation should launch incentive schemes that are more attractive than those of other ports such as discounts and subsidies on loading/unloading charges and tractor costs to attract cargo.

Third, the capabilities of TIPC to market its services and attract customers need to be further strengthened. The company should place more importance on the capabilities of operation and therefore it should choose the "wholly-owned, operate-outsourced" business model in the short to medium term (during the first 10 years of operation). In the long term (after the first 10 years of operation), the "wholly-owned, operate-owned" model should be adopted. A container terminal operator with a longer history may already possess operational capabilities and its operation objectives should no longer be limited to making profits but should include making a contribution to the national economy and the society. In this regard, the "wholly-owned, operate-owned" model will be more feasible.

Finally, it was found from the interviews conducted in this research that the port corporation is subject to the restrictions set out by the law of tender of state-owned business. Moreover, the restrictions set out by a number of laws such as the Commercial Port Law, Law for Promotion of Private Participation in Infrastructure

Projects and Act Governing Relations between the People of the Taiwan Area and the Mainland Area will exert influences on the policy decision process relevant to the operation of the container terminal. These restrictions will have major influences on the timeline for implementing the business model. The port corporation should therefore focus its efforts on the planning of activities relevant to the operation of the container terminal to avoid delays.

5.5.2 Limitations and Future Research

This research on the choice of business models of container terminal operations does have its limitations. First, the scope of the research was limited to the choice of a business model from the perspective of TIPC in Taiwan. It will be worthwhile to assess the same choice from a global terminal operator's perspective such as PSA, DP World, and HIT. Second, the number of senior managers or above and academics possessing an understanding of the business models of container terminal operations is limited, which resulted in a small number of respondents. Third, this study specifically focused on the container terminal operations. Future studies could apply the assessment criteria identified in this study to other sectors, such as airports, air cargo terminals, dry bulk terminals and international distribution centers. Despite its limitations, the study still managed to examine the important criteria for choosing a business model of container terminal operations. It also provides the basis for future research in other sectors and other countries. Finally, methodologically, an analytic hierarchy process was employed to identify the weights of the criteria. Other methods to assess the effects of criteria on the business model choice based on a cause-and-effect relationship might provide additional insights.

Acknowledgments We would like to thank Dr. T.L. Yip, Associate Director of the C.Y. Tung Internatinal Centre for Maritime Studies and Dr. Johnny Wan at the Department of Logistics and Maritime Studies, The Hong Kong Polytechnic University, for their comments and suggestions which helped to improve the quality of this research.

References

Alderton, P. M. (2008). *Port management and operation* (3rd ed.). Informa: London.

Baird, A. J. (1995). Privatization of trust ports in the United Kingdom: Review and analysis of the first sales. *Journal of Transport Policy, 2*(2), 135–143.

Chao, S. L., & Lin, Y. J. (2011). Evaluating advanced quay cranes in container terminals. *Transportation Research Part E: Logistics and Transportation Review, 47*(4), 432–445.

Chen, S. L. (2009). Port administrative structure change worldwide: Its implication for restructuring port authorities in Taiwan. *Transport Reviews, 29*(2), 163–181.

Chen, J. H., Lee, D. H., & Cao, J. X. (2012). A combinatorial benders' cuts algorithm for the quayside operation problem at container terminals. *Transportation Research Part E: Logistics and Transportation Review, 48*(1), 266–275.

Cheon, S. H., Dowall, D. E., & Song, D. W. (2010). Evaluating impacts of institutional reforms on port efficiency changes: ownership, corporate structure, and total factor productivity changes of world container ports. *Transportation Research Part E: Logistics and Transportation Review, 46*(4), 546–561.

Cheung, R.K., Li, C. L., & Lin, W. (2002). Interblock crane deployment in container terminals. *Transportation Science, 36*(1), 79–93.

David, P. A., & Stewart, R. D. (2010). *International logistics—the management of international trade operations* (3rd ed.). USA: Cengage Learning.

Dekker, S., & Verhaeghe, R. J. (2012). *The blackwell companion to maritime economics* (1st ed.). USA: Blackwell.

Dolan, J.G. (2008). Shared decision-making—transferring research into practice: The analytic hierarchy process (AHP). *Patient Education and Counseling, 73*(3), 418–425.

Drewry Maritime Research. (2014). Global container terminal operators annual review and forecast 2013. http://www.drewry.co.uk/news.php?id=232. Accessed 27 June 2014.

Farrell, S. (2012). The ownership and management structure of container terminal concessions. *Maritime Policy and Management, 39*(1), 7–26.

Fransoo, J. C., & Lee, C. Y. (2013). The critical role of ocean container transport in global supply chain performance. *Production and Operations Management, 22*(2), 253–268.

Goodchild, A. V., & Daganzo, C. F. (2007). Crane double cycling in container ports: Planning methods and evaluation. *Transportation Research Part B: Methodological, 41*(8), 875–891.

Gupta, G., & Gupta, H. (1992). Outsourcing the IS function—is it necessary for your organization? *Information System Management, 9*(3), 44–50.

Heaver, T., Meersman, H., & Moglia, F., Van De Voorde, E. (2000). Do mergers adn alliances influence European shipping and port competition? *Maritime Policy and Mangement, 27*(4), 363–373.

Ho, W. (2008). Integrated analytic hierarchy process and its applications—A literature review. *European Journal of Operational Research, 186*(1), 211–228.

Jin, J. G., Lee, D. H., & Cao, J. X. (2014). Storage yard management in maritime container terminals. *Transportation Science*. (article in advance)

Kim, K. H., & Kim, H. B. (2002). The optimal sizing of the storage space and handling facilities for import containers. *Transportation Research Part B: Methodological, 36*(9), 821–835.

Kim, K. H., & Bae, J. W. (2004). A look-ahead dispatching method for automated guided vehicles in automated port container terminals. *Transportation Science, 38*(2), 224–234.

Kolstad, I., & Villanger, E. (2008). Determinants of foreign direct investment in services. *European Journal of Political Economy, 24*(2), 518–533.

Lu, C. S., Liao, C. H., Li, C. W., Tai, H. H., Yang, C. C., & Li, C. C. (2012). Development of transshipment market and operations models of Container Berth 75 and Container Berth 120 in the Port of Kaohsiung, Taiwan International Ports Corporation.

McFarlan, F. W., & Nolan, R. L. (1995). How to manage an IT outsourcing alliance? *Sloan Management Review, 36*(2), 8–23.

Meisel, F., & Bierwirth, C. (2013). A framework for integrated berth allocation and crane operations planning in seaport container terminals. *Transportation Science, 47*(2), 131–147.

Musso, E., Ferrari, C. & Benacchio, M. (2006). Port investment: Profitability, economic impact and financing. In K. Cullinane & W. Talley (Eds.), *Port economics: Research in transport economics* (Vol 16, pp. 171–218) Amsterdam: Elsevier.

Ng, A., & Yu, K. (2006). Assessing the attractiveness of ports in the North European container transhipment market: An agenda for future research in port competition. *Maritime Economics and Logistics, 8*(3), 234–250.

Notteboom, T. E., & Winkelmans, W. (2001). Structural changes in logistics: How will port authorities face the challenge? *Maritime Policy and Management, 28*(1), 71–89.

Notteboom, T. E., & Rodrigue, J. P. (2010). The corporate geography of global container terminal operator. *Maritime Policy and Management, 39*(3), 249–279.

Pallis, A. A., Vitsounis, T. K., & Langen, P. W. D. (2010). Port economics, policy and management: Review of an emerging research field. *Transport Reviews, 30*(1), 115–161.

Petering, M. E. H. (2011). Decision support for yard capacity, fleet composition, truck substitutability, and scalability issues at seaport container terminals. *Transportation Research Part E: Logistics and Transportation Review, 47*(1), 85–103.

Quinn, J. B., & Hilmer, F. G. (1994). Strategic outsourcing. *Sloan Management Review, 35*(4), 43–55.

Rao, K., & Young, R. R. (1994). Global supply chains: Factors influencing outsourcing of logistics functions. *International Journal of Physical Distribution and Logistics Management, 24*(6), 11–19.

Rossignol, M. H. L. (2007). *Global terminal operators: An overview*. the Inter-American Committee on Port (CIP): Washington.

Saaty, T. L. (1980). *The analytic hierarchy process*. New York: McGraw-Hill.

Saaty, T. L. (1990). How to make a decision: The Analytic Hierarchy Process. *European Journal of Operational Research, 48*(1), 9–26.

Smith, M. A., Mitra, S., & Narasimhan, S. (1998). Information systems outsourcing: A study of pre-event firm characteristics. *Journal of Management Information Systems, 15*(2), 61–93.

Taleb-Ibrahimi, M., Castilho, B. D., & Daganzo, C. F. (1993). Storage space vs handling work in container terminals. *Transportation Research Part B: Methodological, 27*(1), 13–32.

Tongzon, J., & Heng, W. (2005). Port privatization, efficiency and competitiveness: Some empirical evidence from container ports (terminals). *Transportation Research Part A: Policy and Practice, 39*(5), 405–424.

Tsai, P. L. (1991). Determinants of foreign direct investment in Taiwan: An alternative approach with time-series data. *World Development*, *19*(2/3), 275–285.

United Nations Conference on Trade and Development (UNCTAD). (2013). *Review of maritime transport 2013*. New York: United Nations Publication.

Vaidya, O. S., & Kumar, S. (2006). Analytic hierarchy process: An overview of application. *European Journal of Operational Research*, *169*(1), 1–29.

Verhoeven, P. (2010). A review of port authority functions: Towards a renaissance? *Maritime Policy and Management, 37*(3), 247–270.

Vis, I. F. A., & Carlo, H. J. (2010) Sequencing two cooperating automated stacking cranes in a container terminal. *Transportation Science, 44*(2), 169–182.

Wiegmans, B. W., Ubbels, B., & Rietveld, P., Nijkamp, P. (2002). Investments in container terminals: Public private partnerships in Europe. *Maritime Economics, 4*(1), 1–20.

Zhang, C., Wan, Y. W., Liu, J., & Linn, R. J. (2002). Dynamic crane deployment in container storage yards. *Transportation Research Part B: Methodological, 36*(6), 537–555.

Zhang, C., Liu, J., Wan, Y. W., Murty, K. G., & Linn, R. J. (2003). Storage space allocation in container terminals. *Transportation Research Part B: Methodological, 37*(10), 883–903.

Part II
Shipping Liners: Tactical and Operational Management

Chapter 6
Empty Container Repositioning

Dong-Ping Song and Jing-Xin Dong

Abstract Empty container repositioning (ECR) is one of the most important issues in the liner shipping industry. Not only does it have an economic effect on the stakeholders in the container transport chain, but it also has an environmental and sustainability impact on the society since the reduction of empty container movements will reduce fuel consumption, and reduce congestion and emissions. This chapter first analyzes the main reasons that cause empty container repositioning. Secondly, we provide a literature review with the emphasis on modeling the ECR problem from the network scope, e.g. modeling ECR in seaborne transportation network, modeling ECR in inland or intermodal transportation network, and treating ECR as a sub-problem or a constraint under other decision-making problems. Thirdly, we discuss the solutions to the ECR problems from the logistics channel scope perspective, which are categorized into four groups including organizational solutions, intra-channel solutions, inter-channel solutions, and technological innovations. Fourthly, we discuss the solutions to the ECR problems from the modeling technique perspective, which includes two broad research streams: network flow models and inventory control-based models. We then present two specific models representing the above two research streams, which aim to tackle the ECR problems in stochastic dynamic environments considering both laden and empty container management.

6.1 Introduction

Container ships carry an estimated 52 % of global seaborne trade in terms of value (UN 2013). Container shipping has experienced a rapid development in the last two decades. According to the data from Containerization International (ci-online.co.uk) and United Nations (UN 2008, 2012, 2013), the container traffic has increased from 84.6 million TEUs (20-foot equivalent unit) in 1990 to 485 million TEUs in 2007

D.-P. Song (✉)
School of Management, University of Liverpool, Chatham Street, Liverpool, L69 7ZH, UK
e-mail: Dongping.Song@liverpool.ac.uk

J.-X. Dong
Business School, Newcastle University, Newcastle upon Tyne, NE1 4SE, UK
e-mail: jingxin.dong@ncl.ac.uk

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_6

(before the global economic crisis in 2008), to 602 million TEUs in 2012. The annual growth rate was about 10.8 % in the period from 1990 to 2007, and about 9.3 % in the period from 1990 to 2012. The above growth rates were well above the average world trade growth rate around 6 %, and also showed the continuous growth despite the global economic crisis in 2008. There are several factors that have contributed to the rapid growth of container traffic in the world. Firstly, in the last two decades more and more goods have been containerized, not only the majority of manufactured goods, but also commodities such as coffee and refrigerated cargos (e.g. fruit, meat, fish). Secondly, the size of the containerships has increased dramatically from about 5000 TEUs in 1990s (Post-Panamax vessels) to 18,000 TEUs in 2013 (Maersk's triple-E series, where the Triple-E stands for energy efficiency, economies of scale and environmental improvements). One major shipping line in China, CSCL, placed orders for even larger container ships in 2013, which are scheduled to carry 18,400 TEU and to be delivered in 2014. The deployment of the mega-vessels has reshaped the container shipping networks, e.g. from direct service network to hub-and-spoke systems in many cases, which requires more double-handling (i.e. transshipment) at the hub ports. For example, the share of transshipments in total port throughput has grown from 10 % in 1980 to 27 % in 2007 (UN 2008). Transshipment plays a particularly important role in hub ports such as Singapore, Hong Kong, Busan and Rotterdam. For instance, in Hong Kong port, transshipment cargo movements took up 57 % of port cargo throughput in 2011 (Hong Kong Census and Statistics Department 2012). Thirdly, the world trade becomes more imbalanced and empty container movements have accounted for a significant percentage of port traffic. The last point, empty container repositioning (ECR), is the main topic of this chapter.

The trade imbalance of container shipping and the economic impact of empty container management have been well documented in the literature. In the Europe-Asia and Trans-Pacific trade routes, European ports and American ports have been experiencing a high surplus of empty containers, while Asian ports are facing severe shortages. Drewry Shipping Consultant estimated that about 20 % of all ocean container movements have involved repositioning of empty boxes since 1998 (Mongelluzzo 2004; Drewry 2006). According to the data in the annual reports published by United Nations Conference on Trade and Development (UN 2005, 2008, 2011, 2012, 2013), the container trade volume from Asia to Europe was between twice and three times of the volume in the opposite direction in the last decade. In other words, at least half of the boxes moving westward to Europe were sent back empty. The percentage of empty container movements in inland networks could be higher since empty containers are often stored at ports or depots, which are away from the demand locations. Various reports have shown that the share of empty containers in hinterland transport ranges from 40 to 50 % of all containers transported (e.g. Crainic et al. 1993a; Konings 2005; Braekers et al. 2011).

A number of cost components could be incurred in relation to empty containers including handling and transshipping at the terminals/ports/depots, storage and maintenance at empty warehouses, chassis location for drayage, inland transportation by rail or truck, and seaborne repositioning by vessels. Various sources have provided estimations of the overall cost of empty container repositioning. For example, it was

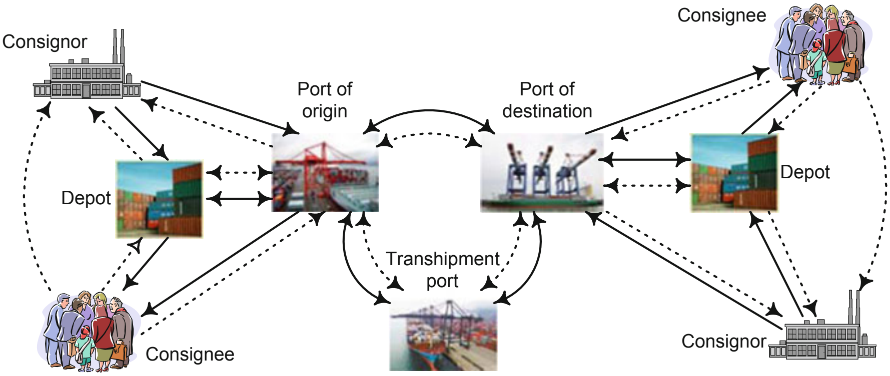


Fig. 6.1 The container transport chain. (Note: solid-lines indicate laden container flows and dashed-lines indicate empty container flows)

reported that the cost of container management inefficiencies in year 2001 reached almost US$ 17 billion (Boile 2006; Theofanis and Boile 2009). Drewry Consultant stated that empty container repositioning costs have reached US$ 20 billion yearly (Veestra 2005). Based on the data for 2002, Song et al. (2005) simulated the global maritime container shipping business and reported that the cost of repositioning empties was just under US$15 billion, which was about 27 % of the total world fleet running cost. It was estimated that shipping companies spent about US$ 110 billion per year in managing their container fleets (e.g. purchase, maintenance, repairs), of which US$ 16 billion (or 15 %) for repositioning empty containers (Rodrigue et al. 2013). It is estimated by the UN (UN 2011) that the cost of seaborne empty container repositioning was about $ 20 billion in 2009. If the cost of landside transportation of empty container repositioning is considered, the total cost would reach $ 30.1 billion and account for 19 % of global industry income in 2009. Although the reported figures of the total cost associated with empty container repositioning in the above sources were slightly different, they lead to the same conclusion, i.e. the cost is huge and has become a burden to the container shipping industry. In particular, the profitability of shipping lines is highly dependent on whether, or not, the empty repositioning cost is redeemable. For example, it was reported that a shipping company after implementing an empty container logistics optimization system has saved cost US$ 81 million in year 2010 (Epstein et al. 2012).

The container transport chain can be broadly described as follows (Fig. 6.1). Consignors (shippers) are regarded as the customers who require empty containers to transport their cargoes. Shipping companies are usually responsible for providing the required empty containers to their customers. Empty containers may be stored in an inland depot or a sea port. After consolidating the cargoes into the containers at the customers' premise (or a depot or port), the laden containers will be transported to the depot or ports waiting for vessels. These laden containers are then lifted on a vessel in the booked shipping service. There may involve a couple of other shipping services for transshipment at sea ports before the laden containers finally reach the port of destination. Then the laden containers will be discharged from the vessel and transported to the consignees (cargo receivers) or a depot for unpacking. After unpacking, the empty container can either be moved/stored in an inland depot or a port for survey and reuse in the future, or be repositioned to other ports in the shipping networks to meet customer demands there. It can be observed that the container transport chain actually involves two supply chains: the forward supply chain of laden container flows, and the backward supply chain of empty container flows. A unique characteristic of container transport chains is that both laden and empty containers have to be moved and stored in the same shipping network using the same resources (e.g. vessels, trucks, trains, and facilities), which implies that these two supply chains are interwoven and difficult to separate. It should be noted that one important difference between laden container flows and empty container flows is that the former is driven externally by the customer demands whereas the latter is driven by the laden container flows and determined internally by the shipping companies themselves.

In a broad sense, the stakeholders in the container transport chain include shipping lines (including feeder operators), terminal operators, port authorities, depot operators, freight forwarders, inland transport operators (rail operators, road hauliers, barge operators), shippers or customers (consignors and consignees), container leasing companies, and others (e.g. associations, residents). In terms of empty container repositioning, the focal player is the shipping lines (ocean carriers), who usually bear the costs of repositioning empty containers and are responsible to transport both laden and empty containers at sea (port-to-port service) or even at inland as well (door-to-door service). It is therefore necessary to explain a bit more about shipping lines' business operations. A shipping line usually operates a number of shipping service routes, which form an inter-connected shipping service network. A shipping service route refers to a fixed sequence of ports, in which a fleet of container vessels is deployed to provide regular service (normally weekly service). These vessels make round-trips (voyages) along the service route repeatedly. A port may be called at more than once in a single round-trip. The shipping lines normally publish their service routes and schedules on the Internet several months before the actual voyages.

The empty container repositioning (ECR) problem concerns arranging the storage and movements of empty containers in the shipping networks in order to better position the movable resources to better satisfy customer demands. Effectively and efficiently repositioning empty containers has been a very important problem in shipping industry. It does not only have significant economic effect, but also an environmental and sustainability impact since the reduction of empty container movements would reduce the emissions along the container transport chain.

The rest of the chapter is organized as follows. In the next section, we discuss the main reasons that cause the empty container repositioning. In Sect. 6.3, we explain the empty container repositioning problem and provide a literature review with the emphasis on ECR models from the network scope, e.g. modeling ECR in seaborne transportation network, modeling ECR in inland or intermodal transportation network, and treating ECR as a sub-problem or a constraint under other decision-making problems. In Sect. 6.4, we discuss the solutions to the ECR problems from the logistics channel scope perspective. More specifically, the ECR problems could be tackled internally within an organization, externally in the vertical logistics channel, externally in the horizontal channels (i.e. collaboration with other shipping companies), and through technological innovations. In Sect. 6.5, we discuss the solutions to the ECR problems from the modeling technique perspective focusing on two main research streams based on recent studies. We present two specific models that represent the above two research streams. The models aim to tackle the ECR problem in stochastic dynamic environments considering both laden and empty container movements simultaneously. Finally, conclusions are drawn in Sect. 6.6.

6.2 Causes of ECR Problems

Empty container repositioning has been an on-going issue since the beginning of containerization. But it has become more prominent in the recent decades due to the rapid growth of container shipping business and the regional difference in economic development. This section will discuss the critical factors that cause empty container movements, which include the trade imbalance, dynamic operations, uncertainties, size and type of equipment, lack of visibility and collaboration within the transport chain, and transport companies' operational and strategic practices (Song and Carter 2009).

The fundamental reason for empty container repositioning is the *trade imbalance*, i.e. the trade in one direction is more than that in the other direction. The trans-Pacific and Europe-Asia routes are prominently imbalanced. Due to the China's economic boom in the last three decades, there is ever-increasing container traffic demand out of China, although the importing volume to China is also increasing. The United Nation publishes an annual review of maritime transport, which lists the estimated container flows on three major trade routes: Europe-Asia, Trans-Pacific (North America-Asia), and Trans-Atlantic (Europe-North America) routes. For example, the annual container trade demands for the years from 2007 to 2012 are summarized in Table 6.1.

It can be seen from Table 6.1 that the trade demands in the Europe-Asia and the Trans-Pacific routes were severely imbalanced. The volume in one direction was more than double of that in the opposite direction. This indicates the scale of empty container movements in the global context since the empty containers have to be moved from surplus areas to deficit areas.

The majority of the existing literature on empty container repositioning has explicitly emphasized the importance of considering trade demand imbalance in empty container allocation/repositioning. For example, Crainic et al. (1993a) indicated that empty containers are often repositioned between depots in order to overcome regional imbalances. They proposed the concept of container flow balancing in the context of inland transportation network between depots, customers and ports. Cheung and Chen (1998) stated that most international trades in liner shipping industry

Table 6.1 Containerized trade demands in three major shipping routes in million TEUs. (UN 2011, 2012, 2013)

	Eur-Asia		Trans-Pacific		Tran-Atlantic	
Year	Eur-Asia	Asia-Eur	Asia-NA	NA-Asia	NA-Eur	Eur-NA
2007	5.0	13.0	13.5	5.3	2.4	3.5
2008	5.2	13.5	13.4	6.9	2.6	3.4
2009	5.5	11.5	10.6	6.1	2.5	2.8
2010	5.6	13.5	12.8	6.0	2.8	3.1
2011	6.2	14.1	12.7	6.0	2.8	3.4
2012	6.3	13.7	13.3	6.9	2.7	3.6

are imbalanced in terms of the numbers of import and export containers due to the different economic needs in different regions. They focused on seaborne shipping network and proposed a single-commodity two-stage stochastic network model, in which the first stage is deterministic and aims to balance the empty container flows (including leasing empty containers) according to exogenous information of empty container supply and demand. Olivo et al. (2005) claimed that empty container movements would not exist in a perfect world as there would always be cargos to fill every container when and where it was emptied, but pointed out that in reality commercial traffic never seems to be in balance either in volume or value. They presented an integer programming model to balance container movements between ports and depots with multiple transport modes. Feng and Chang (2008) stated that the phenomenon of import-export imbalance is unavoidable in world trade and this results empty container problem in liner shipping industry. They studied the container balancing issue in an intra-Asia shipping network. Song and Carter (2009) considered the container balancing problem in three major trade lanes (Trans-Pacific, Trans-Atlantic, Europe-Asia) at the aggregated level, and analyzed four strategies that shipping companies could adopt to balance container flows depending on whether companies are sharing empty containers or coordinating empty containers among routes. Song and Dong (2011a) contrasted two types of container flow balancing policies, a point-to-point balancing policy, and a coordinated balancing policy. The policies were applied to a range shipping service routes with different topological structures to investigate their sensitivity to route structure and to the trade demand pattern.

The *dynamic operation* is the natural characteristic of any transport system since it covers different geographical locations and often requires transit time in weeks or months to access them. The impact of dynamic operations on empty container management may be understood from the perspective of the supply and demand of empty containers. The main supply source of empty containers is at the destinations of laden containers where they are discharged and unpacked and become empty containers for reuse, in particular at those import-oriented regions such as Europe and America. Note that the geographic locations of laden containers change over time in the shipping networks, the supply of empty containers therefore changes over time and space. On the demand side, the requirements of empty containers are driven by the trade demands, which also change over time for various reasons, e.g. seasonal products like agriculture produces, special festivals such as Christmas and Chinese New Year. These demand changes, although they may be predictable to a large degree, result in a dynamic impact on the container transport chain. The demands for empty containers and the arrivals of laden containers to be reused cannot match due to the time and space constraints and the volume difference. As a result, empty containers have either to be accumulated in advance to meet these expected increases in demand, or to be repositioned to the areas where empty containers are needed more urgently. The implication is that even if the overall laden container flows between two regions are balanced in long-term, the dynamic operations of the transportation system could be in favor of repositioning empty containers to order to improve container utilization. The impact of seasonality on the flow of empty containers has been confirmed by the empirical research in the Baltic Sea Region (Wolff et al. 2011).

Most of the existing literature on empty container repositioning has taken into account the dynamic operation explicitly or implicitly. Essentially, all studies that tackle the empty container repositioning problem at the operational planning level consider the dynamic nature of the environment explicitly. Examples of such studies (focusing on dynamic but deterministic situations) include: Shen and Khoong (1995); Lai et al. (1995); Olivo et al. (2005); Erera et al. (2005); Feng and Chang (2008); Erera et al. (2009); Di Francesco et al. (2009); Bandeira et al. (2009); and Song and Dong (2012).

The *uncertainty* is another key characteristic in container shipping, which represents the unpredictable elements that affect the container transport system. Uncertainty may occur during the operations in the container transport chain or during the interfaces with external environment. The former includes equipment breakdown, resource unavailability, port congestion, labor strikes (i.e. industrial action), bad weather (Notteboom 2006; Vernimmen et al. 2007). The latter includes random customer demands for empty containers and the instability of the political and economic environment (e.g. the financial crisis in 2008). The impact of uncertainty on empty container management may be explained as follows. For example, industrial action at a port may result in containers piled at the port and/or force container vessels to change their schedule. Weather conditions and traffic congestion may increase the transport time. As a result, these types of uncertainties cause either laden containers not to be delivered to customers on time, or empty containers not to be repositioned timely so as to meet the demands. Therefore, the movements of containers deviate from the plan and often incur extra container movements and costs. On the other hand, customer demand uncertainty probably has a more fundamental impact on container shipping operations. It is often the case that when shippers book the container in advance, often the day of pick-up is unpredictable. At present, liners tend to set up long-term contracts with big shippers, e.g. Maersk line with Argos. However, normally only the total volume within a period (e.g. a year) is specified in the contract whereas the detailed pick-up times of the shipments are unknown. Moreover, in the highly competitive shipping market, shippers have more choices and become more demanding. Therefore, it is extremely difficult for shipping lines to forecast the demands accurately. To accommodate the uncertainty in demands, shipping lines have to invest spare capacity, build up safety stocks, and reposition empties more efficiently. It has been illustrated that even in an overall balanced trade route, if the trade demands are uncertain, efficient empty repositioning could reduce the total cost significantly (Song 2007b).

Crainic et al. (1993a) is probably the first paper that addressed the empty container repositioning problem under stochastic/uncertain situations. Since then, a large number of studies have emerged in this line, for examples, Cheung and Chen 1998; Li et al. 2004, 2007; Lam et al. 2007; Song 2007a; Song and Dong 2008; Dong and Song 2009; Chou et al. 2010; Song and Dong 2011a; Yun et al. 2011; Epstein et al. 2012; Di Francesco et al. 2013.

Container *size and type* also affect the empty container repositioning. There are several different types of container that vary in their dimensions as well as the cargos they are designed to carry. The shortage of empty containers could happen because the size or types of available empty containers do not match customer requirements. Some regions such as Thailand may have a higher imbalance of reefer containers than dry containers. Even for dry containers, it has different grades including food grade, general purpose, and flexible grade. In shipping practice, normally 20-foot container is used for accommodating cargos with high volumetric mass density, while 40-foot one is used for cargoes with low volumetric mass density. Moreover, a full 40-foot container should not be 1.5 times heavier than a full 20-foot one in general. It has been observed that although some trade routes may not have significant trade imbalances, the need to transport empty containers may still be quite significant. One reason is that most types of cargo require, or it is more convenient to use, a specific type of containers (Branch 2000). Wolff et al. (2011) mentioned that the imbalance of container equipment could result from the fact that different goods types demanding for different equipment distinguished by dimension (e.g. TEU, FEU, high cube, pallet wide) and the specific application possibilities (e.g. reefers, tankers). Monios and Wilmsmeier (2013) analyzed highly disaggregated empirical data on container type movements and identified the container type diversification at UK ports, e.g. the use of high-cube and 45 ft pallet-wide maritime containers.

Rather limited literature has explicitly considered the size and type of containers when dealing with the empty container repositioning problem. Chang et al. (2008) allowed container substitution between different types in order to reduce the cost of empty container interchange under the street-turn and depot-direct schemes (i.e. empty containers can be directly distributed among customers without necessarily passing through container terminals). They considered a relatively compact transportation network, i.e. the Los Angeles/Long Beach port area. Wang (2013), which formulated a mixed-integer linear programming model for shipping network design and fleet deployment that took into account multi-type containers and empty container repositioning.

The *lack of visibility* of containers in the transport chain associated with the lack of collaboration between channel members in the supply chain is another reason to cause inefficiency in empty container management. International Asset Systems

(IAS) used the term, blind spot, to describe the situation when containers are moving via rail or truck, or while they are in inland terminals or at shipper/consigner premises (Song and Carter 2009). Blind spots in the transport chain may prevent shipping lines from tracking each container's location and status in real-time, thereby challenging liners' efforts to improve container utilization. In other words, without having timely and accurate information of container status and location, shipping lines are unable to manage their container fleet in the most effective way.

In the last decade, with the development of information and communication technology, auto-ID systems (e.g. barcodes, optical character recognition (OCR), radio frequency identification (RFID)) started being applied to maritime containers. For example, Savi Technology and Hutchison Port Holdings (HPH) formed a partnership in 2005 to use active RFID technology to track ocean shipping containers. It was reported 40 terminals worldwide were outfitted with Savi readers placed on cranes that load and unload ships, and at gates to track the movement of containers. The data were uploaded to a database hosted by Savi Networks (Roberti 2005). The main objective is to secure container terminals (using electronic seals) and meanwhile to add business value (providing information to shippers). Nevertheless, the Savi Networks was shut down in November 2010 according to WorldCargo news online. A few research papers have reported the application of auto-ID technologies to track and secure containers in container yards and terminals at ports (e.g. Lirn and Chiu 2009; Chao and Lin 2010; Rizzo et al. 2011; Acciaro and Serra 2013); however, there were little discussion on such technologies in relation to empty container repositioning.

Transport *companies' strategies and operational practices* actually determine empty container movements. Unlike the laden container movements that are largely determined by the shippers' requirements, empty container repositioning is an endogenous activity determined by shipping companies. Inappropriate or inefficient practices would lead to unnecessary empty container movements. It is not unusual that empty containers may be re-repositioned due to the vessel capacity constraints and the priority of laden container movements. Some shipping lines form an alliance in which they may share vessel slots. Willingness to exchange or share resources with other carriers can provide more opportunities for container reuse and reducing empty repositioning. Transport companies' strategies and operational practices, on the one hand, affect the actual movements of empty containers; on the other hand, act as the potential tools that the empty container repositioning problem could be tackled appropriately. This is the area that has attracted much attention in the last decade and extensive research has been carried out, which will be discussed in more detail in the next section.

Among all of the above factors, the trade imbalance is the root cause and accounts for the largest share for requiring empty container repositioning. This indicates that it is impossible to eliminate empty container movements completely in the real world. However, it has been recognized and demonstrated that through the development of innovative strategies and effective empty container repositioning policies, the costs and impacts associated with empty container repositioning can be reduced significantly. For example, Epstein et al. (2012) reported a cost saving of US$ 81 million for a shipping company after implementing the empty container logistics optimization system.

6.3 ECR Problems and Relevant Literature

Empty container repositioning problem aims to reposition empty containers efficiently and effectively in order to minimize the relevant costs while meeting customer demands for empty containers. This section will provide a literature review with the emphasis on modeling the ECR problems from the transport network scope.

An earlier literature review on empty container transportation was given in Dejax and Crainic (1987). They noted that “in spite of some very interesting problems with important practical applications, work in empty container allocation has still not integrated the latest methodologies created for the other modes (e.g. rail and truck) and has not yet generated any truly innovative modeling approach”. However, since 1990s, particularly in the last decade, numerous studies have been carried out in the area of empty container repositioning. Our literature will concentrate on those from 1990s.

A natural way to classify the empty container repositioning problems is based on the research scope of the underlying container transport networks. For example, the relevant literature may be classified into three groups according to the research context associated with the transport modes. The first group addresses empty container repositioning in seaborne shipping networks; the second group focuses on inland or intermodal transportation networks; whereas the third group tackles the empty repositioning problem as a sub-problem or a constraint under other decision-making problems.

In the first group, some studies consider a single shipping service route or a service network with specific route structure. For example, Lai et al. (1995) used a simulation model and some heuristic search methods to find cost-effective ways to reposition empty containers from Middle East ports to Far East ports in a Europe-Asia service route. Du and Hall (1997) proposed a threshold control policy to allocate empty equipment in a hub-and-spoke transport network. Li et al. (2004) and Song and Zhang (2010) established the optimality of the threshold-type inventory-based control policy in a single port subject to uncertain demands. Song (2007a), Lam et al. (2007) and Shi and Xu (2011) investigated the optimal empty container repositioning policies in two-port systems. Song and Dong (2008) developed threshold-type policies to reposition empties in cyclic service routes with uncertain demands. Li et al. (2007) and Zhang et al. (2014) extended the threshold control policy to multiple port systems. Feng and Chang (2008) presented a two-stage linear programming model for an intra-Asia shipping service route. Dong and Song (2009) employed the simulation-based optimization method and an inventory control based policy to deal with the joint optimization problem of container fleet sizing and empty container repositioning, in which the movements of both laden and empty containers and the constraints of vessel capacities are explicitly modeled. Chou et al. (2010) considered the empty container allocation problem in a single service route. A two-stage model is formulated. At stage one, a fuzzy backorder quantity inventory decision making model is proposed to determine the optimal quantity of empty container at a port; at stage two, an optimization mathematical programming network model is proposed to determine the optimal number of empty containers to be allocated between ports. Song and Dong (2011a) presented flow balancing-based empty repositioning policies in shipping service routes with typical topological structures. One advantage of focusing a specific structure of the service route is to provide opportunities to design optimal or near-optimal repositioning policies in stochastic situations. However, specific structure or a single service route simplifies the routing decisions and excludes the transshipment operations, which is an important phenomenon in container shipping operations.

On the other hand, some studies consider more general shipping networks. For example, Shen and Khoong (1995) optimized the flow of empty containers in a network with multiple ports over a planning horizon, in which vessels are not explicitly modeled. Cheung and Chen (1998) proposed a two-stage stochastic network model to allocate empty containers over a shipping network. They considered the random residual capacity for containers on the ships. Cheang and Lim (2005) developed a decision support system using a minimum cost flow model to distributing empty containers over a shipping network dynamically. The above three papers did not explicitly consider the topological structure of service routes and the regularity of vessel schedules. Erera et al. (2009) developed a robust optimization framework for dynamic empty repositioning problems modeled using time-space networks. They established the feasibility conditions of a repositioning plan and the recovery actions in response to uncertainties arising from forecasts of future container supplies and demands at different time epochs. Di Francesco et al. (2009) addressed the repositioning of empty containers in a scheduled maritime network. A multi-scenario multi-commodity time-extended optimization model is presented to minimize inventory, handling and transportation costs while meeting demand and supply requirements in every port. Moon et al. (2010) considered the empty container repositioning together with purchasing and short-term leasing options in a seaborne network. The problem is formulated as a deterministic multi-commodity model. A linear programming-based genetic algorithm and a hybrid genetic algorithm are proposed to solve the problem. Brouer et al. (2011) considered the laden container allocation and empty container repositioning for a liner shipping company. A multi-commodity time-expanded arc-flow model is formulated, which is then decomposed and solved with a delayed column generation algorithm. The model is able to handle large scale of shipping networks. Their work focuses on tactical planning without considering the details of transshipment between services. Song and Dong (2012) dealt the laden container routing and empty container repositioning at the operational level. A shortest-path based integer programming method and a heuristic-rules based integer programming method are proposed to solve the problem. The model assumes that there are at most twice transshipments for a laden shipment in the shipping network. Epstein et al. (2012) developed an empty container logistics optimization system (ECO) to support repositioning and stocking empty containers in a large shipping company. More specifically, the multi-commodity network flow model manages the repositioning problem, whereas an inventory model determines the safety stock required at each location. Long et al. (2012) formulated a two-stage stochastic programming model for the empty container repositioning problem in a maritime shipping network with uncertainties. The sample average approximation method and the progressive hedging strategy are applied to solve the optimization problem. Di Francesco et al. (2013) addressed the ECR problem in maritime networks under possible port disruptions. The problem is modeled by a time–space network and approximated by a multi-scenario model incorporating the non-anticipativity conditions.

In the second group, the studies focus on the empty container-repositioning problem in inland networks or intermodal transportation networks. The majority of the studies in this group focused on a regional scale. Braekers et al. (2011) conducted a comprehensive literature review on empty container management problems with the focus on the regional level, i.e. the empty container repositioning between importers, exporters, inland depots and ports within a small geographical area. More specifically, Crainic et al. (1993a, b) investigated the empty container allocation problem in the inland transport network in the vicinity of a seaport. Erera et al. (2005) developed a dynamic deterministic multicomodity network flow model for an intermodal transport network. They considered integrated container booking and routing decisions including empty repositioning. Olivo et al. (2005) proposed an integer programming model for empty container flows between container ports and depots across inland transportation network; Choong et al. (2002) investigated the effect of planning horizon length on empty container repositioning for an intermodal transport network. Bourbeau et al. (2000) presented a branch-and-bound parallelization strategy for the depot location and container allocation problems. Bandeira et al. (2009) proposed a heuristic method for integrated distribution of empty and full containers in an intermodal network. Yun et al. (2011) applied the (s, S)-type inventory control policy to reposition empty containers in an inland area between customers and terminals with random demands for empties. Simulation-based optimization tool is applied to find the near optimal (s, S) policy. Dang et al. (2013) extended the above work to a port area with multiple depots considering three types of decisions: repositioning empties from overseas ports, inland repositioning between depots, and leasing from lessors. The parameterized threshold policies are adopted for empty container repositioning and a simulation-based genetic algorithm is developed to optimize the threshold parameters. Lee et al. (2012) considered the joint empty container repositioning and container fleet sizing problem in a multi-port system, in which a single-level threshold policy is used to control the inventory and flow of empty containers among ports. Infinitesimal perturbation analysis method is applied to improve the computational efficiency. Because the formulation assumes that the travel time for each pair of ports is less than one period length and the shipping service routes are not explicitly considered, the model may be more appropriately regarded as a regional (inland or intermodal) network.

As intermodal networks are usually more complicated than seaborne shipping networks and the time-scale for inland transportation and sea transportation are significantly different, most of the above studies either focus on regional intermodal system (which is essentially an inland intermodal network) or treat container movements as flows and neglect individual vessels and their schedules.

The third group treats the empty container repositioning as a constraint or deals with it as a sub-problem within other decision making problems, e.g. dynamic empty container reuse (Jula et al. 2006; Chang et al. 2008), container fleet sizing with implicit empty container repositioning (Imai and Rivera 2001), transport market pricing and competition (Zhou and Lee 2009), shipping service route design (Shintani et al. 2007; Imai et al. 2009; Meng and Wang 2011a, Song and Dong 2013; Braekers et al. 2013; Wang 2013), ship fleet planning (Meng and Wang 2011b), and ship fleet deployment (Wang and Meng 2012).

Those joint optimization problems are often complicated. Most of them either use heuristics/meta-heuristics to tackle the problems or model the empty container repositioning in less detail to make it analytically tractable. Note that the motivation and focus of the studies in this group are often not directly from the empty container repositioning viewpoint; they might have been addressed in other chapters of the book.

6.4 ECR Solutions—the Logistics Channel Scope Perspective

In a broad sense, empty container repositioning problem covers any issues with the aim of mitigating the causes and the impacts of empty container movements and storage. From the logistics channel scope perspective, the ECR problems could be tackled internally within the shipping company, externally in the vertical logistics channel, externally in the horizontal channel (i.e. collaboration with other shipping companies), and through technological innovations. Accordingly, this section presents the solutions to the ECR problems under the following headings: organizational solutions, intra-channel solutions, inter-channel solutions, and technological solutions.

6.4.1 Organizational Solutions

Container fleet is a critical asset for an ocean carrier, which represents a large amount of capital. Empty container repositioning is a key component of the container fleet management, which includes a range of decisions such as fleet sizing, container leasing in/off, laden container routing, and empty repositioning. These decisions are highly related. For example, on one hand, increasing the number of owned containers, leasing extra containers and effectively repositioning empty containers can improve container's utilization and therefore equivalently increase the container fleet capacity. On the other hand, larger fleet size incurs capital and maintenance costs; container leasing-in and off-leasing incur extra leasing costs; while repositioning incurs additional handling and transportation cost. The interaction between laden container routing and empty container repositioning is obvious due to the facts that the laden container movements essentially drive the empty container movements, and both laden and empty containers are transported over the same network and carried by the same vehicles (vessel, train and truck).

Shipping lines are the focal company in the container transport chain, who often takes the responsibility to manage the empty container transportation. It is therefore understandable that the majority of the literature focusing on the ECR seeks internal organizational solutions from a single company perspective (explicitly or implicitly). Most of the literature in Sect. 6.3 belongs to this category. In the following, we try to link the literature on ECR to other components of the container fleet management such as fleet sizing, container leasing and laden container routing.

Container fleet sizing aims to determine how many owned containers should be kept in the fleet, which is a long-term decision since the life-time of a container is about 15 years. Mainly due to the different time scale and the complexity, only a few papers consider the combined problem of fleet sizing and empty container repositioning. Imai and Rivera (2001) presented an analytical model to address the fleet size problem for refrigerated containers where empty container movements are implied. Crainic et al. (1993a) investigated the container fleet sizing and empty allocation by focusing on the inland part of container transportation. Dong and Song (2009) optimized the container fleet size and the inventory-based empty repositioning policy simultaneously in a seaborne shipping network with zero inland transport time. Dong and Song (2012a) investigated the container fleet sizing problem in liner shipping services with uncertain customer demands and inland travel times, and quantified the impact of inland transport time on container fleet size.

Container leasing mainly concerns when and where to lease in/off empty containers, which itself is a complicated issue. Note that the ownership of the world container fleet is mainly split over ocean carriers and leasing companies (called lessors). The data from Containerization International shows that about 50–60 % of the world container fleet was owned by ocean carriers in the period from year 2001 to 2007 (Dong and Song 2012b).

There are generally two types of container leasing arrangement: master lease and term lease. The master lease is more of a service arrangement than a lease in which the customers can pick up and drop off containers according to agreed limits and locations without regard for how long the specific container has been under its control (Transport Trackers 2008). The leasing company is responsible for the full management of the containers including repositioning, storage, and maintenance. The idea behind master leasing is that the leasing company may turn around and re-lease a returned container to other parties quickly and the lessee can avoid repositioning costs. Term leases have fixed length of leases including short, medium and long terms, ranging from a single trip lease (also called spot leasing) up to eight-year terms. Under this type of arrangements, the lessee has the responsibility for repositioning and maintenance of the leased containers before reaching the fixed lease term and returning them to the lessor. Theofanis and Boile (2009) pointed out that there is a tendency that ocean carriers prefer long term leases over master leases so that they can integrate leased containers with their own equipment. Transport Trackers (2008) confirmed the decline in the use of the master lease and stated the main reason for the shift from master lease to term lease is that the premium for master leases plus the costs to lessees associated with off-leasing began to exceed the cost of hauling them back.

Although many studies have addressed the container leasing issues together with empty container repositioning, most of them consider it in an implicit way with the focus on empty container repositioning. For example, a common assumption is that: containers can be leased from lessors whenever owned containers are out of stock to meet customer demands; after leasing in, the leased containers are treated as the same as owned containers or can be returned to lessors at any future time (Crainic et al. 1993a; Lai et al. 1995; Cheung and Chen 1998; Lam et al. 2007; Song 2007a; Moon et al. 2010).

Laden container (or cargo) routing concerns the efficient flows of laden container in the shipping network to meet customer requirements. The origins and destinations of cargos are externally determined by the customers, but the physical path from the origins to the destinations could be either specified by shippers/freight forwarders or determined by the ocean carriers. Intuitively, the traditional shortest path methods could be applied to deal with cargo routing problem. Particularly, for simple networks such as a single specific route, the decision on cargo routing is straightforward and the laden container movements are often implied in the relevant ECR literature (c. f. the literature in the first group in Sect. 6.3). However, as the complexity of the shipping network increases, e.g. involving more service routes with multiple voyages, the cargo routing and its interaction with empty container repositioning become more complicated. For example, Crainic et al. (1993a) recognized the desirability of jointly optimizing laden and empty container allocation in a single mathematical model, but argued that it would be infeasible to solve given the intrinsic complexity of the problem. Most of the ECR literature dealing with general shipping networks (cf. the literature in the first group in Sect. 6.3) generally ignored the laden container routing and movements. Nevertheless, with the advance in linear and integer programming and the development of computing power in the last two decades, Erera et al. (2005) argued that the joint optimization of loaded and empty container allocation became feasible for a reasonable size of problems. A couple of papers have started to address the laden container routing and empty container repositioning simultaneously, which are discussed below.

Erera et al. (2005) formulated a large-scale multi-commodity flow model for global tank container operator by integrating container routing and empty repositioning in a single model. They confirmed the economic benefit of simultaneously considering laden and empty containers. However, their model did not consider the details of the shipping service routes and the vessel capacity was not modeled (assuming infinite shipping capacity). Bell et al. (2011) presented a frequency-based assignment model to allocate full and empty containers over shipping services by minimizing the sailing time plus container dwell time at the original port and any intermediate transshipment ports. Again the vessel capacity was not explicitly modeled. Brouer et al. (2011) studied the laden and empty container dynamic allocation problem for a liner shipping company explicitly considering the vessel capacity. A time-expanded multi-commodity flow model with additional inter-balancing constraints to control repositioning of empty containers was proposed. The aim is to maximize the profit of transported cargo subject to the cost of transport both laden and empty containers, leasing empties and rejecting demands. Their model captured the essential characteristics of the shipping networks at the tactical level, although the details of transshipment between services and the inventory of empty containers were not modeled. They demonstrated the feasibility of solving large-scale problems with simultaneously optimizing laden and empty container movements. The computational results confirmed the economic benefit of the joint planning. Song and Dong (2012) focused on dynamic operational-level planning and addressed the cargo routing and empty container repositioning in a multi-service multi-voyage shipping network in more details. The objective is to minimize the total relevant costs in the planning horizon including: container lifting on/off costs at ports, customer demand backlog costs, the demurrage (or waiting) costs at the transshipment ports for temporarily storing laden containers, the empty container inventory costs at ports, and the empty container transportation costs. Two solution methods are proposed to solve the optimization problem. The first is a two-stage shortest-path based integer programming method, which combines a cargo routing algorithm with an integer programming of the dynamic system. The second is a two-stage heuristic-rules based integer programming method, which combines an integer programming of the static system with a heuristic implementation algorithm in dynamic system. They assumed that the laden container routing from the original port to the destination port is limited with at most three service routes in order to reduce the complexity of the cargo routing sub-problem.

6.4.2 Intra-Channel Solutions

The container transport chain consisting of consignor, shipping line, terminal operator, inland transport operator, depot operator, and consignee can be regarded as a vertical channel from the supply chain viewpoint. Intra-channel solutions emphasize on the coordination (including improving visibility, planning collaboratively, and achieving intermodalism) across different players in the vertical channel, which is a natural extension to the organizational solutions.

The literature in the second group in Sect. 6.3, to some extent, attempts to seek intra-channel solutions explicitly or implicitly using modeling techniques. They mainly focus on the coordination of empty container management in a regional area among terminals, depots, and customers with the assumption that information visibility can be realized and a single objective can be defined (e.g. Crainic et al. 1993a, b; Bourbeau et al. 2000; Olivo et al. 2005; Choong et al. 2002; Bandeira et al. 2009; Yun et al. 2011; Dang et al. 2013).

Apart from the modeling research, empirical concepts and practices of intra-channel solutions have also emerged in the last decade. “Street turns” or “Empty reuse” refers to reusing import containers for export loads at the consignee’s site or in its proximity where direct exchange of empty containers between consignee and consignor can be realized. The potential benefits of street turn include: (i) truck trips to and from the port can be saved; (ii) the haulier can generate more revenue in less time; (iii) the ocean carrier can save paperwork and improve the container utilization;

(iv) the export customer gets the empty container sooner; (v) environmental impact can be reduced, i.e. traffic, congestion, noise, and emissions (Tioga Group 2002). However, there are some challenges and barriers to implement street turns such as: (i) the haulier must identify the opportunity for reuse and communicate the opportunity to the driver; (ii) the agreement between the haulier and the ocean carrier must allow for such reuse and the ocean carrier must be able to track and document the interchange between parties; (iii) the place of the emptied import container should be reasonably close to the next exporter, and its available time should match the loading time window for exporting; (iv) the emptied import container must be in good condition and suitable for the export load, and the container/chassis combination must be acceptable at the terminal used by the export vessel (Tioga Group 2002). From the theoretical aspect, Jula et al. (2006) analyzed the potential cost and congestion reductions through the reuse of empty containers in the Los Angeles and Long Beach port area.

The concept of "off-dock empty return depot" refers to establish a neutral point to serve as buffer storage for container interchange and reuse. Empty containers would first accumulate at an off-dock empty return depot for cleaning, maintenance and repair, and then be reused for local exports or sorted and returned to a marine terminal at off-peak hours. This concept would add extra capacity to the maritime terminal and facilitate empty returns when terminal gates are closed (Tioga Group 2002; Hanh 2003).

Another concept is "depot-direct off-hire", which refers to the process of off-hiring and repositioning an empty container to the leasing company at an inland depot directly before returning to the maritime terminal. This concept would cut at least one truck trip from each off-hiring and repositioning cycle when considering the trips of container and chassis movements among consignee, maritime terminal and inland depot (Tioga Group 2002). While "street turn" and "off-dock empty return depot" emphasize on the coordination between customers, shipping lines, depot and terminal operators, the concept of "depot-direct off-hire" focuses on the coordination between hauliers, depot operators, shipping lines and leasing companies.

The contractual relationship between ocean carriers and inland transport companies in terms of repositioning empty containers can take quite different formats. Lopez (2003) investigated the organizational choices of ocean carriers to reposition their empty containers in the USA. Four organizational formats were discussed including spot contract with road hauliers, one-year contract with rail operators, renewable contracts with road hauliers, and renewable contracts with intermodal marketing companies. It is observed that ocean carriers do not think about transaction costs, but they do adopt some mechanisms (e.g. renewable contracts) to control and to adjust their transactions in order to reduce those costs.

Van Der Horst and De Langen (2008) discussed the coordination issues among the players in the hinterland transport chain including shipping lines, terminal operators, forwarders, hinterland transport companies, and inland depot operators. They found that the development of the coordination in practice was hindered by a lack of contractual relationships, information asymmetry, and a lack of incentives for co-operation. They proposed four coordination mechanisms including introduction of incentives, creation of an inter-firm alliance, changing the scope of the relation and management, creating collective action. One benefit of the coordination between the terminal operators, hinterland depots, and shipping lines is to reduce empty movements.

Wolff et al. (2012) conducted a questionnaire survey to gain an empirical picture of different players in container transport chain dealing with empty containers in the Baltic Sea Region. It was found that the share of "street turns" in practice was in a range of 5–10 % in Hamburg. In terms of backhaul of empty containers, shipping lines prefer to have empty inventories and even depot services directly on the terminal so that they can move their container fleets more flexibly and decrease the throughput time. A range of measures to tackle empty container management were identified including: managerial and organizational measures (e.g. using spare capacities on vessel/vehicle of the own fleet; searching for return cargo; use container pooling; use spare capacities on the vessel/vehicle of other operators' fleet; network design of empty container depots), pricing measures (e.g. selling empties in the surplus and buy new in the deficit area; freight rate surcharge on the high demand transport leg), ICT measures (e.g. use RFID to track and trace containers; use virtual container yards; use online market), and technological measures (e.g. implement foldable containers). It is concluded that no one single measure has a crucial positive impact on empty container management, a combination of measures is more promising, the success and choice of measures are highly player dependent.

In the past decades the container terminal industry has gone through the vertical integration process. For example, shipping lines have invested in terminal operations directly or through parent companies. Most global shipping lines have now owned the dedicated container terminals in various regions, which enables them not only managing the ships more effectively but also the empty container logistics. Therefore, establishing dedicated container terminals could be regarded as an intra-channel strategic measure to tackle empty container repositioning problems.

6.4.3 Inter-Channel Solutions

In container shipping industry, many container transport chains co-exist. For example, there were more than 400 shipping lines in the world (Song et al. 2005) and each of them may be involved in multiple container transport chains. The container management strategies across parallel container transport chains are classified as inter-channel solutions.

Container shipping industry is very unique in terms of the popularity of horizontal integration. Although shipping lines are the competitors as service providers, they also collaborate in various formats such as alliances, slot exchange, and resource pooling.

In the last decade, we have seen the emergence of external collaboration among carriers to achieve effectiveness of container operations and reduce costs. A few third or fourth logistics parties emerged to provide internet-based support. These systems can serve as a neutral platform to facilitate container sharing among shippers, forwarders, and shipping lines. The idea is gaining increasing popularity, however "*There are still pockets of resistance, but the search to reduce costs outweighs the resistance to sharing containers*" (Mongelluzzo 2004). A few examples are introduced below.

SynchroNet, founded in 1996, has developed a neutral global container management tool, termed "s|InterChange". The system enables the registered shipping companies to interchange containers between parties on an inter-continental or intra-theater level and reposition surplus containers economically to deficit areas (www.synchronetmarine.com).

International Asset System (IAS) developed a neutral platform (termed IAS InterChange) that enables ocean carriers, container lessors and NVOCC (Non-Vessel Operating Common Carrier) to interchange containers in surplus and deficit locations. The registered customers provide ISA with the data of their equipment inventory and the InterChange will match between equipment suppliers and receivers in order to avoid costly repositioning. ISA also developed another service product, called SlotXchange. This tool is able to match empty containers with available slot space on ocean-going vessels. With SlotXchange, equipment owners can quickly reposition empty containers to the destination location, whereas the vessel operators with empty space can generate additional freight revenue by offering the empty slots. (www.interasset.com).

From the modeling aspect, Song (2007b) provided a theoretical analysis to a collaborative strategy in shuttle transport systems with uncertain demands. The dynamic programming model quantifies the cost saving of the collaborative strategy under different container dispatching policies. It is identified that the factors such as the container fleet size, the variance of demands, the demand patterns (balanced or imbalanced), and the container dispatching policy have significant impacts on the performance of the collaborative strategy. For example, the collaborative strategy can achieve more cost saving in situations with smaller fleet size or higher degree of uncertainty. It is reported that the cost savings are greater than 20 % in many cases, particularly when two companies have complementing demand patterns. On the other hand, if two companies have relatively large fleet sizes, low degrees of demand uncertainty, and similar patterns of imbalanced demands, then the collaborative strategy can only achieve rather limited cost saving. This might be one of the reasons that major shipping lines are reluctant to share containers with others in severely imbalanced routes such as Asia-Europe and Trans-Pacific.

Song and Carter (2009) further analyzed the inter-channel strategies to balance container flows at the global scale. According to whether shipping lines are coordinating the container flows over different service routes and whether they are willing to share container fleets with other companies, four strategies are defined for empty container repositioning: container-sharing and route-coordination; container-sharing without route-coordination; route-coordination without container-sharing; and neither container-sharing nor route-coordination. Here route coordination refers to ocean carriers acting as a single firm to balance its container flows across different service routes. Container sharing refers to pooling container fleets among different ocean carriers. The results show that route coordination offers more opportunities to reduce empty repositioning costs than container sharing in the container industry, which may further explain the reluctance of large carriers to adopt container-sharing practices.

Vojdani et al. (2013) formulated a space-time network model to evaluate the economic benefit of container pooling by several container carriers and container leasing companies. Numerical examples with three carriers, multiple routes, and multiple ports are provided to illustrate the positive influence on cost reduction compared with non-cooperative scenarios.

Liu et al. (2013) proposed a multi-commodity network flow model in a multi-carrier scenario and provided a cooperative game for container sharing among carriers. The issue of the cost/profit allocation mechanisms is addressed in relation to the format of container sharing mechanism.

Container transport chain is closely related to other supply chains such as manufacturing and purchasing channel, recycle channel, and secondary market channel. The International Institute of Container Lessors (IICL), whose member companies represent approximately 90 % of the container leasing industry and about 40 % of the world's chassis, reported that the amount of container dispositions in 2009 was 530,485 TEUs and the estimated new purchase in 2010 were approximately 600,000 TEUs (IICL 2010). Inter-channel solutions can also be developed by linking the empty container repositioning issue with the management of those supply chains.

6.4.4 Technological Solutions

Technology development and innovations facilitate the development of organizational solutions, intra-channel solutions, and inter-channel solutions. On the other hand, technological innovations could offer a complete new set of solutions to the ECR problems, which may contribute directly to the cost reduction of the empty container transportation.

Note that the solutions to ECR problems from the previous few sections (particularly intra-channel solutions and inter-channel solutions) all depend on the support of information communication and technology. To enable channel members to collaborate to deal with the ECR problem together, a pre-requisite is to ensure the container logistics visibility to the relevant channel members. In practice, various players in the container transport chain have their own tracking system. For example, RFID technology has been used in maritime terminal to track the movement of containers inside the terminal (e.g. Roberti 2005; Lirn and Chiu 2009; Chao and Lin 2010; Rizzo et al. 2011; Acciaro and Serra 2013). Container haulage companies have GPS systems attached to their trucks to identify their locations and the containers they are carrying. Shipping lines have GIS/GPS systems to track the geographic location of the ships and the containers on board. Therefore, in theory it is possible to know whether a container is on board, in maritime terminal, in inland depot or at customers' premises. This would help shipping lines to remove the blind spot in the inland transport chain. However, because of the concern that the release of the data may be misused by other parties and may not be advantageous, companies usually keep the information proprietary. The visibility of container logistic flow is still low in the current practice.

Supply chain integration either vertically or horizontally can only be achieved by the application of information technologies. As a higher level of control over container flows are established, the need for electronic data interchange (EDI) becomes essential. Timely and accurately information exchange between supply chain members can reduce the degree of uncertainty and offer more opportunities to manage the container fleet. Recent years have also seen IT become fundamental for security issues (E-manifest) and have incited the industry to move forward as a matter of compliance to advance notice schemes for the cargo being carried (Van Der Horst and De Langen 2008).

Although the ideas behind the internet-based platforms such as "s|InterChange", "IAS InterChange" and "IAS SlotXchange" are essentially intra-channel or inter-channel solutions, their implementation highly relies on technology development.

Foldable (collapsible) container is a technological innovation to move empty containers more efficiently. It could greatly reduce the number of lifts and moves of empty containers at maritime terminals, and storage space on board. Several foldable (collapsible) container designs have been developed. Fallpac AB developed a Fallpac container in which four units can be folded and stacked inside a fifth erected unit. This means that a package of five empty containers occupies the space of a single standard container (Konings and Thijs 2001; Moon et al. 2013). The Six-in-One Container Company introduced a six-in-one container where six containers can be folded, bundled and interlocked to the exact dimensions of a single standard container (Konings and Thijs 2001). This implies that six empty containers can be treated as one container when loading/unloading at terminals and storing on board. Staxxon has designed a folding shipping container that can be folded vertically, shrinking to as much as one-fifth their normal size. Set side by side, five containers occupy the space of a single standard container. Staxxon is starting to test its model at terminals and believes that it has the potential to be the folding container that finally convinces shippers to start switching over (http://staxxon.com/). Moon et al. (2013) reported that foldable containers are currently under development by Holland Container Innovations and Cargoshell in the Netherlands and Compact Container Systems in the US.

Theoretically, several studies have been conducted to analyze and evaluate the potential application of foldable containers in the real world. Konings and Thijs (2001) discussed several conditions that are necessary for the successful commercial applications of foldable containers. Relevant issues include the folding/unfolding complexity and cost, the production cost, the technical features of foldable containers, the choice of the logistic concept, and product marketing. Konings (2005) further analyzed the opportunities for the commercial application of foldable containers and performed more detailed cost-benefit analysis in four logistic conceptual scenarios of using foldable containers to improve empty container repositioning: port-to-port, continent-to-continent, export depot-to-import depot, door-to-door scenarios. It is reported that the use of foldable containers can lead to substantial net benefits in the total container transport chain, but also much depends on the additional costs that foldable containers may incur such as the cost of folding/ unfolding, additional exploitation costs and any additional transport to places where folding and unfolding can take place. Shintani et al. (2010) evaluated the cost savings of using foldable container in the hinterland to reposition empty containers. Based on the possible movement of empty containers and the locations available for folding and unfolding activities, three unique scenarios were proposed for investigation. Moon et al. (2013) further explored the potential cost savings by using foldable containers for repositioning empty containers at sea transport networks.

Other aspects of technological innovations such as using more efficient quay cranes and new materials to constructing containers may also contribute to the cost reduction of empty container repositioning.

6.5 ECR Solutions—the Modeling Technique Perspective

Broadly speaking, the ECR modeling studies may be categorized into two research streams according to the applied modeling techniques and the type of the proposed solutions. The first stream adopts the network flow models and often applies mathematical programming to produce a set of arc-based matrices. The element in each matrix is a numerical value representing the quantity of empty containers to be moved on an arc (i.e. from one node to another node) in the network. Examples of the studies in this group include: the application of linear programming (Dejax and Crainic 1987; Shen and Khoong 1995; Bourbeau et al. 2000; Choong et al. 2002; Erera et al. 2005; Olivo et al. 2005; Cheang and Lim, 2005; Song and Carter, 2009; Song and Dong 2011b), stochastic programming (Crainic et al., 1993a; Cheung and Chen 1998; Erera et al. 2009), scenario-based linear programming (Di Francesco et al. 2009), sample average approximation based linear programming (Long et al. 2012), and multi-scenario mixed-integer programming (Di Francesco et al. 2013).

The second stream aims to develop effective state-feedback control policies which often uses inventory control, dynamic programming, and simulation-based optimization methods (e.g. Li et al. 2004; Song 2005; Song 2007a; Lai et al. 1995; Li et al. 2007; Lam et al. 2007; Song and Dong 2008; Dong and Song, 2009; Yun et al. 2011; Lee et al. 2011; Song and Dong 2011b; Dang et al. 2013; Lee et al. 2012; Zhang et al. 2014). The solutions of these empty container repositioning polices are similar to those in inventory control in production systems, and they normally consist of a number of decision-making rules associated with system dynamic states such as inventory levels of empty containers. By applying the rules at a decision epoch, the number of empty containers that need to be repositioned out or into a node can be determined dynamically. Several inventory-based control policies have been proposed in the literature; e.g., the double threshold policy (Li et al. 2004; Li et al. 2007; Song and Dong 2008; Dong and Song 2009; Song and Zhang 2010; Zhang et al. 2014), the dynamic port-to-port balancing policy (Dong and Song 2009; Song and

Dong 2012), the coordinated (s, S) repositioning policy (Dang et al. 2013), and the single-level threshold policy (Song 2005; Song 2007b; Lee et al. 2012). It needs to be noted that each inventory-based control policy could have a number of variations depending on the way of splitting empty containers over ports. For example, Song and Dong (2011b) has proposed two variations of the double threshold policy termed as flexible destination port policy and determined destination port policy. The concept of flexible destination port repositioning was also adopted in Di Francesco et al. (2013).

There are also a couple of attempts to combine both the inventory model and mathematical programming model to solve the ECR problems. Chou et al. (2010) proposed a mixed inventory decision-making and mathematical programming model for dealing with the ECR problem. In stage one this paper proposes a fuzzy backorder quantity inventory model for determining the optimal quantity of empty containers at a port considering stochastic import and export at the same time. In stage two, an optimization mathematical programming network model is proposed for determining the optimal number of empty containers to be allocated between ports, which is based on the results for the fuzzy backorder quantity inventory model in stage one. The utilization of the proposed model is demonstrated with a case of trans-Pacific liner route in the real world. However, they focus on a single service route. Epstein et al. (2012) initially planned to develop a single, integrated, and robust optimization model that would address the ECR optimization problem with uncertainties, but realized that the time required finding an optimal solution was too long even for small instances. They then opted for developing a two-stage solution approach, which combines a network flow model and an inventory model, termed empty container optimization (ECO) tool. The ECO tool is based on two decision models supported by a forecasting system. At stage one, an inventory model takes into account the uncertainty in container supply and demand and determines the safety stock for each node in the network. At stage two, a multi-commodity multi-period network flow model addresses the imbalance problem and supports daily empty container repositioning and inventory levels. The service level is managed by imposing the safety stock as constraints in the network flow model with the assumption of normal distributions of the forecast demand. In addition, the ECO tool uses a collaborative web-based optimization framework to address the coordination problem among multiple agents with local objectives. However, both papers Chou et al. (2010) and Epstein et al. (2012), only focused on empty container logistics. The movements and routing of laden containers were not considered.

According to earlier discussions, the most important three reasons to cause empty container movements are probably trade imbalance, dynamic operations, and uncertainties. In particular, trade imbalance is the fundamental reason. Therefore, to model the empty container repositioning problem appropriately, it is desirable to model both laden container routing and empty container repositioning in the transport network simultaneously, because trade imbalance is represented by laden container movements whereas laden container movements are determined by the laden container routing.

In the remainder of this section, we present a few specific mathematical models for empty container repositioning problems (with a focus on maritime transport networks), which represent the above mentioned modeling techniques. All models consider both laden and empty container movements simultaneously. Some of them can also handle dynamic operations and stochastic environments. We make two common assumptions:

Assumption 1 all the containers and customer demands are measured in TEUs. One FEU (forty-foot equivalent unit) is treated as two TEUs.

Assumption 2 the vessels deployed in the same service route have the similar carrying capacity.

6.5.1 *Time-Space Multi-Commodity Network Flow Model for Laden and Empty Container Management*

This section introduces a time-space multi-commodity network flow model to deal with empty container repositioning problem, which is mainly based on Brouer et al. (2011). The model considers both laden and empty container flows in the shipping network over a given planning horizon. The customer demands are deterministic, but can take different values at different time periods.

We introduce the following notations for the model in this sub-section:

P	the set of ports;
R	the set of shipping routes;
T	the planning horizon;
G	a capacitated directed acyclic graph, $G: = (N, A)$;
N	the set of nodes, $N = \{p^t \mid p \in P; 0 \le t \le T\}$;
A	the set of arcs, $A := A_G \cup A_R$;
A_G	the set of uncapacitated ground arcs, $A_G := \{(p^t, p^{t+1}) \mid p^t \in N; p^{t+1} \in N\}$;
A_R	the set of capacitated sea leg arcs $A_R: = \{(p^t, q^{t+\tau_{pq}}) \mid p^t \in N; q^{t+\tau_{pq}} \in N; p \neq q; u(p^t, q^{t+\tau_{pq}}) > 0\}$, where τ_{pq} is the travel time from port p to port q; and $u(.,.)$ represents the aggregated capacity of the corresponding arc to be defined a bit later.
Cap_r	the vessel capacity in the route $r \in R$;
A^r	the set of sea leg arcs in the route $r \in R$ over the planning horizon, i.e. $A^r: = \{(p^0, q^{0+\tau_{pq}}), \ldots, (r^t, o^{t+\tau_{ro}}) \mid p, q, r, o \in P\}$, where $t + \tau_{ro}$ can be regarded as T (more precisely, it refers to the latest time period before T when one of the vessels deployed in route r is berthing at a port);
$u(i, j)$	the aggregated shipping capacity of the sea leg arc $(i, j) \in A_R$, i.e. $u(i, j): = \sum_{r \in R} \sum_{(i,j) \in A^r} Cap_r$. Namely, $u(i, j)$ is the accumulated vessel capacity of all service routes that have a voyage covering the sea leg arc (i, j).
K	the set of commodities to be transported in the shipping network; a commodity $k \in K$ is represented by (O_k, D_k, d_k), where $O_k \in N$ denotes the origin node, $D_k \in N$ denotes the destination node, and d_k denotes the volume of the commodity (i.e. the number of containers)

y_{ij}^k	the number of laden containers of commodity k on arc (i, j);
x_{ij}	the number of empty containers on arc (i, j);
C_{ij}^k	unit cost of arc (i, j) for commodity $k \in K$;
C_{ij}^e	unit cost of arc (i, j) for empty containers;
C_p^k	unit penalty cost for lost-sales of commodity $k \in K$;

The objective is to minimize the sum of the laden container transportation costs, the empty container transportation costs, and the lost-sale penalty cost. The decision variables include the laden container flows, i.e. y_{ij}^k, and the empty container flows, i.e. x_{ij}.

$$\min_{y_{ij}^k, x_{ij}} \left\{ \sum_{k \in K} \sum_{(i,j) \in A} C_{ij}^k y_{ij}^k + \sum_{(i,j) \in A} C_{ij}^e x_{ij} + \sum_{k \in K} C_p^k \left[d_k - \sum_{j \in N, i = O_k} (y_{ij}^k - y_{ji}^k) \right] \right\} \tag{6.1}$$

subject to

$$\sum_{j \in N} y_{ij}^k - \sum_{j \in N} y_{ji}^k \quad \leq d_k, \text{for } i = O_k, k \in K; \tag{6.2}$$

$$\sum_{j \in N} y_{ij}^k - \sum_{j \in N} y_{ji}^k = \sum_{j \in N} y_{jm}^k - \sum_{j \in N} y_{mj}^k, \text{for } i = O_k, m = D_k, k \in K; \tag{6.3}$$

$$\sum_{j \in N} y_{jm}^k = \sum_{j \in N} y_{mj}^k, \text{ for } m \in N, m \neq O_k, m \neq D_k, k \in K; \tag{6.4}$$

$$\sum_{k \in K} \sum_{j \in N} y_{ij}^k + \sum_{j \in N} x_{ij} = \sum_{k \in K} \sum_{j \in N} y_{ji}^k + \sum_{j \in N} x_{ji}, \text{for } i \in N; \tag{6.5}$$

$$x_{ij} + \sum_{k \in K} y_{ij}^k \leq u(i, j), \text{for } (i, j) \in A; \tag{6.6}$$

$$y_{ij}^k \geq 0, x_{ij} \geq 0, \text{for } k \in K, (i, j) \in A; \tag{6.7}$$

The constraint (6.2) represents the satisfied demand of commodity k cannot exceed the volume d_k. Constraint (6.3) represents that the same amount of commodity k will be moved out of node O_k and moved into node D_k. Constraint (6.4) represents the flow conservation of commodity k at a node m that is neither O_k nor D_k. Constraint (6.5) represents flow balancing at any node considering both laden and empty container movements. Constraint (6.6) ensures that the total flows including both laden containers and empty containers on any arc do not exceed the shipping capacity of the arc. Constraint (6.7) is the non-negative requirements for the decision variables. More accurately, we should let the decision variables only take integers. Therefore, the above model is a linear integer programming model, which can be solved using commercial software such as IBM ILOG CPLEX.

The advantages of the above model include: (i) the formulation of the model is relatively simple and easy to understand; (ii) the empty container movements are derived from laden container movements, which reflects the reality; (iii) the model can handle variable demands over different time periods because a planning horizon is introduced; (iv) the lifting-on costs at the commodity's origin port, and lifting-off costs at the commodity's destination port can be easily incorporated. However, there are some drawbacks with the model: (i) although transshipments are modelled over the time-space shipping network, the associated costs are not included in the objective function. This may result in unnecessary or uneconomical transshipment in the solutions; (ii) the actual path that the commodity moves in the shipping network (including the information such as which service routes to use, which ports to transship) is not easy to identify. Brouer et al. (2011) reformulated the problem into a path-based network flow model, in which a path of commodity k consists of a sequence of arcs that connect from node O_k to node D_k. This helps to identify the flow of commodity on the arcs from its original port to destination port. Nevertheless, because the arcs are not associated with service routes, it is still not obvious to identify which specific service routes that carry the commodity in the path; (iii) the number of commodities could be very large in realistic scenarios, which may become computationally intractable.

6.5.2 *Origin-Link Based Network Flow Model for Laden and Empty Container Management*

As transshipment is a very important phenomenon in container shipping industry, particularly for the hub ports (such as Singapore, Hong Kong, Rotterdam, Busan), where transshipment traffic could account over 50 % of their total throughput, this section presents another network flow model that takes into account the transshipment costs and manages both laden and empty containers simultaneously.

We make the following assumptions in this section: (i) all service routes are of weekly frequency; (ii) the weekly demands for any O-D pair are constant; (iii) it is at the tactical planning level.

We adapt the origin-link-based linear programming model to managing the flows of both laden and empty containers in a shipping network. The idea of the origin-link-based linear programming model has been applied to shipping network design and ship deployment (e.g. Alvarez 2009; Wang and Meng 2012; Wang 2014).

The following notations are introduced for the model in this sub-section.

P	the set of ports;
R	the set of shipping routes;
R_p	the set of routes that call at port $p \in P$;
N_r	the number of portcalls in the route $r \in R$;
I_r	the set of portcall indices in the route $r \in R$, i.e. $I_r := \{1, 2, \ldots, N_r\}$;
p_{ri}	the port that corresponds to the ith portcall in route r;
$I_{r,p}$	the set of portcall indices corresponding to port p in the route $r \in R$, i.e. $I_{r,p} := \{i \in I_r \mid p_{ri} = p\}$;

$y^l_{o,ri}$	the number of laden containers originating from $o \in P$ that are loaded at the ith portcall in the route r;
$y^u_{o,ri}$	the number of laden containers originating from $o \in P$ that are unloaded at the ith portcall in the route r;
$y^f_{o,ri}$	the number of laden containers originating from $o \in P$ that are carried on board on the leg i (from ith portcall to $i + 1$th portcall) in the route r;
y_{od}	the fulfilled demands from $o \in P$ to $d \in P$;
x_p	the number of empty containers to be repositioned out of $p \in P$ (into port p if it is negative);
$x^l_{o,ri}$	the number of empty containers originating from $o \in P$ that are loaded at the ith portcall in the route r;
$x^u_{o,ri}$	the number of empty containers originating from $o \in P$ that are unloaded at the ith portcall in the route r;
$x^f_{o,ri}$	the number of empty containers originating from $o \in P$ that are carried on board on the leg i (from ith portcall to $i + 1$th portcall) in the route r;
D_{od}	the weekly demands from $o \in P$ to $d \in P$;
Cap_r	the vessel capacity in the route $r \in R$;
C^l_p	unit cost of loading containers at port $p \in P$;
C^u_p	unit cost of unloading containers at port $p \in P$;
$C^{t,l}_p$	unit cost of transshipping laden containers at port $p \in P$;
$C^{t,e}_p$	unit cost of transshipping empty containers at port $p \in P$;
C^p_{od}	unit penalty cost for lost-sales from $o \in P$ to $d \in P$;
$C^l_{r\,i}$	unit cost of transporting laden containers on vessel in leg i in the route $r \in R$;
$C^e_{r\,i}$	unit cost of transporting empty containers on vessel in leg i in the route $r \in R$;

The objective is to minimize the sum of the laden and empty container loading (lifting-on) cost, the laden and empty container unloading (lifting-off) cost, the laden and empty container transshipment cost, the lost-sale penalty cost, the laden container transportation cost on vessel, the empty container transportation cost on vessel. The decision variables include the laden container flows, i.e. $y^l_{o,ri}$, $y^u_{o,ri}$, $y^f_{o,ri}$, y_{od}, and the empty container flows, i.e. x_p, $x^l_{o,ri}$, $x^u_{o,ri}$, $x^f_{o,ri}$.

To simplify the narrative, we introduce a few intermediate variables. Let y^l_p, y^u_p, y^t_p denote total number of laden container loading (including export from the port and the transshipment), total number of laden container unloading (including import into the port and the transshipment), and the number of laden container transshipment at port p. Similarly, let x^l_p, x^u_p, x^t_p, denote total number of empty container loading (including repositioning out from the port and the transshipment), total number of empty container unloading (including repositioning into the port and the transshipment), and the number of empty container transshipment at port p. The linear programming model is given by,

$$\min_{\substack{y_{od},y^l_{o,ri},y^u_{o,ri},y^f_{o,ri},y^l_p,y^u_p,y^t_p\\ x_p,x^l_{o,ri},x^u_{o,ri},x^f_{o,ri},x^l_p,x^u_p,x^t_p}} \left\{ \sum_{p\in P}\left[C^l_p(y^l_p+x^l_p)+C^u_p\left(y^u_p+x^u_p\right)+C^{t,l}_p y^t_p+C^{t,e}_p x^t_p\right] \right.$$
$$+\sum_{r\in R}\sum_{i\in I_r}\left(C^l_{ri}\sum_{o\in P}y^f_{o,ri}+C^e_{ri}\sum_{o\in P}x^f_{o,ri}\right)$$
$$\left. +\sum_{o\in P}\sum_{d\in P}C^p_{od}(D_{od}-y_{od})\right\} \tag{6.8}$$

subject to

$$y_{od}\le D_{od}, \text{for any o}, d\in P; \tag{6.9}$$

$$\sum_{r\in R_o}\sum_{i\in I_{r,o}}y^u_{o,ri}=0, \text{for any o}\in P; \tag{6.10}$$

$$\sum_{r\in R_o}\sum_{i\in I_{r,o}}y^l_{o,ri}=\sum_{p\in P}y_{op}, \text{for any } o\in P; \tag{6.11}$$

$$\sum_{r\in R_p}\sum_{i\in I_{r,p}}(y^u_{o,ri}-y^l_{o,ri})=y_{op}, \text{for any o}, p\in P, o\ \neq p; \tag{6.12}$$

$$y^l_p=\sum_{r\in R_p}\sum_{i\in I_{r,p}}\sum_{o\in P}y^l_{o,ri}, \text{for any } p\in P; \tag{6.13}$$

$$y^u_p=\sum_{r\in R_p}\sum_{i\in I_{r,p}}\sum_{o\in P}y^u_{o,ri}, \text{for any } p\in P; \tag{6.14}$$

$$y^t_p=y^l_p-\sum_{d\in P}y_{pd}=y^u_p-\sum_{o\in P}y_{op}, \text{for any } p\in P; \tag{6.15}$$

$$y^f_{o,ri}=y^f_{o,ri-1}-y^u_{o,ri}+y^l_{o,ri}, \text{for any } o\in P, r\in R, i\in I_r; \tag{6.16}$$

$$x_p=\sum_{o\in P}y_{op}-\sum_{d\in P}y_{pd}, \text{for any } p\in P; \tag{6.17}$$

$$\sum_{r\in R_o}\sum_{i\in I_{r,o}}x^u_{o,ri}=0, \text{for any } o\in P; \tag{6.18}$$

$$\sum_{r\in R_p}\sum_{i\in I_{r,p}}\sum_{o\in P}(x^l_{o,ri}-x^u_{o,ri})=x_p, \text{for any} p\in P; \tag{6.19}$$

$$x^l_p=\sum_{r\in R_p}\sum_{i\in I_{r,p}}\sum_{o\in P}x^l_{o,ri}, \text{for any} p\in P; \tag{6.20}$$

$$x_p^u = \sum_{r \in R_p} \sum_{i \in I_{r,p}} \sum_{o \in P} x_{o,ri}^u, \text{for any } p \in P; \tag{6.21}$$

$$x_p^t = \sum_{r \in R_p} \sum_{i \in I_{r,p}} \sum_{o \in P \backslash p} x_{o,ri}^l, \text{for any } p \in P; \tag{6.22}$$

$$x_{o,ri}^f = x_{o,ri-1}^f - x_{o,ri}^u + x_{o,ri}^l, \text{for any } o \in P, r \in R, i \in I_r; \tag{6.23}$$

$$\sum_{o \in P} (y_{o,ri}^f + x_{o,ri}^f) \leq Cap_r, \text{for any } r \in R, i \in I_r; \tag{6.24}$$

$$y_{o,ri}^l \geq 0, y_{o,ri}^u \geq 0, y_{o,ri}^f \geq 0, y_{od} \geq 0, y_p^l \geq 0, y_p^u \geq 0, y_p^t \geq 0, \tag{6.25}$$

$$x_{o,ri}^l \geq 0, x_{o,ri}^u \geq 0, x_{o,ri}^f \geq 0, x_p^l \geq 0, x_p^u \geq 0, x_p^t \geq 0 \tag{6.26}$$

Constraint (6.9) represents that the fulfilled demands are no more than customer demands. Equation (6.10) indicates that laden containers will not be unloaded at their original ports. Equation (6.11) represents the total fulfilled demands from a port. It must be equal to the number of the laden container loaded from this port and originating from this port. Equation (6.12) represents the fulfilled demands from one port to another. They must be unloaded at the destination port. Equations (6.13)–(6.15) represented the total laden containers that are loaded, unloaded, and transshipped at port p. Equation (6.16) represents the flow balancing for laden containers. Equation (6.17) represents the requirements of repositioning empty container out/into port p. Equation (6.18) represents that the empty containers originating from a port will not be unloaded at this port. Equation (6.19) represents that the requirements of empty containers to be repositioned out/into a port have to be satisfied. Equations (6.20)–(6.22) represent the total empty containers that are loaded, unloaded, and transshipped at port p. Equation (6.23) represents the flow balancing for empty containers. Constraint (6.24) represents the vessel capacity constraints at each leg for all routes. Constraints (6.25)–(6.26) represent the non-negative of the relevant decision variables.

The above model has advantages: (i) the flows on arcs (links) are explicitly associated with service routes; (ii) the lifting-on and lifting-off activities are associated with port-of-calls within a service route; (iii) the transshipment activities and costs can be reasonably modelled; (iv) the model with realistic sizes of the problems is computationally affordable, because the model is static and does not involve the time dimension. The disadvantages include: (i) the model assumes constant weekly demands for individual port-pairs; therefore seasonality requires additional treatment; (ii) although transshipment lifting-on/off costs are included, the demurrage costs of transshipment cargos and transshipment empty containers are accurately modelled because of the lack of operational information. Nevertheless, since shipping services are often weekly services, it is reasonable that the transshipment dwell times may be estimated to be in the range between one day and seven days; (iii) some constraints such as vessel capacity may be satisfied at the tactical planning level, but not satisfied at the operational planning level due to the dynamic operations.

6.5.3 Two-Stage Path-Based Network Flow Model for Laden and Empty Container Management

This section introduces a formulation that combines the path-based network flow model with heuristic-rules based implementation algorithm. The purpose is to model the laden and empty container management at a detailed operational level, but still applicable to large-scale planning problems in terms of computational complexity. The model presented in this section is mainly based on Song and Dong (2012).

The model consists of two stages. Stage one formulates a path-based network flow model, which is a static lower-dimension integer programming model. Stage two is to implement and adjust the solution from stage one in the dynamic system using a set of dynamic decision-making rules.

We make the following assumptions in this section: (i) after laden containers are unloaded from vessels at their destination ports, they become empty and can be reused or repositioned to other ports. The inland transportation is not considered explicitly; (ii) the shipping network consists of multiple service routes, and any two ports in the shipping network can be connected by at most three service routes. The laden container routing from the original port to the destination port is limited with at most three service routes; (iii) all service routes are of weekly frequency; (iv) the weekly demands for any O-D pair are stable (it allows variations on daily basis, and even stochastic).

We introduce the following notations for the model in this sub-section:

P	the set of ports in the system.
R	the set of shipping service routes in the system.
R	a shipping service route (consisting of a sequence of ports) that belongs to R. For simplicity, it also represents the set of ports in this service.
Cap_r	the carrying capacity in TEUs of the service route r.
D_{ij}	the average cumulative customer demands from port i to port j within a week.
C_i^f	unit lifting-off cost (per laden or empty container) at port i.
C_i^o	unit lifting-on cost (per laden or empty container) at port i.
C_{ij}^b	unit backlog cost of a customer demand from port i to port j per unit per period (day).
C_i^d	unit demurrage (or waiting) cost of a transhipment (laden) container at port i per unit per period (day).

At stage one, the laden container routing and empty repositioning problem is treated as assigning the weekly demands of laden container movements and the derived requirements of empty container movements over the given shipping network subject to vessel capacity constraints and flow balancing (i.e. total containers flow out of a port should be equal to the total containers flow into the port). The idea is similar to the model in the previous section. However, here we adopt the path-based network flow model, which includes more information about the container flows in relation to service routes so that the solutions can be relatively easily implemented at the second stage (in the dynamic operational environments). We introduce a few definitions below to explain relevant concepts in our context.

Definition 1 A port pair (p_i, p_j) is called a leg in service route r from p_i to p_j if $p_i \in r$, $p_j \in r$, and p_j is the next port-of-call immediately after p_i on the route r. If p_j is not the next port-of-call immediately after p_i on the route r, then (p_i, p_j) should be understood as a set of legs connecting port p_i to port p_j in the service route r.

Definition 2 For $r \in R$, $p_i, p_j \in r$, the port sequence $p_i, p_0, p_1, \ldots, p_n, p_j$, denoted as (p_i, r, p_j), is defined as the shortest path on the route r from p_i to p_j if the following conditions are met: (i) $p_0, p_1, \ldots, p_n \in r$; (ii) (p_i, p_0), (p_l, p_{l+1}) and (p_n, p_j) are legs in the service route r for $l = 0, 1, \ldots, n\text{-}1$; (iii) $p_i \notin \{p_0, p_1, \ldots, p_n\}$ and $p_j \notin \{p_0, p_1, \ldots, p_n\}$.

Definition 3 (i) $(p_i, r_1, p_l, r_2, p_j)$ is defined as the shortest path from port p_i to p_j with a single transhipment port at p_l using two services r_1 and r_2 if (p_i, r_1, p_l) is the shortest path on the route r_1 from p_i to p_l, and (p_l, r_2, p_j) is the shortest path on the route r_2 from p_l to p_j. (ii) Similarly, $(p_i, r_1, p_l, r_2, p_m, r_3, p_j)$ is defined as the shortest path from p_i to p_j with two transhipment ports at p_l and p_m using three services r_1, r_2, and r_3 if (p_i, r_1, p_l) is the shortest path on the route r_1 from p_i to p_l, (p_l, r_2, p_m) is the shortest path on the route r_2 from p_l to p_m, and (p_m, r_3, p_j) is the shortest path on the route r_3 from p_m to p_j.

From the assumptions at the beginning of this section, we only consider three types of paths for any given O-D port-pair (i, j) of customer demands, i.e. direct service path from original port to destination port (i, r, j), two different services path with a single transhipment (i, r_1, l, r_2, j), three different services path with two transhipments $(i, r_1, l, r_2, m, r_3, j)$.

To simplify the narrative, we introduce three sets of paths. Let Q_0 denote the set of all paths with direct shipment in the shipping network, i.e. $Q_0 := \{(i, r, j) \mid r \in R, i, j \in r\}$; Q_1 denote the set of all paths with a single transhipment; Q_2 denote the set of all paths with two transhipments; and $Q := Q_0 \cup Q_1 \cup Q_2$ representing the set of all paths for any port-pair in the shipping network (with no more than two transhipments).

The above sets of paths can be generated in a number of ways, e.g. from shipping company's experience and preference, or from a more systematic way. Song and Dong (2012) provided a path generation algorithm that enumerates all feasible paths for each set.

To formulate the path-based network flow model, we introduce the following notations to facilitate the narrative.

$O(n)$	the original port of the path $n \in Q$;
$D(n)$	the destination port of the path $n \in Q$;
$C(n)$	the transportation cost per container using the path $n \in Q$;
$y(n)$	the flow volume of laden containers using the path $n \in Q$;

$x(n)$	the flow volume of empty containers using the path $n \in Q$;
$r(n)$	the service route in the path $n \in Q_0$;
$T(n)$	the transshipment port in the path $n \in Q_1$;
$r_1(n)$	the first service route in the path $n \in Q_1$;
$r_2(n)$	the second service route in the path $n \in Q_1$;
$W(n)$	the waiting time at the transshipment port in the path $n \in Q_1$;
$T_1(n)$	the first transshipment port in the path $n \in Q_2$;
$T_2(n)$	the second transshipment port in the path $n \in Q_2$;
$r_3(n)$	the third service route in the path $n \in Q_2$;
$W_1(n)$	the waiting time at the first transshipment port in the path $n \in Q_2$;
$W_2(n)$	the waiting time at the second transshipment port in the path $n \in Q_2$;

Using the notation in Definitions 2 3, we can observe that a direct shipment path $n \in Q_0$ is characterized by $(O(n), r(n), D(n))$. A single-transhipment path $n \in Q_1$ is characterized by $(O(n), r_1(n), T(n), r_2(n), D(n))$. A twice-transhipment path $n \in Q_2$ is characterized by $(O(n), r_1(n), T_1(n), r_2(n), T_2(n), r_3(n), D(n))$.

The first stage path-based network flow model is to seek the optimal assignment of laden and empty containers onto the paths in Q. Namely, we want to find the optimal assignment $\{y(n), x(n), n \in Q\}$ by minimizing the following total cost:

$$\text{Min } J = J_o + J_f + J_t + J_b + J_d \tag{6.27}$$

Where the cost elements include: container lifting-on costs J_o, container lifting-off costs J_f, container transportation cost J_t, customer demand backlog costs J_b, and transhipment demurrage cost J_d. Here the demand backlog cost can be interpreted as the lost-sale cost in the previous two sections. However, in multi-period planning problem, backlogged demands could be satisfied at later periods, whereas lost-sales will be lost permanently. The above cost elements are defined as,

$$\begin{aligned} J_o = {} & \sum_{n \in Q_0} (y(n) + x(n)) \cdot C^o_{O(n)} + \sum_{n \in Q_1} (y(n) + x(n)) \cdot (C^o_{O(n)} + C^o_{T(n)}) \\ & + \sum_{n \in Q_2} (y(n) + x(n)) \cdot (C^o_{O(n)} + C^o_{T_1(n)} + C^o_{T_2(n)}) \end{aligned} \tag{6.28}$$

$$\begin{aligned} J_f = {} & \sum_{n \in Q_0} (y(n) + x(n)) \cdot C^f_{D(n)} + \sum_{n \in Q_1} (y(n) + x(n)) \cdot (C^f_{D(n)} + C^f_{T(n)}) \\ & + \sum_{n \in Q_2} (y(n) + x(n)) \cdot (C^f_{D(n)} + C^f_{T_1(n)} + C^f_{T_2(n)}); \end{aligned} \tag{6.29}$$

$$J_t = \sum_{n \in Q} (y(n) + x(n)) \cdot C(n) \tag{6.30}$$

$$J_b = \sum_i \sum_j \left(D_{ij} - \sum_{n \in Q, O(n)=i, D(n)=j} y(n) \right) \cdot C^b_{ij} \cdot 7; \tag{6.31}$$

$$J_d = \sum_{n\in Q_1} y(n)\cdot W(n)\cdot C^d_{T(n)} + \sum_{n\in Q_2} y(n)\cdot (W_1(n)\cdot C^d_{T_1(n)} + W_2(n)\cdot C^d_{T_2(n)}); \tag{6.32}$$

Subject to

$$\sum_{n\in Q, O(n)=i} (y(n)+x(n)) = \sum_{n\in Q, D(n)=i} (y(n)+x(n)), \text{for any port } i; \tag{6.33}$$

$$\begin{aligned}
&\sum_{n\in Q_0, r(n)=r, (i,j)\subseteq(O(n),D(n))} (y(n)+x(n)) + \sum_{n\in Q_1, r_1(n)=r, (i,j)\subseteq(O(n),T(n))} (y(n)+x(n))\\
&\sum_{n\in Q_1, r_2(n)=r, (i,j)\subseteq(T(n),D(n))} (y(n)+x(n)) + \sum_{n\in Q_2, r_1(n)=r, (i,j)\subseteq(O(n),T_1(n))} (y(n)+x(n))\\
&\sum_{n\in Q_2, r_2(n)=r, (i,j)\subseteq(T_1(n),T_2(n))} (y(n)+x(n)) + \sum_{n\in Q_2, r_3(n)=r, (i,j)\subseteq(T_2(n),D(n))} (y(n)+x(n))\\
&\le Cap_r, \text{for any leg } (i,j) \text{ in any service route } r;
\end{aligned} \tag{6.34}$$

$$\sum_{n\in Q, O(n)=i, D(n)=j} y(n) \le d_{ij}, \text{for any port-pair from port } i \text{ to port} j. \tag{6.35}$$

Equation (6.33) represents that at each port the total number of containers (laden and empty) flowing into it is equal to the total number of containers flowing out of it. Constraint (6.34) ensures that the total number of containers (laden and empty) carried on each leg for any service route does not exceed the vessel capacity (because we assumed all vessels deployed in the same service route are of the similar size). Constraint (6.35) indicates that the satisfied laden containers are no more than the customer demands. The unmet demands in the current week are backlogged and charged for a backlog cost.

It should be pointed out that the purpose of the above static integer programming model is to find the optimal $\{y(n), x(n), n\in Q\}$, which represent the assignment plan of laden and empty containers (aggregated over a week) onto the paths in Q.

The second stage aims to determine the container flows and storage dynamically over multiple periods in the planning horizon based on the weekly plan obtained at the first stage. We introduce the following dynamic variables first:

K	the planning horizon.
k	the time period (e.g. day).
V	the set of vessels in the system.
v	a vessel that belongs to V.
r_v	the service route that vessel v is deployed.
$\xi_{ij}(k)$	the customer demands from port i to j arrived in period k.
$\xi_{ij}^{r_1}(k)$	the customer demands from port i to j arrived in period k allocated to service r_1.

$\xi_{ilj}^{r_1r_2}(k)$	the customer demands from port i to j arrived in period k allocated to services r_1 and r_2 transhipped at port l.
$\xi_{ilmj}^{r_1r_2r_3}(k)$	the customer demands from port i to j arrived in period k allocated to services r_1, r_2, and r_3 transhipped at port l and m.

The above three sets of variables, $\xi_{ij}^{r_1}(k)$, $\xi_{ilj}^{r_1r_2}(k)$, $\xi_{ilmj}^{r_1r_2r_3}(k)$, represent demand routing variables (they represents the demands generated at period k to be delivered on the specified path).

$s_i(k)$	the inventory level of empty containers at port i at the end of period k.
$d_{ij}^{r_1}(k)$	the cumulative demands from port i to j at time k to be satisfied using service r_1.
$d_{0ilj}^{r_1r_2}(k)$	the cumulative demands from port i to j at time k using service r_1 and r_2 transhipped at port l, which are waiting to be satisfied at original port i at time k.
$d_{0ilmj}^{r_1r_2r_3}(k)$	the cumulative demands from port i to j at time k to be satisfied using service r_1, r_2, and r_3 transhipped at port l and m, which are waiting at original port i to be satisfied at time k.

The above three sets of variables, $d_{ij}^{r_1}(k)$, $d_{0ilj}^{r_1r_2}(k)$, $d_{0ilmj}^{r_1r_2r_3}(k)$, represent the demand states at their original ports:

$d_{1ilj}^{r_1r_2}(k)$	the cumulative laden containers from port i to j at time k using service r_1 and r_2 transhipped at port l, which are waiting at the transhipment port l to be served by r_2 at time k.
$d_{1ilmj}^{r_1r_2r_3}(k)$	the cumulative laden containers from port i to j at time k to be satisfied using service r_1, r_2, and r_3 transhipped at port l and m. which are waiting at the first transhipment port l to be satisfied at time k.
$d_{2ilmj}^{r_1r_2r_3}(k)$	the cumulative laden containers from port i to j at time k to be satisfied using service r_1, r_2, and r_3 transhipped at port l and m, which are waiting at the second transhipment port m to be satisfied at time k.

The above three sets of variables, $d_{1ilj}^{r_1r_2}(k)$, $d_{1ilmj}^{r_1r_2r_3}(k)$, $d_{2ilmj}^{r_1r_2r_3}(k)$, represent the transhipment states at transshipment ports:

$y_{ij}^{v}(k)$	the number of laden containers from port i to j on board of vessel v at time k. In other words, $y_{ij}^{v}(k)$ represents the number of the laden containers on board of vessel v at time k, whose original port is i and destination port is j. It should be pointed out that those containers may be loaded onto vessel v at a time earlier than k.
$y_{ilj}^{vr_2}(k)$	the number of laden containers from port i to l on board of vessel v at time k to be further transported to port j using service r_2 transhipped at port l.
$y_{ilmj}^{vr_2r_3}(k)$	the number of laden containers from port i to l on board of vessel v at time k to be further transported to port j using service r_2 and r_3 transhipped at port l and m.
$y_{ilj}^{r_1v}(k)$	the number of laden containers from port l to j on board of vessel v at time k, which has been transported from port i to port l using service r_1.

$y_{ilmj}^{r_1vr_3}(k)$	the number of laden containers from port l to m on board of vessel v at time k, which has been transported from port i to port l using service r_1, and is to be further transported from port m to port j using service r_3.
$y_{ilmj}^{r_1r_2v}(k)$	the number of laden containers from port m to j on board of vessel v at time k, which has been transported from port i to port l using service r_1, and from port l to port m using service r_2.

The above six sets of variables, $y_{ij}^{v}(k)$, $y_{ilj}^{vr_2}(k)$, $y_{ilmj}^{vr_2r_3}(k)$, $y_{ilj}^{r_1v}(k)$, $y_{ilmj}^{r_1vr_3}(k)$, $y_{ilmj}^{r_1r_2v}(k)$, represent laden container shipments on vessels:

$x_{ij}^{v}(k)$	the number of empty containers from port i to j on board of vessel v at time k.

Considering the multiple periods in the planning horizon, at the second stage we need to determine: (i) the dynamic demand (cargo) routing variables $\{\xi_{ij}^{r_1}(k), \xi_{ilj}^{r_1r_2}(k), \xi_{ilmj}^{r_1r_2r_3}(k)\}$, demand variables at original ports $\{d_{ij}^{r_1}(k), d_{0ilj}^{r_1r_2}(k), d_{0ilmj}^{r_1r_2r_3}(k)\}$, transhipment-at-port variables $\{d_{1ilj}^{r_1r_2}(k), d_{1ilmj}^{r_1r_2r_3}(k), d_{2ilmj}^{r_1r_2r_3}(k)\}$, and shipment-on-vessel variables $\{y_{ij}^{v}(k), y_{ilj}^{vr_2}(k), y_{ilmj}^{vr_2r_3}(k), y_{ilj}^{r_1v}(k), y_{ilmj}^{r_1vr_3}(k), y_{ilmj}^{r_1r_2v}(k)\}$; (ii) the dynamic empty container inventory variables at ports $\{s_i(k)\}$; and (iii) the dynamic empty container-on-vessel variables $\{x_{ij}^{v}(k)\}$.

Song and Dong (2012) presented a heuristic algorithm to implement the static assignment plan in a dynamic multiple period situations. It is reasonable to assume that laden containers have the priority over empty containers. We summarize the heuristic algorithm below:

A heuristic implementation algorithm

Step 1: Initialisation. Note that laden containers have priority in the dynamic assignment, we can determine the laden container routing variables $\{\xi_{ij}^{r_1}(k), \xi_{ilj}^{r_1r_2}(k), \xi_{ilmj}^{r_1r_2r_3}(k)\}$ for each period based on the static information in $\{y(n), n \in Q\}$. Other decision variables at period $k = 0$ are initialised to be zero except the empty container inventories at ports, which represent the initial distribution of the container fleet over ports.

Step 2: Let $k = k + 1$.

Step 3: For any port i: (i) Update the demand variables at original ports $\{d_{ij}^{r_1}(k), d_{0ilj}^{r_1r_2}(k), d_{0ilmj}^{r_1r_2r_3}(k)\}$ by accumulating the newly generated demands and subtracting the recently satisfied demand; (ii) update the transhipment variables at ports $\{d_{1ilj}^{r_1r_2}(k), d_{1ilmj}^{r_1r_2r_3}(k), d_{2ilmj}^{r_1r_2r_3}(k)\}$ by accumulating the newly generated transhipments and subtracting the ones that are recently transhipped out of the port.

Step 4: For any port i: (i) for any vessel v arriving at port i at period k, the empty container inventory variables at port $\{s_i(k)\}$ is updated by adding all the laden and empty containers designated to port i from vessel v; (ii) for any vessel v departing from port i at period k, the empty container inventory variables at port $\{s_i(k)\}$ is further updated by reducing the number of empty containers that are moved away from port i via vessel v (either being used to meet customer demands or repositioned out of the port).

Step 5: For any vessel v to be departing from port i at period k, if the vessel v has spare capacity on board, then

Step 5.1: Meet customer demands and load designated transhipments on vessel v as many as possible;

If there are enough residual capacity on vessel v and enough empty containers at port i, all the relevant demands at the port i can be satisfied and loaded on vessel v; Otherwise, the assignment of empty containers to customer demands and the assignment of vessel spare space to laden containers will be performed according to priority rules (e.g. whether transhipping containers have a priority; whether larger volume customers should be satisfied first); the assignment procedure terminates when either there is no more space available on board, or all relevant demands and transhipment have been loaded on the vessel. The shipment-on-vessel variables $\{y^{v}_{ij}(k),\ y^{vr_2}_{ilj}(k),\ y^{vr_2r_3}_{ilmj}(k),\ y^{r_1v}_{ilj}(k),\ y^{r_1vr_3}_{ilmj}(k),\ y^{r_1r_2v}_{ilmj}(k)\}$ are updated, and the demand and transhipment states at port i are also updated accordingly.

Step 5.2: The dynamic pushed amount of empty containers

Let E_1 denote the dynamic planned empty containers to be repositioned out of port i by vessel v after Step 5.1 based on the optimal empty repositioning plan in stage one $\{x(n),\ n\in Q\}$. It is given by $E_1 = \min\{AE_i,\ PE^{r_v}_i,\ RC^{v}_i\}$, where AE_i represents the available empty containers at port i; $PE^{r_v}_i$ represents the optimal planned empty container flows out of port i via the service route r_v, obtained from $x(n)$; and RC^{v}_i represents the available residual capacity of vessel v. Since $x(n)$ has been determined at Stage one, E_1 may be regarded as the pushed amount of empty containers to be repositioned out of port i. Those empty containers can then be proportionally split among the relevant destination ports, denoted as $\{x'^{v}_{ij}(k)\}$.

Step 5.3: The dynamic pulled amount of empty containers

Let E_2 denote the maximum additional empty containers that are able to be repositioned out of port i by vessel v after Step 5.2, by taking into account the available amount of empty containers at port i and the residual capacity of vessel v. Let $U_j(k)$ denote the requirement for empty containers at port j at time k after considering the backlogged demands, the current inventory level, the empty containers en-route, and the laden containers en-route. The amount, $\min\{E_2,\ \sum_j U_j(k)\}$, is the additional planned empty containers to be repositioned out of port i by vessel v. Note that $U_j(k)$ represents the dynamic requirements of empty containers from other relevant ports. The amount $\min\{E_2,\ \sum_j U_j(k)\}$, can be regarded as the pulled amount of empty containers to be repositioned out of port i. This amount is then further split proportionally among all relevant destination ports using the paths obtained from Stage one, denoted as $\{x''^{v}_{ij}(k)\}$.

Step 5.4: The total empty containers to be repositioned from port i to j via vessel v at period k is given by: $x^{v}_{ij}(k) = x'^{v}_{ij}(k) + x''^{v}_{ij}(k)$for $j\in r_v$.

Step 6: If $k < K$, go to Step 2; otherwise, terminate the algorithm.

In summary, the above heuristic implementation algorithm is able to determine the dynamic cargo routing variables, empty container inventory variables at ports, demand variables at original ports, transhipment variable at ports, shipment-on-vessel variables, and empty container-on-vessel variables over the multiple periods in the planning horizon.

The advantages of the model in this section are: (i) the path-based network flow model at stage one has detailed information about the shipment movements including transhipment ports, the involved service routes and the transhipment waiting time at transhipment ports; therefore, transhipment costs (lifting costs and demurrage costs) can be more accurately modelled; (ii) because we only allow maximum twice transhipments, the sizes of the path sets are relatively limited even for realistic large-scale problems; hence, the static integer programming at stage one can be solved rather quickly. Note that the decision-making rules at stage two are executed dynamically on event driven basis, e.g. only when a vessel arrives at or departs from a port. They do not require complicated iterations or searching processes. Therefore the second stage is also computationally efficient; (iii) the heuristic-rules based method at the second stage can actually be applied to stochastic situations. The combination of the push and pull mechanisms in the heuristic implementation algorithm can reasonably handle the impact of uncertain demands and adjust the empty container repositioning dynamically.

The disadvantages of the model are: (i) if we allow more than twice transhipment in the path sets, the number of paths could increase exponentially. Nevertheless, in practice it is rare for a laden container shipment to have more than two transhipments. The main reason is that transhipment will incur additional lifting-on/off costs which are quite significant among the total transport cost. (ii) at the second stage, the dynamic operational model assumes that laden containers become empty immediately after being unloaded from vessels at their destination ports. Further research is required to incorporate the inland transportation into the model.

6.5.4 Apply Solutions From Mathematical Programming Models to Stochastic Situations

In general, mathematical programming models such as linear programming or integer programming are often limited within deterministic situations. The solutions may be regarded as arc-based matrices to represent the plan of laden and empty container flows over the shipping network. As the model is deterministic, the direct application of its solution in practice may not be easily achieved due to the discrepancies between the model and the reality, especially in a stochastic dynamic environment. Multi-scenario-based method could reduce the discrepancy between the plan and the reality, but cannot eliminate the discrepancy.

A commonly used approach to implement the deterministic solution into stochastic dynamic environment is the rolling horizon policy. Namely, arc-based planning decisions are generated from the optimization models for all the periods of the planning horizon, but only the decisions in the first period of the planning horizon are implemented. Then, in the next period, when new information becomes available, some forecasts are updated and deterministic models are solved again to produce new decisions in the next planning horizon (Long et al. 2012; Di Francesco et al. 2013). However, Di Francesco et al. (2009) stated that there is no paper quantifying what is actually lost in terms of operations efficiency and profitability by using deterministic models used in a rolling-horizon fashion.

Another approach to implement the deterministic solution into stochastic dynamic environment is based on operational rules to make the solution feasible. Dong et al. (2013) presented two types of operational rules. The first type attempts to follow the deterministic solution whenever possible, e.g. assign the flows to the arcs as specified in the solution, but may assign less amount than the solution if it is not able to (e.g. if there are no enough empty containers that are available to be repositioned out due to the uncertainties in the system, the unsatisfied part of the repositioning plan will be neglected). For instance, the plan requires repositioning 1000 empty containers out of the port according to the arc-based matrices, but the total number of empty containers on hand at the port is only 900; in this case, only 900 empty containers will be repositioned out, and the 100 unsatisfied requirements in the plan will be disregarded. The second type includes a compensation mechanism during the course of solution implementation. A shadow matrix for each arc-based container flow matrix is created to store the cumulative unsatisfied flow requirements. Whenever a vessel calls at a port, both the cumulative unsatisfied flow requirements specified in the shadow matrix and the current flow requirements specified in the arc-based matrix will be tried to meet. Under these operational rules, the arc-based matrices are used as guidance to move the laden and empty containers over the shipping network in response to the dynamically changing environment. However, again it is an open question whether the above implementation of the optimal solution from the deterministic models is near optimal in stochastic situations and how to measure the degree of the closeness.

6.5.5 Inventory Control-Based Simulation Model for Laden and Empty Container Management

This section presents an inventory-based simulation model to address the management of both laden and empty containers in stochastic and dynamic situations.

It is noted that there are a rather limited number of papers that addressed the operations of container carriers between ports using simulation. Rensburg and He (2005) stated that they found only one reference (i.e. Lai et al. 1995) to a simulation model of ocean container carrier operations in the literature. Lai et al. (1995) developed a simulation model to optimize a type of heuristic allocation policy for a shipping company to transport empty containers from the Middle East to ports in the Far East. Li et al. (2007) used simulation to compare the performance of their threshold policies in three-port and four-port shipping routes. The above two papers focused on developing empty repositioning policies and showing their effectiveness in specific shipping routes. The purpose of the simulation was not designed for policy evaluation in general shipping networks. Rensburg and He (2005) described a generic simulation model of ocean container carrier operations including transporting containers from depots to customers according to requirements and from port to port according to vessels' schedules. However, their focus was not on the performance evaluation of empty container repositioning policies and no numerical results were reported. Song et al. (2005) simulated the global container-shipping network focusing on business competition between ocean carriers, in which the empty container repositioning was modeled implicitly rather than explicitly.

In the following, we present an event-driven simulation tool that can serve as a platform to evaluate and optimize inventory control-based empty container repositioning policies taking into account the stochastic nature and dynamic operations of the container shipping industry (Dong et al. 2008).

The key components of container maritime transport system include containers, vessels, ports/terminals, shipping networks and customer demands. These components interact with each other and form a dynamic container shipping system. Although individual shipping companies may manage and operate their systems differently, the basic and essential parts are similar. We make the following assumptions: (i) shipping services are on weekly basis; (ii) container unloading occurs at the vessel arrival event epoch and container loading occurs at the vessel departure epoch; (iii) laden containers that are unloaded at the destination ports will become empty containers and available for reuse after a number of weeks (many literatures assuming immediately available).

Suppose that a fleet of container vessels travel on a shipping network according to a pre-determined schedule, and we then observe a sequence of vessel arrival events into ports and vessel departure events out of ports in chronological order. These events essentially drive the evolution of the dynamic system. Robinson (2004) pointed out that: "*each event occurs at an instant in time and marks a change of state in the system*". In our case, the containers on board of vessels (both laden and empty) will not change until an event occurs. This is because no lifting-on or lifting-off activities are performed between two consecutive events.

With respect to a vessel arrival event, when a vessel arrives at a port, both the laden and empty containers onboard that are destined to, or transshipped at, the current port are usually unloaded from the vessel. The unloaded empty containers are immediately available for reuse; while unloaded laden containers at their destination ports may become empty after a number of weeks. This time varies for different ports, which represents the aggregated inland transportation time. For a transshipment container, it will be staying at the port waiting for a vessel in another service route to continue its journey.

With respect to a vessel departure event, customer demands at the current port are accumulated from the time when the last vessel (in the same direction) departed from this port to the time one day before the current vessel's departure. The one-day in advance reflects the fact that a certain period of time is required at port to prepare for loading (Song et al. 2005). The accumulated demands will be satisfied using the empty containers in inventory. If they are not sufficient, extra empty containers may be leased from lessors to meet demands. However, due to the physical constraints of vessel capacity, some demands may not be able to be carried by the current vessel; in which case customers may turn their business to other shipping companies and therefore the demands could be lost. This reflects the high competitive business environment of the container shipping industry and the customers' emphasis on just-in-time delivery. On the other hand, shipping lines may try to persuade their customers to wait for the vessel arriving next week; in which case the customer demands could be regarded as backlogged and delayed by one or more weeks. Apart from meeting customer demands, the shipping company has to make decisions on repositioning empty containers. If there are stocks of empty containers remaining at the port after meeting customer demands (e.g. those ports in west-coast America which have many more imports than exports), the operational decisions include how many empty containers need to be repositioned out of the current port, which destination ports to go and in what proportion.

The decisions on empty container repositioning are based on parameterized rules, usually represented by threshold policies. For example, in the literature, double threshold policy, or (s, S)-type inventory control policy, has been used to determine the number of empty container flow in/out of a port/depot dynamically (e.g. Li et al. 2004; Li et al. 2007; Song and Dong 2008; Dong and Song 2009; Song and Zhang 2010; Song and Dong 2011b; Dang et al. 2013; Zhang et al. 2014). The simplest threshold policy is using a single threshold-level at each port/depot to control the inventory and flows of empty containers (e.g. Song 2005; Song 2007b; Lee et al. 2012).

Apart from the input data and output data, the simulator includes the following key modules: Simulation Manager module (it controls the simulator, which takes input information from the Input Data module and sets up a the running environment for the simulator), Inventory Control Policy module (it selects customer demand satisfaction and laden container routing rules, and selects an empty container repositioning policy from a list of inventory-based control policies), Simulation Processing module (it handles the vessel arrival and departure events and the laden and empty container loading/unloading activities), and Cost Calculation module, as shown in Fig. 6.2.

The inventory control-based simulation model offers a great flexibility in handling dynamic and stochastic situations because the specific empty container repositioning decisions are determined dynamically rather than in advance. However, two issues deserve more research. Firstly, what types of inventory control-based repositioning policies are appropriate in complicated shipping networks. Secondly, how the control parameters used in these inventory control-based policies can be determined efficiently.

6.6 Conclusions and Further Research

Empty container repositioning is an important phenomenon in the container shipping industry. It has been an on-going issue since the beginning of containerization and will remain as a key issue in the future due to the nature of the industry. The critical factors that cause empty container movements include the trade imbalance, dynamic operations, uncertainties, size and type of equipment, lack of visibility and collaboration within the transport chain, and transport companies' operational and strategic practices. Among these factors, we believe that the trade imbalance, dynamic operations and uncertainties are probably the most important factors, whereas

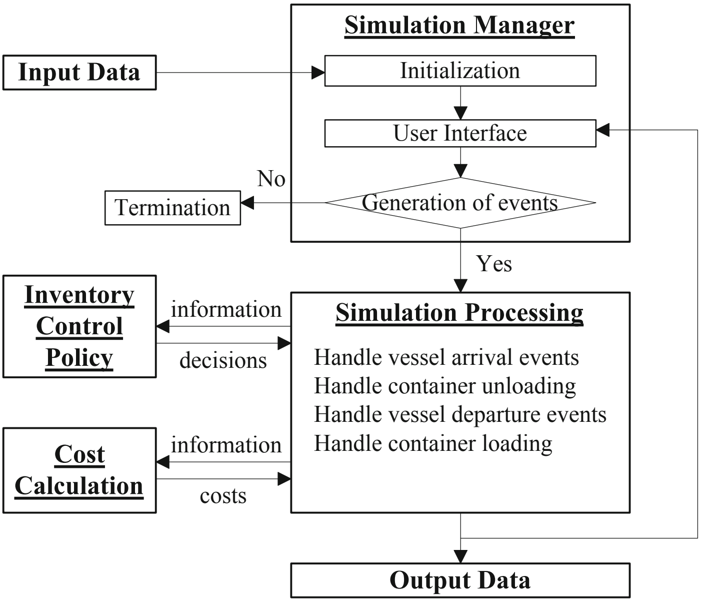


Fig. 6.2 Flowchart of an inventory control-based simulation model (based on Dong et al. 2008)

the lack of visibility and collaboration and the transport companies' practices provide opportunities for tackling the ECR problems.

By understanding the process of container storage and flows in the container transport chain, it can be seen that shipping lines are not the only ones that are affected by the ECR problem and should tackle the ECR problem, but also other players associated with the transport chain may be affected and are able to contribute to the solutions to the ECR problems. A large number of studies have been conducted to deal with the ECR problems from different angles using different methods in the last three decades. We classified the solutions to the ECR problems into four categories according to the logistics channel scope: organizational solutions, intra-channel solutions, inter-channel solutions, and technological solutions.

Due to the importance of the first three causes to the ECR (i.e. trade imbalance, dynamic operations and uncertainties), we believe that it is desirable to build models by taking into account all these three factors. In particular, trade imbalance could be more realistically modeled by considering both laden and empty container movements in a single model. We present three mathematical programming models, a time-space multi-commodity network flow model, an origin-link based network flow model, and a two-stage path-based network flow model. The third model includes a second stage to implement the static assignment plan into dynamic operation situations. We then discuss the common approached to incorporate the solutions from mathematical programming models into dynamic stochastic environments. We also present an inventory control-based simulation model, which is flexible to model the laden and empty container movements in complex dynamic stochastic environments. To some extent, the above models reflect the recent advances in the ECR modeling techniques in two broad research steams: network flow models and inventory control-based models.

As empty container repositioning problem is closely related to other issues in the container shipping, further research is required to integrate ECR with other decisions such as network design and vessel management. Apart from continuing pursuing more efficient and effective organizational solutions to the ECR problems, it is also interesting to seek appropriate intra-channel, inter-channel, and technological solutions since empty container repositioning will affect all stakeholders associated with the container transport chain.

References

Acciaro, M., & Serra, P. (2013). Maritime supply chain security: A critical review. Proceeding of the international forum on shipping, ports and airports, 3–5 June 2013, Hong Kong.

Alvarez, J. F. (2009). Joint routing and deployment of a fleet of container vessels. *Maritime Economics & Losgistics, 11*(2), 186–208.

Bandeira, D. L., Becker, J. L., & Borenstein, D. (2009). A DSS for integrated distribution of empty and full containers. *Decision Support Systems, 47,* 383–397.

Bell, M. G. H., Liu, X., Angeloudis, P., Fonzone, A., & Hosseinloo, S. H. (2011). A frequency-based maritime container assignment model. *Transportation Research Part B, 45*(8), 1152–1161.

Boile, M. (2006). *Empty intermodal container management.* Report FHWA—NJ—2006-005 for the NJ DOT. https://cait.rutgers.edu/files/FHWA-NJ-2006-005.pdf. Accessed 15 Aug 2013.

Bourbeau, B., Crainic, T. G., & Gendron, B. (2000). Branch-and-bound parallelization strategies applied to a depot location and container fleet management problem. *Parallel Computing, 26*(1), 27–46.

Braekers, K., Janssens, G. K., & Caris, A. (2011). Challenges in managing empty container movements at multiple planning levels. *Transport Reviews, 31*(6), 681–708.

Braekers, K., Caris, A., & Janssens, G. K. (2013). Optimal shipping routes and vessel size for intermodal barge transport with empty container repositioning. *Computers in Industry, 64,* 155–164.

Branch, A. (2000) *Export practice and management* (4th Ed.). London: Thomson Learning.

Brouer, B. D., Pisinger, D., & Spoorendonk, S. (2011). Liner shipping cargo alloca-tion with repositioning of empty containers. *Information Systems and Opertional Research, 49*(2), 109–124.

Chang, H., Jula, H., Chassiakos, A., & Ioannou, P. (2008). A heuristic solution for the empty container substitution problem. *Transportation Research Part E, 44*(2), 203–216.

Chao, S. L., & Lin, Y. L. (2010). Applying fuzzy AHP to select an automatic container number identification system in port terminals. 8th International conference on supply chain management and information systems, 6–9 Oct. 2010, Hong Kong, pp. 1–7.

Cheang, B., & Lim, A. (2005). A network flow based method for the distribution of empty containers. *International Journal of Computer Applications in Technology, 22*(4), 198–204.

Cheung, R. K., & Chen, C. Y. (1998). A two-stage stochastic network model and solution methods for the dynamic empty container allocation problem. *Transportation Science, 32*(2), 142–162.

Choong, S. T., Cole, M. H., & Kutanoglu, E. (2002). Empty container management for intermodal transportation networks. *Transportation Research Part E, 38*(6), 423–438.

Chou, C. C., Gou, R. H., Tsai, C. L., Tsou, M. C., Wong, C. P., & Yu, H. L. (2010). Application of a mixed fuzzy decision making and optimization programming model to the empty container allocation. *Applied Soft Computing, 10*(4), 1071–1079.

Crainic, T. G., Gendreau, M., & Dejax, P. (1993a). Dynamic and stochastic models for the allocation of empty containers. *Operations Research, 41*(1), 102–126.

Crainic, T. G., Gendreau, M., Soriano, P., & Toulouse, M. (1993b). A tabu search procedure for multicommodity location/allocation with balancing requirements. *Annals of Operations Research, 41*(4), 359–383.

Dang, Q. V., Nielsen, I. E., & Yun, W. Y. (2013). Replenishment policies for empty containers in an inland multi-depot system. *Maritime Economics and Logistics, 15*(1), 120–149.

Dejax, P., & Crainic, T. (1987). A review of empty flows and fleet management models in freight transportation. *Transportation Science, 21* (4), 227–247.

Di Francesco, M., Crainic, T. G., & Zuddas, P. (2009). The effect of multi-scenario policies on empty container repositioning. *Transportation Research Part E, 45*(5), 758–70.

Di Francesco, M., Lai, M., & Zuddas, P. (2013). Maritime repositioning of empty containers under uncertain port disruptions. *Computers & Industrial Engineering, 64*(3), 827–837.

Dong, J. X., & Song, D. P. (2009). Simulation-based optimisation for container fleet sizing and empty repositioning in liner shipping systems. *Transportation Research Part E, 45*(6), 860–877.

Dong, J. X., & Song, D. P. (2012a). Quantifying the impact of inland transport times on container fleet sizing in liner shipping services with uncertainties. *OR Spectrum, 34*(1), 155–180.

Dong, J. X., & Song, D. P. (2012b). Lease term optimization in container shipping systems. *International Journal of Logistics Research and Applications, 15*(2): 87–107.

Dong, J. X., Song, D. P., Dinwoodie, J., & Roe, M. (2008). Performance evaluation of control policies for container fleet management using simulation. LRN Annual Conference 2008: Supply Chain Innovations: People, Practice and Performance, University of Liverpool, UK, 10–12 September 2008, pp 635–640.

Dong, J. X., Xu, J. J., & Song, D. P. (2013). Assessment of empty container repositioning policies in maritime transport. *International Journal of Logistics Management, 24*(1): 49–72.

Drewry. (2006). *Annual container market review & forecast 2006/07*. London: Drewry Shipping Consultants Ltd.

Du, Y., & Hall, R. (1997). Fleet sizing and empty equipment resistribution for center-terminal transportation networks. *Management Science, 42*(2), 145–157.

Epstein, R., Neely, A., Weintraub, A., Valenzuela, F., Hurtado, S., González, G., Beiza, A., Naveas, M., Infante, F., Alarcón, F., Angulo, G., Berner, C., Catalán, J., González, C., & Yung, D. (2012). A strategic empty container logistics optimization in a major shipping company. *Interfaces, 42*(1), 5–16.

Erera, A. L., Morales, J. C., & Savelsbergh, M. (2005). Global intermodal tank container management for the chemical industry. *Transportation Research Part E, 41*(6), 551–566.

Erera, A. L., Morales, J. C., & Savelsbergh, M. W. P. (2009). Robust optimization for empty repositioning problems. *Operations Research, 57*(2), 468–483.

Feng, C. M., & Chang, C. H. (2008). Empty container reposition planning for intra-Asia liner shipping. *Maritime Policy & Management, 35*(5), 469–489.

Hanh, L. D. (2003). The logistics of empty cargo containers in the Southern California region. Final Report, Long Beach, CA.

Hong Kong Census and Statistics Department. (2012). Hong Kong monthly digest of statistics: Port transhipment cargo statistics, 2006–2011. http://www.statistics.gov.hk/pub/B71205FB2012XXXXB0100.pdf. Accessed 15 Aug 2013.

IICL. (2010). 2010 IICL annual leased container fleet survey, instititute of international container lessors. www.iicl.org/news/fleet_survey/2010_Container_Fleet_Survey.pdf. Accessed 15 Aug 2013.

Imai, A., & Rivera, F. (2001). Strategic fleet size planning for maritime refrigerated containers. *Maritime Policy and Management, 28*(4), 361–374.

Imai, A., Shintani, K., & Papadimitriou, S. (2009). Multi-port vs. Hub-and-Spoke port calls by containerships. *Transportation Research Part E, 45*(5), 740–757.

Jula, H., Chassiakos, A., & Ioannou, P. (2006). Port dynamic empty container reuse. *Transportation Research Part E, 42*(1), 43–60.

Konings, J. W. (2005). Foldable containers to reduce the costs of empty transport? A cost-benefit analysis from a Chain and multi-actor perspective. *Maritime Economics and Logistics, 7,* 223.

Konings, R., & Thijs, R. (2001). Foldable containers: a new perspective on reducing container-repositioning costs. *European Journal of Transport and Infrastructure Research, 1*(4), 333–352.

Lai, K. K., Lam, K., & Chan, W. K. (1995). Shipping container logistics and allocation. *Journal of the Operational Research Society, 46*(6), 687–697.

Lam, S. W., Lee, L. H., & Tang, L. C. (2007). An approximate dynamic programming approach for the empty container allocation problem. *Transportation Research Part C, 15*(4), 265–277.

Lee, L. H., Chew, E. P., & Luo, Y. (2011). Inventory-based empty container repositioning in a multi-port system, The first international conference on advanced communications and computation, Barcelona, Spain, October 23–29, 2011, 80–85.

Lee, L. H., Chew, E. P., & Luo, Y. (2012). Empty container management in multi-port system with inventory-based control. *International Journal on Advances in Systems and Measurements, 5*(3–4), 164–177.

Li, J. A., Liu, K., Stephen, C. H. S., & Lai, K. K. (2004). Empty container management in a port with long-run average criterion. *Mathematical and Computer Modelling, 40*(1–2), 85–100.

Li, J. A., Leung, S. C. H., Wu, Y., & Liu, K. (2007). Allocation of empty containers between multi-ports. *European Journal of Operational Research, 182*(1), 400–412.

Lirn, T. C., & Chiu, M. S. (2009). A Study of the SMART container monitoring system in the ocean shipping industry. *Proceedings of the 14th international symposium on logistics*, 5–8 July 2009, Istanbul, Turkey, 351–358.

Liu L. M., Lai, X. F., & Xu, Z. (2013). Empty container sharing and mechanism design in a two-port system. *The 26th european conference on operational research*, Rome, 1–4 July, 2013.

Long, Y., Lee, L. H., & Chew, E. P. (2012). The sample average approximation method for empty container repositioning with uncertainties, *European Journal of Operational Research, 222,* 65–75.

Lopez, E., (2003). How do ocean carriers organize the empty containers reposition activity in the USA? *Maritime Policy & Management, 30*(4), 339–355.

Meng, Q., & Wang, S. (2011a). Liner shipping service network design with empty container repositioning. *Transportation Research Part E, 47*(5), 695–708.

Meng, Q., & Wang, T. (2011b). A scenario-based dynamic programming model for multi-period liner ship fleet planning. *Transportation Research Part E, 47*(4), 401–413.

Mongelluzzo, B. (2004). Thinking inside the box. *The Journal of Commerce, 5*(2), 10–12.

Monios, J., & Wilmsmeier, G. (2013). The operational dynamics of container types in regional British port development strategies. UK: Edinburgh Napier University. http://www.europeanfreightleaders.eu/media/news/files/. Accessed 15 Aug 2013.

Moon I. K., Do Ngoc A. D., & Hur, Y. S. (2010). Positioning empty containers among multiple ports with leasing and purchasing considerations. *OR Spectrum, 32,* 765–786.

Moon, I., Do Ngoc, A. D., & Konings, R. (2013). Foldable and standard containers in empty container repositioning. *Transportation Research Part E, 49,* 107–124.

Notteboom, T. E. (2006). The time factor in liner shipping services. *Maritime Economics & Logistics, 8,* 19–39.

Olivo, A., Zuddas, P., Francesco, M. D., & Manca, A. (2005). An operational model for empty container management. *Maritime Economics and Logistics, 7*(3), 199–222.

Rensburg, J. J., & He, Y. (2005). A computer simulation model of container movement by sea. Proceedings of the 2005 winter simulation conference, Orlando, 1559–1566.

Rizzo, F., Barboni, M., Faggion, L., Azzalin, G., & Sironi, M. (2011). Improved security for commercial container transports using an innovative active RFID system. *Journal of Network and Computer Applications, 34*(2), 846–852.

Roberti, M. (2005). RFID container seals deliver security, value. *The RFID Journal*. http://www.rfidjournal.com/articles/view?1965. Accessed 15 Aug 2013.

Robinson, S. (2004). *Simulation—The practice of model development and use*. Chichester: Wiley.

Rodrigue, J. P., Comtois, C., & Slack, B. (2013). *The geography of transport systems* (3rd ed.). New York: Routledge.

Shen, W. S., & Khoong, C. M. (1995). A DSS for empty container distribution planning. *Decision Support System, 15*(1), 75–82.

Shi, N., & Xu, D. (2011). A Markov decision process model for an online empty container repositioning problem in a two-port fixed route. *International Journal of Operations Research, 8*(2), 8–17.

Shintani, K., Imai, A., Nishimura, E., & Papadimitriou, S. (2007). The container shipping network design problem with empty container repositioning. *Transportation Research Part E, 43*(1), 39–59.

Shintani, K., Konings, R., & Imai, A. (2010). The impact of foldable containers on container fleet management costs in hinterland transport. *Transportation Research Part E, 46*, 750–763.

Song, D. P. (2005). Optimal threshold control of empty vehicle redistribution in two depot service systems. *IEEE Transactions on Automatic Control, 50*(1), 87–90

Song, D. P. (2007a). Characterizing optimal empty container reposition policy in periodic-review shuttle service systems. *Journal of the Operational Research Society, 58*(1), 122–133.

Song, D. P. (2007b). Analysis of a collaborative strategy in container fleet management, the 11th world conference on transport research, University of California, Berkeley, June 24–28, 2007.

Song, D. P., & Carter, J. (2009). Empty container repositioning in shipping industry. *Maritime Policy & Management, 36*(4), 291–307.

Song, D. P., & Dong, J. X. (2008). Empty container management in cyclic shipping routes. *Maritime Economics & Logistics, 10*(4), 335–361.

Song, D. P., & Dong, J. X. (2011a). Flow balancing-based empty container repositioning in typical shipping service routes. *Maritime Economics and Logistics, 13*(1), 61–77.

Song, D. P., & Dong, J. X. (2011b). Effectiveness of an empty container repositioning policy with flexible destination ports. *Transport Policy, 18*(1), 92–101.

Song, D. P., & Dong, J. X. (2012). Cargo routing and empty container repositioning in multiple shipping service routes. *Transportation Research Part B: Methodological, 46*(10), 1556–1575.

Song, D. P., & Dong, J. X. (2013). Long-haul liner service route design with ship deployment and empty container repositioning. *Transportation Research Part B, 55,* 188–211.

Song, D. P., & Zhang, Q. (2010). A fluid flow model for empty container repositioning policy with a single port and stochastic demand. *SIAM Journal on Control and Optimization, 48*(5), 3623–3642.

Song, D. P., Zhang, J., Carter, J., Field, T., Marshall, M., Polak, J., Schumacher, K., Sinha-Ray, P., & Woods, J. (2005). On cost-efficiency of the global container shipping network. *Maritime Policy & Management, 32*(1), 15–30.

Theofanis, S., & Boile, M. (2009). Empty marine container logistics: Facts, issues and management strategies. *GeoJournal, 74*(1), 51–65.

Tioga Group. (2002). Empty ocean logistics study. Technical Report, Submitted to the Gateway cities council of 541 governments, CA, USA.

Transport Trackers. (2008). Is there a global shortage of containers? Independent Transport Research. http://www.lhbco.com/page7/files/TRANSPORT%20TRACKERS%20Apr2008.pdf. Accessed 15 Aug 2013.

United Nations. (2005) *Review of maritime transport*. Geneva: United Nations Publication.

United Nations. (2008) *Review of maritime transport*. Geneva: United Nations Publication.

United Nations. (2011) *Review of maritime transport*. Geneva: United Nations Publication.

United Nations. (2012) *Review of maritime transport*. Geneva: United Nations Publication.

United Nations. (2013) *Review of maritime transport*. Geneva: United Nations Publication.

Van Der Horst, M. R., & De Langen, P. W. (2008). Coordination in hinterland transport chains: a major challenge for the seaport community. *Maritime Economics & Logistics, 10,* 108–129

Veestra, A. W. (2005). Empty container reposition: The Port of Rotterdam case. In S. D. P. Flapper, J. A. E. E. van Nunen, & L. N. Van Wassenhoven (Eds.), *Managing closed—Loop supply chains* (pp. 65–78). Berlin: Springer.

Vernimmen, B., Dullaert, W., & Engelen, S. (2007). Schedule unreliability in liner shipping: origins and consequences for the hinterland supply chain. *Maritime Economics and Logistics, 9,* 193–213.

Vojdani, N., Lootz, F., & Rosner, R. (2013). Optimizing empty container logistics based on a collaborative network approach. *Maritime Economics and Logistics, 15,* 467–493.

Wang, S. (2013). Essential elements in tactical planning models for container liner shipping. *Transportation Research Part B, 54,* 84–99.

Wang, S. (2014). A novel hybrid-link-based container routing model. *Transportation Research Part E, 61,* 165–175.

Wang, S., & Meng, Q. (2012). Liner ship fleet deployment with container transshipment operations. *Transportation Research Part E, 48*(2), 470–484.

Wolff, J., Herz, N., & Flamig, H. (2011). Report on empty container management in the Baltic Sea region: Experiences and solutions from a multi-actor perspective, Hamburg University of Technology, The Baltic Sea region programme 2007–2013.

Wolff, J., Herz, N., & Flamig, H. (2012). Report on case study: Empty container logistics: Hamburg-Baltic Sea region, Hamburg University of Technology, The Baltic Sea region programme 2007–2013.

Yun, W. Y., Lee, Y. M., & Choi, Y. S. (2011). Optimal inventory control of empty containers in inland transportation system. *International Journal of Production Economics, 133*(1), 451–457.

Zhang, B., Ng, C. T., & Cheng, T. C. E. (2014). Multi-period empty container repositioning with stochastic demand and lost sales. *Journal of the Operational Research Society, 65*(2), 302–319.

Zhou, W. H., & Lee, C. Y. (2009) Pricing and competition in a transportation market with empty equipment repositioning. *Transportation Research Part B, 43*(6), 677–691.

Chapter 7
Robust Optimization Approach to Empty Container Repositioning in Liner Shipping

Ho-Tak Tsang and Ho-Yin Mak

Abstract In global container liner networks, the costly operations of empty container repositioning are necessitated by the imbalance of cargo flows across regions. Up to 40 and 60 % of containers shipped from Europe and North America to Asia are empty, respectively. Repositioning costs are sizable, often amounting up to 5–6 % of a shipping lines revenue. Therefore, identifying an optimal repositioning schedule to rebalance empty containers with minimal cost is one of the most critical planning problems in liner shipping. This is often complicated by the stochastic nature of demand and long transportation lead times. In this paper, we formulate a multiple-stage stochastic programming problem for the optimal repositioning of containers for a liner shipping network. As the problem is highly complex, the stochastic programming formulation is not computationally tractable. Therefore, we utilize emerging techniques in robust optimization to provide a tight approximation (bond) on the stochastic version of the problem. The resulting formulation is a second-order cone program (SOCP) and is computationally tractable. With this approximation, we perform computational experiments to evaluate the effectiveness of different repositioning policies.

7.1 Introduction

With rapid development over several decades, ocean transportation now accounts for 90 % of the international trade volume (International Maritime Organization 2012). Ocean container transportation, in particular, has become a dominant mode for freight and the backbone of global supply chain operations. The importance of container shipping can be much attributed to the standardized, re-usable containers that can be handled conveniently with standard port equipment and vessels. However, the

---

H.-T. Tsang (✉) · H.-Y. Mak
Department of Industrial Engineering and Logistics Management, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China
e-mail: ptktsang@connect.ust.hk

H.-Y. Mak
e-mail: hymak@ust.hk

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_7

repeated-use nature has also necessitated the shipment of empty containers along routes with trade imbalance. For example, it is estimated that up to 40 and 60 % of containers shipped from Europe and North America to Asia are empty, respectively (New York Times (2006). In Chapter 6, Song and Dong (2015, this volume) discuss the various causes of empty container repositioning and provide statistics for the three major shipping routes (Europe-Asia, Trans-Pacific and Trans-Atlantic).

Facing random demand for shipping from shippers, ensuring availability of containers at ports is of strategic importance to shipping carriers. As is the case for inventory control problems for other commodity types, the carrier faces the trade-off between purchasing and holding too much inventory of containers at a port, which incurs storage and maintenance costs, and having insufficient containers, which typically leads to increased costs due to engaging in short-term leases from container rental companies. Two major complicating factors involved in the empty container inventory control problem are the multi-port network structure and shipping lead times. These, together with the dynamic and stochastic nature of shipping demand necessitate the joint modeling and optimization of the holding and shipping of containers over time and space. Such modeling poses significant analytical and computational challenges.

As to be reviewed in Sect. 7.2, one of the existing mathematical programming approaches to tackling empty container management model the problem using deterministic linear programs defined on time-space networks. In such networks, nodes represent ports at different time points, and flows on (directed) edges represent shipping of empty or laden containers between ports and carrying of empty containers at the same port over time. While such a modeling approach gives rise to computationally tractable linear programs, it does not capture uncertainty in planning. A natural extension of the time-space modeling approach is to use multi-stage stochastic programming with recourse. In particular, demand for shipping laden containers can be modeled as random variables, and shipping and repositioning decisions are modeled as wait-and-see recourse decisions that are dynamically made over time upon observing realizations of demand. The major challenge lies in the large number of scenarios needed to characterize the dynamic nature of information revelation. As the length of the planning horizon increases, the size of the decision tree explodes exponentially. What further add to the complexity are the non-anticipatory constraints, i.e., the modeling requirement that different scenarios that have the same history up to the current period must induce the same decisions. In the literature, these non-anticipatory constraints are often relaxed due to computational difficulty.

This chapter aims to tackle the empty container inventory control problem using a different approach based on techniques from distributionally-robust optimization. The advantages of our approach are three-fold. First, our approach allows us to develop a computationally-tractable convex optimization model in the form of a second-order cone program, which can be efficiently solved by commercial solvers. Unlike scenario-based stochastic programming formulations, our convex optimization formulations grow in size polynomially as the length of the planning horizon increases. Second, this approach requires only partial information, such as means, variances and supports, on the distribution of random parameters (shipping demand in particular). This property makes our optimization formulation distributionally robust, i.e., it works under alternative probability distributions with the same descriptive statistics. This property makes our approach particularly attractive for the volatile ocean container shipping business, in which demand fluctuates sharply with the business cycle, trade patterns and seasonality factors. In view of such planning uncertainty, past demand data may not be perfectly reliable for generating future forecasts. In the literature, it is known that given limited (reliable) data, fitting the distribution of uncertain parameters can be difficult (e.g., (Levi et al. 2011)); whereas obtaining descriptive statistics requires less of a burden on data availability. Third, from our computational experiments, our approach performs favorably against other heuristic methods in terms of solution quality.

This chapter will be organized as follows. Sections 7.2 and 7.3 will review the related literature and the distributionally-robust optimization methodology to be applied, respectively. Then, in Sect. 7.4, we propose an empty container repositioning model for a two-port system and its mathematical formulation and illustrate the setup using a deterministic model in Sect. 7.4.1. In Sect. 7.4.2, we directly formulate the problem as a multi-stage stochastic program, which is computationally intractable. Then, in Sect. 7.4.3, we discuss how to approximate the stochastic program with a distributionally-robust conic optimization model. Finally, we present results of computational experiments in Sect. 7.5 and conclude the chapter in Sect. 7.6.

7.2 Literature Review

The literature on tactical and operational planning of shipping carriers is quite extensive. Meng et al. (2013) provides an excellent review of the development over the last 30 years and discusses promising future research directions on this topic. Operations of empty container repositioning has long been studied by a number of researchers. In Chapter 6, Song and Dong (2015, this volume) discuss in detail the topic of empty container repositioning, including its causes and various solutions, in the context of the broader container transport system. Reviews on the earlier and more recent works are provided by (Dejax and Crainic 1987) and (Cimino et al. 2010; Liu et al. 2010), respectively. In these works, the typical focus is either to analyze structural policies by considering stylized models, or to propose mathematical programming models for decision support. Quite a number of works study the problems of joint empty container repositioning and the other practical industry problems, for example, ship scheduling, network service design and container fleet sizing (see e.g., (Agarwal and Ergun 2008), (Dong and Song 2009) and (Song and Dong 2012)).

Focusing on a simplified setting with two depots or ports, Song (2005) and Song and Earl (2008) derive the optimal threshold-type control policy for empty container rebalancing. Due to the complexity of the problem, they impose certain simplifying assumptions that give rise to tractable analytical models, in the form of Markov decision processes. They relax some of these assumptions in several sequel papers. In particular, they consider the possibility of leasing containers in Song et al. (2007), hub-and-spoke operations in (Song and Carter 2009), cyclic shipping routes in (Song and Dong 2008) and a liner shipping system in (Dong and Song 2009). In the more general shipping networks, they focus on deriving optimal policies for special cases (e.g., two-port cases), and adapting those policies in the more general settings. The policies are evaluated by simulation.

Many of the earlier mathematical programming models (see, e.g., (Dejax and Crainic 1987) for a review) for empty container and vehicle repositioning are deterministic and do not account for demand and other operational uncertainties. To address planning uncertainties, a number of stochastic programming models have been proposed and studied. Crainic et al. (1993) propose a deterministic formulation for the empty container repositioning problem and extend it to a stochastic programming model. However, due to the complexity of multi-period dynamic control, they propose to solve the problem dynamically as a series of two-stage stochastic programs with restricted recourse in a rolling horizon basis. Later, Cheung and Chen (1998) propose a two-stage stochastic network formulation for the empty repositioning and leasing problem, and propose enhanced stochastic quasi-gradient and hybrid approximation algorithms for the problem, they treat supply and demand of empty container as well as residual (space availability) capacity for empty containers on containership as random variables. Francesco et al. (2009, 2013) study a scenario-based formulation for the problem.

As the number of scenarios typically increases exponentially in the number of planning periods for multi-stage stochastic programs. The stochastic programming approach is not ideal for optimizing dynamic decisions over many decision points over time at which uncertainty (e.g., demand information) are realized. To address this, one possible approach is approximate dynamic programming (e.g., (Lam et al. 2007)). However, it is known in the literature that the difficulty and effectiveness of approximate dynamic programming implementations depends on the form of approximation used (e.g., (Ben-Tal et al. 2005)). A viable alternative is adjustable robust optimization (Ben-Tal et al. 2004; Chen and Zhang 2009), which allows decisions to be adjusted dynamically as uncertainty is realized over time. Using this approach, Erera et al. (2009) propose a robust formulation for the dynamic container rebalancing problem. Starting from a nominal repositioning plan, the allocation of containers can be dynamically readjusted (recovered) for any realization of demand within an uncertainty set. They further prove that the problem is polynomially solvable. However, in their model, the objective is simply the cost incurred by the nominal solution, i.e., the recovery actions do not directly impact the objective as long as feasibility is guaranteed. Furthermore, the definition of uncertainty sets implicitly does not take into account likelihood information of the uncertain parameters (e.g., past demand data).

Recent developments in robust optimization has allowed the adjustable robust optimization formulations to take into account distributional information of uncertain parameters, in the context of distributionally-robust optimization. This modeling approach requires only partial distributional information (e.g., mean, support, variance), but not the precise distribution (Chen et al. 2008; Goh and Sim 2010). A very desirable property is that, very often, the distributionally-robust counterpart (or a tight approximation thereof) can be efficiently computed using convex conic programs. This approach has been applied to several contexts, such as multi-period inventory control (See and Sim 2010) and warehouse management (Ang et al. 2012). Using this approach, we are able to obtain a tractable approximate formulation of the empty container inventory planning problem, which is scalable both in the time dimension and the size of the shipping network.

7.3 Methodology

Stochastic programming has long been used as a standard modeling approach for multi-stage optimization problems under uncertainty. Such optimization problems incoroporate dynamic realization of uncertain parameters over time and recourse decisions, i.e., as uncertainty is resolved incrementally over time, decisions are made in a step-by-step fashion by taking into account the available information that has been observed. In the empty container repositioning problem studied in this chapter, demand for shipping laden containers is observed at the beginning of a period, and the decisions on the number of empty containers to reposition for said period are then made subsequently. As shall be seen in Sect. 7.4, a direct multi-stage stochastic programming formulation of this problem is computationally intractable. Therefore, we utilize recent results developed in the area of distributionally-robust optimization to obtain tractable approximations for the stochastic program. The approximations are built on the notion of truncated affine policies. In this section, we review the key concepts based on a general stochastic linear program. Then in Sect. 7.4, we discuss how this approach can be adapted to the empty container repositioning problem. We use the $\tilde{}$ sign to denote random variables, boldface letters to denote matrices and vectors, and $'$ to denote the transpose of vectors.

First, we consider a multi-stage stochastic linear program of the following form:

$$\min_{\mathbf{x},\mathbf{y}(\cdot)} \quad \mathbf{a}'\mathbf{x} + E_{\mathbb{P}}\left[\mathbf{d}'\mathbf{y}(\tilde{z})\right] + E_{\mathbb{P}}\left[\mathbf{c}'\mathbf{y}(\tilde{z})\right]^{+} \tag{7.1}$$

$$\text{subject to:} \quad \mathbf{A}(\tilde{z})\mathbf{x} + \mathbf{B}\mathbf{y}(\tilde{z}) = \mathbf{b}(\tilde{z}) \tag{7.2}$$

$$\underline{\mathbf{y}} \le \mathbf{y}(\tilde{z}) \le \bar{\mathbf{y}}. \tag{7.3}$$

In the above, $\tilde{z} = (\tilde{z}_1, \cdots, \tilde{z}_M)$ is a random vector defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The vector $\mathbf{x}$ denotes the first-stage decision variables that are determined before the random variables are realized. The vector $\mathbf{y}(\tilde{z})$ denotes the recourse decision variables that are determined after the random variables $\tilde{z}$ are (partially) realized. Because of the wait-and-see nature of these decision variables, they can be written as functions of $\tilde{z}$, whose forms are to be optimized. Note that non-anticipatory relationships can be easily incorporated using this representation, by forcing the functional form of $\mathbf{y}(\cdot)$ to be independent of components of $\tilde{z}$ that cannot be observed prior to making the corresponding decisions. For example, the repositioning quantity at a certain period can only depend on shipping demand observed in the same period or before, but not the future periods. In the objective function 7.1, the expectation of $\left[\mathbf{c}'\mathbf{y}(\tilde{z})\right]^+$, where $[\,\cdot\,]^+ = \max\{\cdot, 0\}$, can be used to represent, for example, shortages and surplus inventory of empty containers. We assume that, while $\tilde{z}$ is uncertain, the model data $(\mathbf{a}, \mathbf{c}, \mathbf{d}, \mathbf{A}, \mathbf{B}, \mathbf{b}, \underline{\mathbf{y}}, \bar{\mathbf{y}})$ are known and deterministic. In particular, the assumption that $\mathbf{B}$ is deterministic is referred to as the fixed recourse condition in stochastic programming.

With the functional forms of $\mathbf{y}(\,\cdot\,)$ to be optimized, the problem is computationally intractable. Dyer and Stougie (Dyer and Stougie 2006) proved that even a two-stage version of the problem (i.e., all random variables are realized at once) is $\#P$-hard to solve, given some particular distribution of $\mathbb{P}$. Ben-Tai et al. (Ben-Tal et al. 2004) consider a version of the problem in which $\mathbb{P}$ is only known to fall within the family of all distributions with given support, and the worst-case expectation is considered in the objective. They show that the problem can be NP-hard. Therefore, they propose the idea of restricting the space of functional forms for $\mathbf{y}(\cdot)$ to a set of tractably parameterized functions, known as *decision rules*, and optimizing the function parameters instead. By choosing appropriate classes (affine functions in particular) of decision rules, the (restricted) problem becomes computationally tractable. Using this idea, Ben-Tal et al. (2005) applied affine decision rules to a multi-period inventory control problem.

Chen et al. (2007, 2008) consider the more general case in which the distribution $\mathbf{P}$ is not known, and falls within a family with given descriptive statistics such as support, mean, (co)variance and new deviation measures known as forward and backward deviations. Then, they show how the problem 7.1, possibly with additional chance constraints, can be approximated using affine and other classes of piecewise linear decision rules using tractable second-order cone programs. See and Sim (2010) further introduce a class of truncated affine decision rules to improve the performance for cases where recourse variables have to lie within certain bounds. We shall briefly review some of these key results below.

First, we review the notion of forward and backward deviations (Chen et al. 2007).

Definition 1 For a univariate, zero-mean random variable $\tilde{z}$, the forward and backward deviations are defined as:

$$\begin{aligned}\sigma_f(\tilde{z}) &= \sup_{\theta>0}\left\{2\log\left(E(\exp\left(\theta(\tilde{z}-E(\tilde{z}))\right))\right)/\theta^2\right\},\\ \sigma_b(\tilde{z}) &= \sup_{\theta>0}\left\{2\log\left(E(\exp\left(-\theta(\tilde{z}-E(\tilde{z}))\right))\right)/\theta^2\right\}.\end{aligned}$$

While the definitions of these new deviation measures may not appear immediately intuitive, one can show that, for many common symmetric distributions, their values are typically quite close to the standard deviation. For example, for the normal distribution, their values are exactly equal to the standard deviation. Furthermore, they also possess many similar mathematical properties as the standard deviation and can be estimated conveniently. See Chen et al. (2007) for more details. The significance of these new deviation measures is that they capture asymmetry of distributions, by reflecting deviations above (forward) and below (backward) the mean.

This allows for the formulation of stronger bounds on expectations and probabilities of functions of the random variable than what can be implied by the standard deviation, in the case where the precise distribution of $\tilde{z}$ is not known. For example, the following probability bounds hold:

Lemma 1 *(Chen et al. (2007)) For any* $\kappa > 0$, *it holds that* $P(\tilde{z} > \kappa\sigma_f(\tilde{z})) \leq \exp(-\kappa^2/2)$ *and* $P(\tilde{z} < -\kappa\sigma_b(\tilde{z})) \leq \exp(-\kappa^2/2)$.

Practically, these probability bounds are typically significantly tighter than the Chebyshev-type bounds implied by only the standard deviation (and mean). These deviation measures will be useful in developing the subsequent approximations involving the decision rules.

To begin the discussion on truncated affine decision rules, we first state the following assumption (Chen and Sim (2007)):

Assumption 1 (Chen and Sim (2007)). We assume that the uncertainties $\{\widetilde{z_j}\}_{j=1:N}$ are zero mean random variables, with positive definite covariance matrix $\sum$. Let W be the smallest convex set containing the support of $\widetilde{\mathbf{z}}$. We denote a subset, J $\subseteq \{1, \cdots,$ N$\}$, which can be an empty set, such that $\widetilde{z_j}$, j $\in$ J are stochastically independent. Moreover, the corresponding forward and backward deviation $p_j = \sigma_f(\widetilde{z_j})$ and $q_j = \sigma_b(\widetilde{z_j})$, respectively, for $j \in J$ and that $p_j = q_j = \infty$ for $j \notin J$.

Assumption 2 We assume that the matrix $\mathbf{A}(\tilde{z})$ and $\mathbf{b}(\tilde{z})$ are affinely dependent on $\tilde{z}$. That is,

$$\mathbf{A}(\tilde{z}) = \mathbf{A}^0 + \sum_{i=1}^{N} \mathbf{A}^i \tilde{z}_i \,, \qquad \mathbf{b}(\tilde{z}) = \mathbf{b}^0 + \sum_{i=1}^{N} \mathbf{b}^i \tilde{z}_i .$$

Then, we consider the following functional form of the recourse variables:

$$y_k(\tilde{z}) = \min\left\{\max\left\{y_k^0 + \sum_{i=1}^{N} y_k^i \tilde{z}_i, \underline{y}_k\right\}, \bar{y}_k\right\}. \tag{7.4}$$

Note that the above decision rule takes the piecewise linear form. In particular, it is obtained by assuming that the recourse variables are simple affine functions $y_k^0 + \sum_{i=1}^{N} y_k^i \tilde{z}_i$, and truncating the values at boundary points of $\underline{\mathbf{y}}$ and $\bar{\mathbf{y}}$ to satisfy 7.3. Then, following the discussion of See and Sim (2010), the following formulation provides an upper bound on 7.1:

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y^0},\mathbf{y^i}} \ \mathbf{a}'\mathbf{x} + E_{\mathbb{P}}\left[\sum_{k=1}^{K} d_k \left(y_k^0 + \sum_{i=1}^{N} y_k^i \tilde{z}_i\right)\right] + E_{\mathbb{P}}\left[\sum_{k=1}^{K} c_k \left(y_k^0 + \sum_{i=1}^{N} y_k^i \tilde{z}_i + \right.\right. \\ \left.\left.\left(y_k^0 + \sum_{i=1}^{N} y_k^i \tilde{z}_i - \bar{y}_k\right)^+ + \left(-y_k^0 - \sum_{i=1}^{N} y_k^i \tilde{z}_i + \underline{y}_k\right)^+\right)\right]^+ \end{aligned} \tag{7.5}$$

$$\text{subject to:} \quad \mathbf{A^0 x} + \mathbf{B y}^0 = \mathbf{b}^0 \tag{7.6}$$

$$\mathbf{A^i x} + \mathbf{B}^i \mathbf{y}^i \quad = \quad \mathbf{b}^i \text{ , for } i = 1, \cdots N. \tag{7.7}$$

In the above, constraints 7.6 and 7.7 impose the equality constraint 7.2 to hold for all corresponding coefficients in the affine mapping $y_k(\tilde{z}) = y_k^0 + \sum_{i=1}^N y_k^i \tilde{z}_i$, ignoring the truncation in 7.4. To reflect this truncation, note that the expectation term in the objective function 7.5 contains two correction terms. The first corrects underestimations of the objective value arising from the (pre-truncation) affine mapping exceeding the upper bound $\bar{y}_k$, and the latter adjusts for the (pre-truncation) affine mapping reaching below $\underline{y}_k$.

With the above, the last remaining step is to obtain tractable formulations of the last expectation term in 7.5 in the form $E[\,\cdot + \sum\{\cdot\}^+]^+$. In general, evaluating and optimizing such expectations is intractable; nevertheless, the results from See and Sim (2010) allow these expectation terms to be tightly bounded above by SOCPs (whose detailed formulations are omitted for brevity - refer to (See and Sim 2010) for details):

Theorem 1 *(See and Sim (2010)) Under the Assumption 1, for any* $(u^0, \mathbf{u})$*, it holds that* $E\left[u^0 + \mathbf{u}'\tilde{\mathbf{z}}\right]^+ \leq \pi(u^0, \mathbf{u})$*, where* $\pi(u^0, \mathbf{u})$ *can be obtained as the optimal objective value of an SOCP.*

Based on Theorem 1., the authors derived an upper bound on general nested sum of expected positive parts of random variables as follows:

Theorem 2 *(See and Sim (2010)) Under the Assumption 1, for any* $(u^0, \mathbf{u}), (v_1^0, \mathbf{v_1}), \cdots, (v_M^0, \mathbf{v_M})$*, it holds that:*

$$E\left[u^0 + \mathbf{u}'\tilde{\mathbf{z}} + \sum_{m=1}^{M} (v_m^0 + \mathbf{v_m}\tilde{\mathbf{z}})^+\right]^+ \leq \eta((u^0, \mathbf{u}), (v_1^0, \mathbf{v_1}), \cdots, (v_M^0, \mathbf{v_M}))$$

where $\eta((u^0, \mathbf{u}), (v_1^0, \mathbf{v_1}), \cdots, (v_M^0, \mathbf{v_M}))$

$$= \min_{w_m^0, \mathbf{w}_m} \left\{ \pi(u^0 + \sum_{m=1}^{M} w_m^0, \mathbf{u} + \sum_{m=1}^{M} \mathbf{w}_m) + \sum_{m=1}^{M} (\pi(-w_m^0, -\mathbf{w}_m) + \pi(v_m^0 - w_m^0, \mathbf{v_m} - \mathbf{w_m})) \right\}$$

Therefore, for any stochastic linear program in the form 7.1, one can utilize Theorems 1 and 2 to obtain tight bounds on the objective value by solving tractable SOCPs. Note that these bounds hold for any distribution $\mathbb{P}$ with the descriptive statistics specified in Assumption 1, and are thus *distributionally robust*. In the next section, we apply this technique to the empty container repositioning problem.

7.4 Empty Container Inventory Planning Problem

We study a shipping route consisting of two ports, j and k. Inventory of empty containers is managed at the two ports for a finite horizon consisting of T periods, indexed by $t = 1, ..., T$. We assume that one vessel departs from each port in each period (e.g., for a weekly service), and takes L periods to finish one leg of the route. These assumptions are made solely for simplicity of illustration. It is straightforward to relax these assumptions to allow asymmetric service frequencies or lead times between the two ports. At the beginning of the planning horizon, the carrier acquires z_j^0 and z_k^0 containers for the two ports. In each period t, shippers demand to ship a random number n_{jk}^t (n_{kj}^t) of laden containers from j to k (k to j). Because the shipping volumes in the two directions are random and may not be balanced, the carrier has the option to ship empty containers on the vessel, subject to the capacity of the vessel, which is assumed to be K containers. We define the decision variables r_{jk}^t (r_{kj}^t) to denote the number of empty containers shipped from j to k (k to j) in period t. As a result of the laden and empty container flows, the empty container inventory available at the end of each period t will vary. We denote the inventory levels by z_j^t and z_k^t.

Due to lead time in hinterland operations, outbound demand can only be met with empty containers available l_e periods in advance. That is, l_e is the duration needed for the empty containers to be picked up at the port by shippers, transported to the shippers' facilities (e.g., factories or warehouses), loaded with cargo, and returned to the port. Similarly, incoming containers at a port only become available for use in subsequent shipments l_i periods after arrival. If there are insufficient containers available at the port to meet demand for shipments, the shipping line engages in a short-term lease for the shortfall with a container leasing company, and incurs a penalty cost of p per container of shortage. The notation is summarized as follows: Parameters:

- $\tilde{n}_{jk}^t, \tilde{n}_{kj}^t$ = Stochastic demand for laden container shipment from j to k and k to j in period t, respectively;
- l_i, l_e = Processing lead times for import and export containers in hinterland, respectively (assumed to be equal and constant over time for both ports for simplicity);
- L = Vessel travel lead time (one way);
- K = Capacity of vessel;
- c = Shipping cost for empty container;
- h = Holding cost for storing empty container at the port;
- p = Penalty cost for having incurring shortages of containers.

Decision Variables:

- z_j^t = Net inventory of empty containers at port j at the end of period t, which is positive if there is leftover inventory and negative if there is shortage;
- i_j^t = Leftover inventory level of empty containers at port j at the end of period t;
- s_j^t = Shortage of empty containers at port j at the end of period t;
- r_{jk}^t = Number of empty containers to be shipped from j to k in period t.

7.4.1 Deterministic Case: Linear Programming Formulation

To begin with, we would use a deterministic case to illustrate the idea of our multi-stage empty container inventory planning problem. We consider the case in which the demands for shipping laden containers are deterministic, i.e., $\tilde{n}^t_{jk} = \mu^t_{jk}$ and $\tilde{n}^t_{kj} = \mu^t_{kj}$. For notational convenience, we define $\tilde{n}^t_{\cdot} \equiv 0$ and $\mu^t_{\cdot} = 0$ for $t < 0$ throughout the paper. Note that the shipping demands can still vary deterministically over time, e.g., due to seasonality patterns. Because all problem inputs are deterministic, the repositioning decisions can be made optimally at the beginning of the planning horizon. This can be done by solving the following linear program:

$$\min \sum_{t=1}^{T} [c(r^t_{jk} + r^t_{kj}) + h(i^t_j + i^t_k) + p(s^t_j + s^t_k)] \tag{7.8}$$

subject to:

$$z^t_j = z^{t-1}_j + \mu^{t-l_i-L}_{kj} - \mu^{t+l_e}_{jk} + r^{t-L}_{kj} - r^t_{jk}, \text{for } t = 1, \cdots, T \tag{7.9}$$

$$z^t_k = z^{t-1}_k + \mu^{t-l_i-L}_{jk} - \mu^{t+l_e}_{kj} + r^{t-L}_{jk} - r^t_{kj}\ , \text{for } t = 1, \cdots, T \tag{7.10}$$

$$0 \le r^t_{jk} \le K - \mu^{t+l_e}_{jk}\ , \text{for } t = 1, \cdots, T \tag{7.11}$$

$$0 \le r^t_{kj} \le K - \mu^{t+l_e}_{kj}\ , \text{for } t = 1, \cdots, T \tag{7.12}$$

$$i^t_j \ge z^t_j, i^t_j \ge 0\ , \text{for } t = 1, \cdots, T \tag{7.13}$$

$$i^t_k \ge z^t_k, i^t_k \ge 0\ , \text{for } t = 1, \cdots, T \tag{7.14}$$

$$s^t_j \ge -z^t_j, s^t_j \ge 0\ , \text{for } t = 1, \cdots, T \tag{7.15}$$

$$s^t_k \ge -z^t_k, s^t_k \ge 0\ , \text{for } t = 1, \cdots, T. \tag{7.16}$$

In the above, the objective 7.8 is to minimize the shipping cost, inventory holding cost and shortage penalty over the planning horizon. The flow conservation constraints 7.9 and 7.10 require that the number of containers leaving and entering a port are equal. Note that, due to the hinterland operations lead times, inbound containers are only counted toward available inventory l_i periods after vessel arrival, and outbound containers depart l_e periods prior to vessel departure. Constraints 7.11 and 7.12 require that the number of laden and empty containers shipped on a vessel must be nonnegative and cannot exceed the vessel capacity. Constraints 7.13 and 7.14 relate the storage quantity variables i^t_j and i^t_k with the net inventory variables z^t_j and z^t_k. Similarly, constraints 7.15 and 7.16 relate the shortage variables s^t_j and s^t_k with z^t_j and z^t_k. Note that, with a minimization objective, these constraints guarantee that $i^t_j = [z^t_j]^+$, $i^t_k = [z^t_k]^+$, $s^t_j = [-\,z^t_j]^+$ and $s^t_k = [-\,z^t_k]^+$ at the optimal solution.

7.4.2 Multi-Stage Stochastic Programming Formulation

We now proceed with formulating the problem of optimizing the inventory levels of empty containers and the flows of empty containers under stochastic demand. With stochastic demands that are revealed dynamically over time (i.e., realizations of $\tilde{n}_{jk}^{t+l_e}$ and $\tilde{n}_{kj}^{t+l_e}$ are observable only in period t when the containers are requested by shippers), rebalancing shipments must be determined dynamically over time as recourse decisions based on the realized demand and inventory levels, unlike in the deterministic case where all repositioning flows can be detemined in one shot. Therefore, we formulate a multi-stage stochastic program for the problem:

$$\min E_\omega \left[\sum_{t=1}^{T} \{ c(r_{jk}^t(\omega) + r_{kj}^t(\omega)) + h(i_j^t(\omega) + i_k^t(\omega)) + p(s_j^t(\omega) + s_k^t(\omega)) \} \right] \tag{7.17}$$

subject to:

$$z_j^t(\omega) = z_j^{t-1}(\omega) + \tilde{n}_{kj}^{t-l_i}(\omega) - \tilde{n}_{jk}^{t+l_e}(\omega) + r_{kj}^{t-L}(\omega) - r_{jk}^t(\omega) \text{ ,for } t = 1, \cdots, T, \omega \in \Omega \tag{7.18}$$

$$z_k^t(\omega) = z_k^{t-1}(\omega) + \tilde{n}_{jk}^{t-l_i}(\omega) - \tilde{n}_{kj}^{t+l_e}(\omega) + r_{jk}^{t-L}(\omega) - r_{kj}^t(\omega) \text{ , for } t = 1, \cdots, T, \omega \in \Omega \tag{7.19}$$

$$0 \le r_{jk}^t(\omega) \le K - \tilde{n}_{jk}^{t+l_e}(\omega) \text{ , for } t = 1, \cdots, T, \omega \in \Omega \tag{7.20}$$

$$0 \le r_{kj}^t(\omega) \le K - \tilde{n}_{kj}^{t+l_e}(\omega) \text{ , for } t = 1, \cdots, T, \omega \in \Omega \tag{7.21}$$

$$i_j^t(\omega) \ge z_j^t(\omega), s_j^t(\omega) \ge 0 \text{ , for } t = 1, \cdots, T, \omega \in \Omega \tag{7.22}$$

$$i_k^t(\omega) \ge z_k^t(\omega), s_k^t(\omega) \ge 0 \text{ , for } t = 1, \cdots, T, \omega \in \Omega \tag{7.23}$$

$$s_j^t(\omega) \ge -z_j^t(\omega), s_j^t(\omega) \ge 0 \text{ , for } t = 1, \cdots, T, \omega \in \Omega \tag{7.24}$$

$$s_k^t(\omega) \ge -z_k^t(\omega), s_k^t(\omega) \ge 0 \text{ , for } t = 1, \cdots, T, \omega \in \Omega \tag{7.25}$$

$$r_{jk}^t(\omega) = r_{jk}^t(\xi), r_{kj}^t(\omega) = r_{kj}^t(\xi) \text{ , for } \xi \in \Omega^t(\omega), t = 1, \cdots, T, \omega \in \Omega. \tag{7.26}$$

In the above formulation, we use $\omega \in \Omega$ to denote the possible realizations, or sample paths, of demands $(\tilde{n}_{jk}^1, \cdots, \tilde{n}_{jk}^T, \tilde{n}_{kj}^1, \cdots, \tilde{n}_{kj}^T)$. To capture recourse decisions, the decision variables now depend on which ω is realized. The objective function 7.17 is expected objective value over all possible sample paths, and constraints 7.18 to 7.25 are required to hold for all sample paths.

Note that, in each $t < T - l_e$, the future demands $(\tilde{n}_{\cdot}^{t+l_e+1}, \cdots \tilde{n}_{\cdot}^{T})$ are not yet observable. That is, two sample paths ω_1 and ω_2 with the same demands from period 1 to period $t + l_e$ (but different in some of the subsequent period(s)) are indistinguishable to the decision maker. We use $\Omega^t(\omega) = \left\{ \xi \in \Omega : \tilde{n}_{\cdot}^1(\xi) = \tilde{n}_{\cdot}^1(\omega), \cdots, \tilde{n}_{\cdot}^{t+l_e}(\xi) = \tilde{n}_{\cdot}^{t+l_e}(\omega) \right\}$ to denote all scenarios with the same observable history prior to making period-t decisions. In the aforementioned example, $\omega_2 \in \Omega^\tau(\omega_1)$ and $\omega_1 \in \Omega^\tau(\omega_2)$ for all $1 \leq \tau \leq t$. Because recourse decisions can only be made based on observed information, decisions under sample paths ω_1 and ω_2, up to period t, must be the same. This requirement is formulated in 7.24, commonly referred to as the non-anticipatory constraints in the literature.

In practice, the formulation 7.17 is very challenging to solve. If the demand in each period follows a discrete distribution, then the number of possible realizations (the size of Ω) increases exponentially in T. Alternatively, if demand follows a continuous distribution, the set Ω has infinite number of elements in general and a practical approach is to consider only a finite sample. However, to accurately capture the problem dynamics, the sample size needed also increases rapidly in T. Therefore, the problem sizes for practical instances are very large, because decision variables and constraints need to be defined for each sample path.

In view of such computational challenges, multi-stage stochastic programming formulations such as the above are typically not solved directly in practice. Instead, one popular approach (e.g., (Crainic et al. 1993; Cheung and Chen 1998)) is to relax non-anticipatory constraints 7.24, which allows the problem to be transformed into a two-stage stochastic program. In this particular case, such an approach involves defining a number of sub-sample paths spanning periods $t = 2, \cdots, T$, each of which consists of one particular realization of $(\tilde{n}^2_{jk}, \cdots, \tilde{n}^T_{jk}, \tilde{n}^2_{kj}, \cdots, \tilde{n}^T_{kj})$. Then, because the realization of demands in all subsequent periods are fixed, it is possible to optimize repositioning flows along the entire sample path in one shot, with a formulation similar to 7.8. These sub-formulations are then linked with the decisions of the current period, $t = 1$, with flow balance constraints 7.18 and 7.19.

Our proposed approach to solve the problem, on the other hand, utilizes decision rules to update decisions from one period to the next, instead of re-solving the remaining sub-problem. To illustrate the concept of decision rules, we provide an equivalent representation of the multi-stage stochastic program. Because decision variables in the stochastic program are recourse decisions made after observing the history of past demand, we can think of the decision variables as functions of past demand realizations, in the form of $z^t_j(n^1_{jk}, ..., n^{t+l_e}_{jk}, n^1_{kj}, ..., n^{t+l_e}_{kj})$ and $r^t_{jk}(n^1_{jk}, ..., n^{t+l_e}_{jk}, n^1_{kj}, ..., n^{t+l_e}_{kj})$. These functions are referred to as decision rules, because they can be interpreted as rules that specify the recourse decisions given past demand realizations. With this representation, solving the stochastic program is equivalent to optimizing these functions over the space of all functions satisfying all constraints of the problem. In summary, we may rewrite the formulation as follows:

$$\begin{aligned}\min \sum_{t=1}^{T} & c(E[r^t_{jk}(n^1_{jk}, ..., n^{t+l_e}_{jk}, n^1_{kj}, ..., n^{t+l_e}_{kj})] + E[r^t_{kj}(n^1_{jk}, ..., n^{t+l_e}_{jk}, n^1_{kj}, ..., n^{t+l_e}_{kj})]) \\ & +h(E[z^t_j(n^1_{jk}, ..., n^{t+l_e}_{jk}, n^1_{kj}, ..., n^{t+l_e}_{kj})]^+ + E[z^t_k(n^1_{jk}, ..., n^{t+l_e}_{jk}, n^1_{kj}, ..., n^{t+l_e}_{kj})]^+) \\ & +p(E[z^t_j(n^1_{jk}, ..., n^{t+l_e}_{jk}, n^1_{kj}, ..., n^{t+l_e}_{kj})]^- + E[z^t_k(n^1_{jk}, ..., n^{t+l_e}_{jk}, n^1_{kj}, ..., n^{t+l_e}_{kj})]^-)\end{aligned} \tag{7.27}$$

subject to:

$$
\begin{aligned}
z_j^t(n_{jk}^1,\ldots,n_{jk}^{t+l_e},n_{kj}^1,\ldots,n_{kj}^{t+l_e}) &= z_j^{t-1}(n_{jk}^1,\ldots,n_{jk}^{t+l_e-1},n_{kj}^1,\ldots,n_{kj}^{t+l_e-1}) + n_{kj}^{t-l_i} - n_{jk}^{t+l_e} \\
&\quad + r_{kj}^{t-L}(n_{jk}^1,\ldots,n_{jk}^{t+l_e-L},n_{kj}^1,\ldots,n_{kj}^{t+l_e-L}) \\
&\quad - r_{jk}^t(n_{jk}^1,\ldots,n_{jk}^{t+l_e},n_{kj}^1,\ldots,n_{kj}^{t+l_e}) \qquad (7.28)
\end{aligned}
$$

$$
\begin{aligned}
z_k^t(n_{jk}^1,\ldots,n_{jk}^{t+l_e},n_{kj}^1,\ldots,n_{kj}^{t+l_e}) &= z_k^{t-1}(n_{jk}^1,\ldots,n_{jk}^{t+l_e-1},n_{kj}^1,\ldots,n_{kj}^{t+l_e-1}) + n_{jk}^{t-l_i} - n_{kj}^{t+l_e} \\
&\quad + r_{jk}^{t-L}(n_{jk}^1,\ldots,n_{jk}^{t+l_e-L},n_{kj}^1,\ldots,n_{kj}^{t+l_e-L}) \\
&\quad - r_{kj}^t(n_{jk}^1,\ldots,n_{jk}^{t+l_e},n_{kj}^1,\ldots,n_{kj}^{t+l_e}) \qquad (7.29)
\end{aligned}
$$

$$0 \le r_{jk}^{t-L}(n_{jk}^1,\ldots,n_{jk}^{t+l_e-L},n_{kj}^1,\ldots,n_{kj}^{t+l_e-L}) \le K - n_{jk}^t \qquad (7.30)$$

$$0 \le r_{kj}^{t-L}(n_{jk}^1,\ldots,n_{jk}^{t+l_e-L},n_{kj}^1,\ldots,n_{kj}^{t+l_e-L}) \le K - n_{kj}^t. \qquad (7.31)$$

The above problem is clearly intractable (Dyer and Stougie 2006), as it attempts to optimize over the space of all functions of the associated random variables. Therefore, we adopt approximations to simplify this problem by restricting the class of functions to the family of truncated affine decision rules discussed in Sect. 7.3.

7.4.3 Distributionally-Robust Optimization Approximation

To obtain high-quality solutions to 7.27, we adopt a pragmatic approach and perform the optimization over a class of restricted functions. In particular, it is natural to set the net rebalancing shipments (the r variables) as affine functions of the realized demand, or:

$$r_{jk}^t(n_{jk}^1,\ldots,n_{jk}^{t+l_e},n_{kj}^1,\ldots,n_{kj}^{t+l_e}) = \bar{r}_{jk}^t + \sum_{\tau=1}^{t+l_e}\left(\hat{r}_{j,jk}^{\tau,t}n_{jk}^{\tau} + \hat{r}_{k,kj}^{\tau,t}n_{kj}^{\tau}\right).$$

One desirable property of the affine functional form is that it is consistent with simple policies, such as setting rebalancing shipments to the expected values, realized values or moving averages of the imbalance levels. In Sect. 7.5, we compare solutions obtained based on our approach with these simple heuristic policies.

Instead of optimizing the functions r_{jk}^t in the general form, we optimize within this class of affine functions by treating $\bar{r}_{jk}^t$, $\hat{r}_{j,jk}^{\tau,t}$, etc., as our decision variables. For example, the simple policy of shipping the expected imbalance level can be written as $\bar{r}_{jk}^t = E[n_{kj}^t - n_{jk}^t]$ and $\hat{r}_{\cdot,jk}^{\cdot,t} = 0$. However, we note that the above affine policies are not guaranteed to always satisfy the nonnegativity and capacity constraints 7.30 and

7.31. Therefore, we should instead consider the following *truncated* affine policies:

$$r^t_{jk}(n^1_{jk},...,n^{t+l_e}_{jk},n^1_{kj},...,n^{t+l_e}_{kj}) = \min\left\{\max\left\{\bar{r}^t_{jk} + \sum_{\tau=1}^{t+l_e} \hat{r}^{\tau,t}_{j,jk} n^\tau_j + \hat{r}^{\tau,t}_{k,jk} n^\tau_k, 0\right\}, K - n^t_{jk}\right\}. \tag{7.32}$$

Under the (truncated) affine policy, we may obtain an approximate formulation. For notational brevity, we further define the following vector notation. Let $\boldsymbol{n} = \{n^1_{jk},...,n^T_{jk},n^1_{kj},...,n^T_{kj}\}$ be the vector of the demand random variables. Then, the affine policy can be written in vector form $r^t_{jk} = \bar{r}^t_{jk} + (\hat{\boldsymbol{r}}^t_{jk})'\boldsymbol{n}$. Let $\boldsymbol{e}^t$ be the unit vector with the t-th component equal to 1 and others equal to 0. Then, similar to (See and Sim 2010), we may approximate the stochastic programming formulation with the following robust optimization formulation:

$$\begin{aligned}
&\min \sum_{t=1}^{T} c(E\left[\bar{r}^t_{jk} + (\hat{\boldsymbol{r}}^t_{jk})'\boldsymbol{n}\right]^+ + E\left[\bar{r}^t_{kj} + (\hat{\boldsymbol{r}}^t_{kj})'\boldsymbol{n}\right]^+)\\
&+\sum_{t=1}^{T} hE\left[\bar{z}^t_j + (\hat{z}^t_j)'\boldsymbol{n} + \sum_{\tau=1}^{t}\left[-\bar{r}^{\tau-L}_{kj} - (\hat{\boldsymbol{r}}^{\tau-L}_{kj})'\boldsymbol{n}\right]^+ + \sum_{\tau=1}^{t}\left[\bar{r}^\tau_{jk} + (\hat{\boldsymbol{r}}^\tau_{jk})'\boldsymbol{n} - (K - n^\tau_{jk})\right]^+\right]^+\\
&+\sum_{t=1}^{T} hE\left[\bar{z}^t_k + (\hat{z}^t_k)'\boldsymbol{n} + \sum_{\tau=1}^{t}\left[-\bar{r}^{\tau-L}_{jk} - (\hat{\boldsymbol{r}}^{\tau-L}_{jk})'\boldsymbol{n}\right]^+ + \sum_{\tau=1}^{t}\left[\bar{r}^\tau_{kj} + (\hat{\boldsymbol{r}}^\tau_{kj})'\boldsymbol{n} - (K - n^\tau_{kj})\right]^+\right]^+\\
&+\sum_{t=1}^{T} pE\left[-\bar{z}^t_j - (\hat{z}^t_j)'\boldsymbol{n} + \sum_{\tau=1}^{t}\left[-\bar{r}^\tau_{jk} - (\hat{\boldsymbol{r}}^\tau_{jk})'\boldsymbol{n}\right]^+ + \sum_{\tau=1}^{t}\left[\bar{r}^{\tau-L}_{kj} + (\hat{\boldsymbol{r}}^{\tau-L}_{kj})'\boldsymbol{n} - (K - n^{\tau-L}_{kj})\right]^+\right]^+\\
&+\sum_{t=1}^{T} pE\left[-\bar{z}^t_k - (\hat{z}^t_k)'\boldsymbol{n} + \sum_{\tau=1}^{t}\left[-\bar{r}^\tau_{kj} - (\hat{\boldsymbol{r}}^\tau_{kj})'\boldsymbol{n}\right]^+ + \sum_{\tau=1}^{t}\left[\bar{r}^{\tau-L}_{jk} + (\hat{\boldsymbol{r}}^{\tau-L}_{jk})'\boldsymbol{n} - (K - n^{\tau-L}_{jk})\right]^+\right]^+
\end{aligned} \tag{7.33}$$

subject to:

$$\begin{aligned}
\bar{z}^t_j &= \bar{z}^{t-1}_j + \bar{r}^{t-L}_{kj} - \bar{r}^t_{jk} \text{ for } t = 1,...,T\\
\bar{z}^t_k &= \bar{z}^{t-1}_k + \bar{r}^{t-L}_{jk} - \bar{r}^t_{kj} \text{ for } t = 1,...,T\\
\hat{z}^t_j &= \hat{z}^{t-1}_j + \boldsymbol{e}^{T+t-l_i-L} - \boldsymbol{e}^{t+l_e} + \hat{\boldsymbol{r}}^{t-L}_{kj} - \hat{\boldsymbol{r}}^t_{jk} \text{ for } t = 1,...,T\\
\hat{z}^t_k &= \hat{z}^{t-1}_k + \boldsymbol{e}^{t-l_i-L} - \boldsymbol{e}^{T+t+l_e} + \hat{\boldsymbol{r}}^{t-L}_{jk} - \hat{\boldsymbol{r}}^t_{kj} \text{ for } t = 1,...,T.
\end{aligned}$$

The objective function consists of six components. The first two correspond to shipping costs, the next two correspond to holding costs and the last two reflect shortage costs. The setup of the penalty cost terms warrants some explanation (the setup of the holding cost terms is analogous):

$$\sum_{t=1}^{T} pE\left[-\bar{z}^t_j - (\hat{z}^t_j)'\boldsymbol{n} + \underbrace{\sum_{\tau=1}^{t}\left[-\bar{r}^\tau_{jk} - (\hat{\boldsymbol{r}}^\tau_{jk})'\boldsymbol{n}\right]^+}_{\text{Correction for Negative Outbound Flow from j}}\right.$$

$$+\underbrace{\sum_{\tau=1}^{t}\left[\bar{r}_{kj}^{\tau-L}+(\hat{\boldsymbol{r}}_{kj}^{\tau-L})'\boldsymbol{n}-(K-n_{kj}^{\tau-L})\right]^{+}}_{\text{Correction for Inbound Flow to j due to Vessel Capacity}}\Bigg]^{+}$$

and

$$\sum_{t=1}^{T} pE\Bigg[-\bar{z}_k^t-(\hat{z}_k^t)\boldsymbol{n}+\underbrace{\sum_{\tau=1}^{t}\left[-\bar{r}_{kj}^{\tau}-(\hat{\boldsymbol{r}}_{kj}^{\tau})\boldsymbol{n}\right]^{+}}_{\text{Correction for Negative Outbound Flow from k}}$$

$$+\underbrace{\sum_{\tau=1}^{t}\left[\bar{r}_{jk}^{\tau-L}+(\hat{\boldsymbol{r}}_{jk}^{\tau-L})\boldsymbol{n}-(K-n_{jk}^{\tau-L})\right]^{+}}_{\text{Correction for Inbound Flow to k due to Vessel Capacity}}\Bigg]^{+}$$

The two components represent the overall expected shortage costs of empty containers with variable adjustments at port j and k, respectively. Because these two terms are symmetric, we will just discuss the physical interpretation of the three terms in first component as follows.

To begin, note that $\left(-\bar{z}_j^t-(\hat{z}_j^t)'\boldsymbol{n}\right)$ reflects, following a affine decision rule, the negative part of net inventory level at port j at time t. When the net inventory level is negative, this term reflects the shortage of empty containers at the port, which leads to penalty costs. However, this term assumes that shipments exactly follow the affine decision rule, and does not account for the truncation applied in 7.32 due to constraints 7.30. Therefore, whenever the affine repositioning shipments violate the nonnegativity constraints or the capacity constraints, correction terms need to be applied.

The second term $\sum_{\tau=1}^{t}\left[-\bar{r}_{jk}^{\tau}-(\hat{\boldsymbol{r}}_{jk}^{\tau})'\boldsymbol{n}\right]^{+}$ accounts for the adjustments needed when the affine decision rule stipulates that $r_{jk}^t<0$, which underestimates the outbound repositioning flows from j to k, and thus the shortages at j. Finally, when the affine decision rule suggests a repositioning flow from k to j that exceeds the vessel capacity, the excess amount is truncated following 7.32. This potentially leads to underestimations of shortages, which are corrected by the term $\sum_{\tau=1}^{t}\left[\bar{r}_{kj}^{\tau-L}+(\hat{\boldsymbol{r}}_{kj}^{\tau-L})'\boldsymbol{n}-(K-n_{kj}^{\tau-L})\right]^{+}$.

The constraints of the robust model are derived from the flow conservation constraints 7.28 and 7.29. In particular, flow balance holds when conservation holds for individual coefficients of the affine decision rule. This formulation effectively approximates the stochastic programming formulation 7.27. The methodology is discussed in detail in (See and Sim 2010) and is briefly reviewed in Sect. 7.3. In particular, we use the following result, which can be derived following the proof of Theorem 2 (available in (See and Sim 2010)):

Proposition 1. *The optimal solution to 7.33 gives a feasible solution to 7.27. Furthermore, the optimal objective value of 7.33 is an upper bound on that of 7.27.*

The difficulty of solving 7.33 mainly arises from evaluating the expectation terms in the objective function, which reflect the expected shortages. Using Theorem 2, the expectation terms can be tightly bounded above using a convex, second-order cone programming (SOCP) formulation, given the mean, variance and support of the demand distribution. Such a formulation is *distributionally-robust* in the sense that the upper bound holds for any distribution of demand with the given descriptive statistics. Therefore, we may obtain a SOCP approximation for the stochastic rebalancing problem, which can be solved using commercial solvers such as CPLEX or MOSEK.

7.5 Computational Experiments

In this section, we evaluate the performance of the distributionally-robust formulation 7.33 by performing computational experiments. Because the original stochastic programming formulation is computationally intractable, we use several sensible but easy-to-implement policies as benchmarks. To evaluate the performances of individual policies, we simulate demand sequences and compute the resulting shortages of applying them. The SOCP instances are solved using the CPLEX solver.

7.5.1 Benchmark Policies and Experimental Setup

We first give the descriptions of the benchmark policies. The first policy, *demand difference* (DD) refers to shipping a repositioning flow equal to the net difference between the realized shipping demand of laden containers in the two directions. That is, we set $r^t_{jk} = [n^t_{kj} - n^t_{jk}]^+$ and $r^t_{kj} = [n^t_{jk} - n^t_{kj}]^+$. This policy is simple, and effectively offsets imbalance of demand in the same period. The second policy, *moving average* for N periods of demand difference (MA-N) refers to repositioning the N-period moving average of the demand surplus. That is, $r^t_{jk} = \sum_{\tau=0}^{N-1} [n^{t-\tau}_{kj} - n^{t-\tau}_{jk}]^+/N$ and $r^t_{kj} = \sum_{\tau=0}^{N-1} [n^{t-\tau}_{jk} - n^{t-\tau}_{kj}]^+/N$. This policy uses a moving average to project the amount of imbalance to be countered by repositioning flows. Finally, the *expected difference* (ED) policy corresponds to shipping a repositioning flow equal to the expected demand surplus. That is, $r^t_{jk} = (E[n^t_{kj}] - E[n^t_{jk}])^+$ and $r^t_{kj} = (E[n^t_{jk}] - E[n^t_{kj}])^+$. This policy attempts to maintain balanced flow over the long run.

We also test the policy derived from the robust optimization formulation. In particular, the *robust policy* (RO) refers to the optimal solution to 7.33, in which the

Table 7.1 Parameter settings for data sets

Mean demand at j	200
Mean demand at k	150
Initial inventory at j	600
Initial inventory at k	300
Capacity of vessel	500
Period of planning horizon	15
Lead time	3
Forward & backward deviation at j	60
Forward & backward deviation at k	45
Coefficient of variation	0.3
Shipping cost (c)	1 to 5
Shortage penalty (p)	c to 8
Inventory holding (h)	1

repositioning flows are determined from the all uncertainty (demand at all ports) terms.

In our parameter settings, we set the means of laden container shipping demand in the two directions to be different. This helps evaluate the performances of policies with underlying trade imbalance in demand.

The performances of all aforementioned policies are evaluated under 50 replications (samples) of randomly-generated demand data. The planning horizon T in all cases is fixed to 15 periods. The shipping lead time L is set to 3 periods, and the hinterland operations lead times l_i and l_e are assumed to be 0 for simplicity. For MA-N, we use $N = 2, 5$ and 8 for illustration purposes. For all instances, we assume demand to be normally distributed with coefficient of variation of 0.3, and we vary the means. We normalize the holding (storage) cost for laden containers at ports as 1. Note that the shipping cost for empty containers is naturally higher than storage costs (otherwise, leftover containers will be shipped just for avoiding storage costs); thus we vary c between 1 and 5. Similarly, the penalty cost for shortage is higher than the shipping cost (otherwise, it is never desirable to reposition containers to avoid shortage), and thus we vary p between the value of c to 8.

Based on the above setting, we perform numerical experiments on 65 data sets. In particular, we vary cost of shipping, shortage penalty and inventory holding. The parameter settings are summarized in Table 7.1.

7.5.2 Results and Discussion

The experimental results are reported in Table 7.2. In particular, the third column presents the expected total cost of the robust policy, RO; the next five columns report the expected total costs of the three (shipping difference, moving averages

Table 7.2 Total shipping, inventory holding and shortage penalty cost

c	p	RO	DD	MA-2	MA-5	MA-8	ED	Min	Max
1	4	17324	22218	21071	21500	22184	24362	17324	24362
1	5	24007	29474	27911	28844	29282	28930	24007	29474
1	6	26727	31151	30446	33406	35385	33851	26727	35385
1	7	34783	37561	35454	37029	38372	38664	34783	38664
1	8	39458	42890	41135	43950	45965	46345	39458	46345
2	4	18754	21608	20688	21910	23192	19875	18754	23192
2	5	22075	28451	26453	25859	26188	29337	22075	29337
2	6	29689	32876	31480	31475	32242	33540	29689	33540
2	7	31748	36458	34624	36198	37452	38611	31748	38611
2	8	37766	43997	42589	44173	45201	41063	37766	45201
3	4	25101	26119	24502	23973	24170	27188	23973	27188
3	5	23339	29480	27092	26203	26347	30416	23339	30416
3	6	28440	36285	33728	34658	36002	33424	28440	36285
3	7	22100	37113	33130	31664	31030	43644	22100	43644
3	8	30327	43768	40008	36873	35770	59832	30327	59832
4	4	19997	24820	22842	22641	23180	24528	19997	24820
4	5	31563	33549	31593	30538	30366	27361	27361	33549
4	6	26780	35715	33016	32059	31888	35796	26780	35796
4	7	36654	42001	39445	40459	41981	39387	36654	42001
4	8	40655	51204	48348	46480	46182	46969	40655	51204
5	5	25330	30903	27911	27369	28063	31157	25330	31157
5	6	36856	40899	38670	38342	39047	36457	36457	40899
5	7	35003	41545	39251	40493	41343	41634	35003	41634
5	8	42707	50713	46398	43796	44179	45300	42707	50713

and expected difference) benchmark policies; and the last two columns report the minimum (best) and maximum (worst) cost out of the policies tested. Furthermore, to evaluate the relative performances of the individual policies against the best policy, we compute the relative percentage errors of the expected costs relative to the best achieved by any policy in the scenario. The results are reported in Table 7.3.

From the results, we observe that the RO model, i.e., the optimal solution to 7.33, performs the best in almost all scenarios tested. It achieves the best (smallest) expected cost in most instances. Among the heuristic policies, DD performs the relatively poorly on average. The average percentage difference between DD and the best policy is over 20 % and can be up to 67 %. This suggests that attempting to directly eliminate the demand imbalance in each period by immediately shipping an offsetting flow is undesirable. While this is an adaptive policy that accounts for

Table 7.3 Relative percentage deviation to minimum cost

c	**p**	**RO (%)**	**DD (%)**	**MA-2 (%)**	**MA-5 (%)**	**MA-8 (%)**	**ED (%)**
1	4	0.00	28.25	21.63	24.11	28.05	40.63
1	5	0.00	22.77	16.26	20.15	21.97	20.51
1	6	0.00	16.55	13.91	24.99	32.39	26.65
1	7	0.00	7.99	1.93	6.46	10.32	11.16
1	8	0.00	8.70	4.25	11.38	16.49	17.46
2	4	0.00	15.22	10.31	16.83	23.66	5.98
2	5	0.00	28.88	19.83	17.14	18.63	32.89
2	6	0.00	10.74	6.03	6.01	8.60	12.97
2	7	0.00	14.83	9.06	14.01	17.97	21.62
2	8	0.00	16.50	12.77	16.96	19.69	8.73
3	4	4.71	8.95	2.21	0.00	0.82	13.41
3	5	0.00	26.32	16.08	12.28	12.89	30.33
3	6	0.00	27.58	18.59	21.86	26.59	17.52
3	7	0.00	67.93	49.91	43.27	40.40	97.48
3	8	0.00	44.32	31.92	21.58	17.95	97.29
4	4	0.00	24.12	14.23	13.22	15.92	22.66
4	5	15.36	22.62	15.47	11.61	10.98	0.00
4	6	0.00	33.36	23.29	19.71	19.07	33.67
4	7	0.00	14.59	7.61	10.38	14.53	7.46
4	8	0.00	25.95	18.92	14.33	13.60	15.53
5	5	0.00	22.00	10.19	8.05	10.79	23.01
5	6	1.10	12.19	6.07	5.17	7.10	0.00
5	7	0.00	18.69	12.13	15.68	18.11	18.94
5	8	0.00	18.75	8.64	2.55	3.45	6.07
	Average:	0.88	22.41	14.64	14.91	17.08	24.25

random realizations of demand, it tends to overreact to imbalances. With significant shipping lead times, this policy causes a large number of containers being tied up on vessels without being available to meet demand at ports. On the other hand, the non-adaptive policy of ED, which does not react to random fluctuations of demand and ships a repositioning flow equal to the expected difference of demands, performs similarly poorly. The average performance gap versus the best policy is over 20 % and the worst case is over 97 %. Therefore, it is desirable for the repositioning policy to be moderately adaptive to demand realizations.

We observe that this is indeed the case for the moving average policies, which responds to random realizations of demand in the current and recent periods. Taking moving averages helps smooth out extreme fluctuations of demand, and avoids overreacting to such realizations. While these partially adaptive policies perform better, their performances are still significantly worse than the robust policies on average, nevertheless. This suggests that the mathematical programming approach can help obtain desirable responsive policies that are adaptive and does not overreact.

7.6 Conclusions and Future Research

In this chapter, we propose a distributionally-robust optimization formulation for the two-port empty container inventory control problem. This approach allows us to reformulate the computationally-intractable multi-stage stochastic program into a tractable second-order cone program that can be efficiently solved with commercial solvers. Based on computational experiments, it is demonstrated that solution of robust model outperforms other simple policies, by achieving lower total shortages. Relatively, solution of the ED policy is the worst among the simple policies.

For future research, we would extend current model to multiple ports on single and (or) multiple routes. We also plan to compare the performances of our approach against other approximations and heuristics proposed in the literatures.

References

Agarwal, R., & Ergun, O. (2008). Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science, 42*(2),175–196.

Ang, M., Lim, Y. F., & Sim, M. (2012). Robust storage assignment in unit-load warehouses. *Management Science, 58,* 2114–2130.

Ben-Tal, A., Goryashko, A., Guslitzer, E., & Nemirovski, A. (2004). Adjustable robust solutions of uncertain linear programs. *Mathematical Programming, 99,* 351–376.

Ben-Tal, A., Golany, B., Nemirovski, A., & Vial, J. P. (2005). Retailer-supplier flexible commitments contracts: A robust optimization approach. *Manufacturing & Service Operations Management, 7*(3), 248–271.

Chen, W., & Sim, M. (2007). Goal-driven optimization. *Operations Research, 57*(2), 342–357.

Chen, X., & Zhang, Y. (2009). Uncertain linear programs: Extended affinely adjustable robust counterparts. *Operations Research, 57*(6), 1469–1482.

Chen, X., Sim, M., & Sun, P. (2007). A robust optimization perspective on stochastic programming. *Operations Research, 55*(6), 1058–1071.

Chen, X., Sim, M., Sun, P., & Zhang, J. (2008). A linear decision-based approximation approach to stochastic programming. *Operations Research, 56*(2), 344–357.

Cheung, R. K., & Chen, C. Y. (1998). A two-stage stochastic network model and solution methods for the dynamic empty container allocation problem. *Transportation Science, 32* (2), 142–162.

Cimino, A., Diaz, R., Longo, F., & Mirabelli, G. (2010). Empty containers repositioning: A state of the art overview. *Proceedings of the Spring Simulation Multi-Conference.*

Crainic, T. G., Gendreau, M., & Dejax, P. (1993). Dynamic and stochastic-models for the allocation of empty containers. *Operations Research, 41,* 102–126.

Dejax, P. J., & Crainic, T. G. (1987). A review of empty flows and fleet management flows in freight transportation. *Transportation Science, 21*(4), 227–247.

Dong, J. X., & Song, D. P. (2009). Container fleet sizing and empty repositioning in liner shipping systems. *Transportation Research Part E, 45*(6), 860–877.

Dyer, M., & Stougie, L. (2006). Computational complexity of stochastic programming problems. *Mathematical Programming Series A, 106,* 423–432.

Erera, A. L., Morales, J. C., & Savelsbergh, M. (2009). Robust optimization for empty repositioning problems. *Operations Research, 57*(2), 468–483.

Francesco, M. D., Crainic, T. G., & Zuddas, P. (2009). The effect of multi-scenario policies on empty container repositioning. *Transportation Research Part E, 45,* 758–770.

Francesco, M. D., Lai, M., & Zuddas, P. (2013). Maritime repositioning of empty containers under uncertain port disruption. *Computers & Industrial Engineering, 64,* 827–837.

Goh, J., & Sim, M. (2010). Distributionally robust optimization and its tractable approximations. *Operations Research, 58*(4), 902–917.

International Maritime Organization. (2012). *International Shipping Facts and Figures V Information Resources on Trade, Safety, Security, Environment.*

Lam, S. W., Lee, L. H., & Tang, L. C. (2007). An approximate dynamic programming approach for the empty container allocation problem. *Transportation Research, Part C*, 265–277.

Levi, R., Perakis, G., & Uichanco, J. (2011). The data-driven newsvendor problem - new bounds and insight. Working Paper, Massachusetts Institute of Technology.

Liu, C., Jiang, Z., Chen, F., Liu, X., Liu, L., & Xu, Z. (2010). Empty container repositioning—A review. *Proceedings of the 8th World Congress on Intelligent Control and Automation.*

Meng, Q., Wang, S., Andersson, H., & Thun, K. (2013). Containership routing and scheduling in liner shipping: Overview and future research directions. *Transportation Science*, Article in Advance

New York Times. (29. January 2006). *China trade unbalances shipping*.

See, C. T., & Sim, M. (2010). Robust approximation to multiperiod inventory management. *Operations Research, 58*(3), 583–594.

Song, D. P. (2005). Optimal threshold control of empty vehicle redistribution in two depot service systems. *IEEE Trans. On Automatic Control, 50*(1), 87–90.

Song, D. P., & Carter, J. (2009). Empty container repositioning in shipping industry. *Maritime Policy & Management, 36*(4), 291–307.

Song, D. P., & Dong, J. X. (2008). Empty container management in cyclic shipping routes. *Maritime Economics & Logistics, 10,* 335–361.

Song, D. P., & Dong, J. X. (2012). Cargo routing and empty container repositioning in multiple shipping service routes. *Transportation Research Part B, 46,* 1556–1575.

Song, D. P., & Dong, J. X. (2015). Empty Container Repositioning. In C.-Y., Lee & Q., Meng (Eds.), *Handbook of Ocean Container Transport Logistics*. Springer.

Song, D. P., & Earl, C. F. (2008). Optimal empty vehicle repositioning and fleet-sizing for two-depot services systems. *European Journal of Operational Research, 185,* 760–777.

Song, D. P., Dinwoodie, J., & Roe, M. (2007). Integrated vehicle fleet-sizing, leasing and dispatching policy in a shuttle service system. *International Journal of Logistics Research and Applications, 10*(1), 29–40.

Chapter 8
Disruption Management for Liner Shipping

Xiangtong Qi

Abstract In many operations management problems, including vessel scheduling in liner shipping, people need to make and announce an operations plan in advance, with tremendous efforts being paid to optimize the plan. When the plan is executed in real time, however, it is constantly subject to different unexpected disruptions, making the original plan sub-optimal or even infeasible. Therefore we have the need of dynamically revising the operations plan at the execution stage, a problem often referred to as disruption management. In the context of liner shipping, disruption events may include bad weather, unusual port congestion and even port closure, etc., with the direct consequence of delaying the vessels from their schedules. In this chapter, we will study how disruptions can be effectively managed in liner shipping. We will show how to model and formulate such problems, and present a few key results of the solution schemes and managerial insights observed.

8.1 Introduction

In operations research and management, one major challenge is how to effectively handle various uncertainties that exist ubiquitously. For example, a retailer needs to make an ordering decision before the demand is revealed, which leads to the classic newsboy problem. Even in a so-called make-to-order environment with a known deterministic demand, a production manager needs to ensure enough raw materials or components on stock because some suppliers may have a random lead-time or yield rate for delivery. In the service sector, the customer arrival process and service times are always stochastic. In the transportation and logistics industry, people constantly face delays of trains, buses, and flights.

Uncertainties originate from many sources with different forms, frequencies, and degrees of significance. In general, they can be classified from different perspectives. Here we consider two categories of uncertainties based on the frequency of their occurrence and people's knowledge to them, namely, *recurring and regular*

X. Qi (✉)
Department of Industrial Engineering and Logistics Management, The Hong Kong University of Science and Technology, Kowloon, Clearwater Bay, Hong Kong
e-mail: ieemqi@ust.hk

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_8

uncertainties, and *rare and irregular uncertainties* which will be referred to as *disruptions.* For the former category of regular uncertainties, because they occur frequently, people can learn from historical data and establish some probabilistic models to quantify and predict their occurrence. For the latter category of disruptions, it is virtually impossible to obtain any probability-based models to describe their occurrence, though in some situations it is known that they will happen at some time. As a consequence, the way of coping with these two types of uncertainties would be different. For example, when making operational plans, people usually can proactively incorporate the impact of regular uncertainties, but hardly consider those disruptions due to insufficient information.

We may explain these concepts by a simple example in the daily life. Suppose that someone goes to work by driving a private car. Then the major regular uncertainty may be the possible traffic jam on the road. It is a common sense that some extra time should be planned against the traffic in a proactive way when we decide the time to leave home. However, no one will plan any extra time for disruptions such as failing to start up the car in the morning, though everyone knows that this may happen someday. Instead, people rely on real-time reactive decisions after the occurrence of the disruption, for example, to take a bus or call for a taxi, depending on the urgency of a specific day.

People start to recognize the importance of disruption management from the airline industry. Due to the fierce competition, airlines operate on highly, if not over, optimized schedules that aim to use the resources in the most efficient way. Unfortunately, the airline industry is also the one that is very sensitive to various disruptions such as bad weather conditions, not to mention the potential terrorist attacks. Clearly the impact of such disruptions cannot be taken into account in making a flight schedule, and airlines need a solution that can quickly adjust the flight schedules after knowing the occurrence of a specific disruption. For the research on airline disruption management, we refer to Thengvall et al. (2000), Clausen et al. (2010), and the work cited therein.

Following the concept of disruption management in airlines, people have also started to study similar problems in other areas, for example, production planning and scheduling, project scheduling, and supply chain management. For details, we refer to Yu and Qi (2004) for a comprehensive introduction. Nevertheless, how to effectively deal with disruptions is still a largely unexplored topic in liner shipping, with only a few work done recently (Brouer et al. 2013; Li et al. 2014a). The purpose of this chapter is to provide a brief overview for these problems.

The rest of this chapter is organized as follows. In Sect. 8.2, we present a few cases showing how disruptions may affect the liner shipping operations. In Sect. 8.3, we review the related work at the tactical level, how to design a vessel routing network that can tolerate certain regular uncertainties. In Sects. 8.4 and 8.5, we present models for handling disruptions at the operational level, for the case of a single vessel and the case of multiple vessels in a network, respectively. In Sect. 8.6, we make a few discussions.

8.2 Uncertainty and Disruption in Liner Shipping

The basic characteristic of liner shipping is that each vessel has a pre-announced schedule, in terms of arrival and departure times, to visit a series of ports of call. All stakeholders of liner shipping, including not only the carriers but also terminal operators, cargo shippers and freighter forwarders, expect that a vessel follow its schedule whenever possible. Unfortunately, in the practice of liner shipping, deviation from schedule has been a norm rather than exception (Notteboom 2006). For example, according to a research and advisory organization for the maritime sector, Drewry Maritime Research (http://www.drewry.co.uk), the average deviation between the schedule and the actual arrival is about 0.6 days.

While in the past carriers were mainly competing on vessel capacity and speed, in recent years major carriers have also started to pay efforts on improving their service reliability. In Drewry *Carrier Performance Report*, the shipping on-time score stayed in the 70–80 % range for the whole year of 2012, a significant improvement compared with the fact that it had not exceeded 69 % in any quarter since 2005, the time when Drewry started the statistics.

Vernimmen et al. (2007) discuss different practical causes of the schedule instability in liner shipping. The same as in other sectors, a vessel may be delayed by both regular uncertainties and disruptions. For example, some terminals are known to be easily congested where it is possible to draw a distribution of the random congestion delay from historical statistics. Then a shipping line can plan a preventive time buffer against the delay when designing the vessel schedule. On the other hand, a typhoon may close a port for one or even multiple days, causing a typical disruption. It is not possible to forecast the arrival of a typhoon, as well as its severity, at any port when a shipping line makes a vessel schedule that will be used for a couple of months. Hence we have to rely on real-time operation decisions to recover the delay caused by a disruption.

Only in the past two years, the maritime logistics industry has experienced a number of severe disruptions.

- In January 2012, high winds caused breaks in services in Felixstowe and Southampton, the two largest container hubs in UK.
- In February 2012, the pilot at Antwerp stopped working from 0800 h to 1700 h on a day by day basis against a change in national pension regulations, affecting MSC's container services over 21 vessels.
- In November 2012, a hurricane forced container terminal operations at New York-New Jersey closed for 1 week.
- In November 2012, the worst flooding in more than five decades caused severe disruption in container traffic at Thailand's main port Laem Chabang.
- In spring 2013, a 40-day labor strike at Hong Kong terminal reduced the dock operations capacity by 20 %, but that already delayed many vessels by 2–4 days, even making some vessels skipping calls at Hong Kong.
- In September 2013, a failure of a quay-crane's gearbox at the DP World Port Botany terminal caused sudden and unforeseen slot cancellations and significant terminal disruption.

- In early 2014, the snow storms in northeast America closed ports in NJ/NY multiple times, not only during the storm hits but also the aftermath, reporting on average a week or ten days delay to deliveries.

The above facts underscore the importance of effective disruption management, a well taken practice in airlines but still relatively new to container shipping lines. Since liner shipping and airline industry both are in the transportation sector, we can observe a few similarities between them; for example, the plan is for a set of vehicles (airplanes and vessels) each having a sequence of places to visit, and the consequence of a disruption is to delay the visiting time at each place. As a result, people may assume that they can follow similar strategies to handle disruptions, which however is not true. There are still a number of major differences between the two industries.

First, airlines commonly use the option of swapping aircraft for rescheduling flights. This is feasible because an airplane becomes empty after arriving at an airport. However, swapping vessels is impossible in liner shipping because most containers stay on the vessel at a port.

Second, airlines also extensively cancel some flights and are able to resume the remaining flights, partially due to the flexibility of aircraft swapping. In liner shipping, the decision of skipping a port is similar to flight cancellation, but it has to be the same vessel to continue the remaining journey.

Third, speeding up an airplane is not regarded as effective in airline disruption management except for the case of a short delay because all airplanes have been planned at high speeds. On the contrary, speeding up the vessels would be a major strategy in liner shipping, especially when nowadays most shipping lines take the practice of slow steaming (c.f. Cariou 2011; Corbett et al. 2009; Psaraftis and Kontovas 2013).

Fourth, crews are subject to strict legal and contractual constraints regarding their working hours. As a result, airlines often need to consider several interconnected problems such as scheduling aircraft and crew members separately. In liner shipping, such constraints do not exist.

Because of the above differences, we need to develop new disruption management models and solution schemes for container liner shipping.

8.3 A Brief Review of Robust Network Design for Liner Shipping

We first have an overview for the problems of designing liner shipping networks, with the emphasis on models specified for handling various regular uncertainty factors. For comprehensive literature surveys, we refer to Agarwal and Ergun (2008), Christiansen (2013), Meng et al. (2014), Tran and Haasis (2013), and Wang et al. (2013a) on network design, and Fransoo and Lee (2013) on the critical role of ocean container transportation in the global supply chains.

To be concise, our review focuses on the modeling aspect and interesting insights, while skipping the technical solution details.

In the literature, there are different models with different scopes, but virtually all models share the following assumptions: When a vessel is delayed after departing from a port or in the sea, the vessel will speed up as much as possible, with the aim of reducing the delay or even completely eliminating the delay at the immediate next port. As we are to show later, such an assumption overlooks a few other possible decisions in practice such as skipping a port when necessary. In addition, sometimes it may be better to strategically allow some delay at a port to save fuel consumption, even if it is possible to completely eliminate the delay if the vessel takes its maximum speed. Therefore, even for a vessel schedule that has already be designed at the tactical level to tolerate some uncertainty, there is still opportunity for real-time adjustment of the schedule at the operational level, e.g., when the vessel is sailing on the sea.

Having said that, however, we should point out that it is probably the only practical approach to take the always-speed-up assumption for the tactical network design problem, at least due to the following reasons. First, the problem would be too complicated to tackle if too much real-time decisions are considered. Second, always-speed-up would be the best, also broadly taken in practice, real-time decision if a vessel faces a small delay. In other words, at the network design level people may be able to handle regular uncertainties by good planning, but have to leave the decision after an irregular disruption to real time. As a matter of fact, early work on liner shipping network design mainly focuses on deterministic optimization (see review of Agarwal and Ergun 2008). Only recently uncertain has been introduced in network design.

Wang and Meng (2012a) present a mixed integer nonlinear convex stochastic programming formulation to design a liner shipping network. The objective is to optimize the network cost, including the ship deployment cost and expected fuel cost, while requiring a perfect no-delay schedule at any port. Specifically, the model assumes that the uncertain service time at each port has an upper bound, and in designing the vessel schedule, the model enforces a vessel to arrive at a port on time even for the worst case where the vessel experiences the longest possible service time at the immediate preceding port. To achieve such a goal, a sea contingency time is designed between two consecutive ports, which gives more flexibility for a vessel to adjust its sailing speed. Using a case study with 46 ports and 11 routes, the paper shows that the sea contingency time may affect the vessel speed and deployment in a very complicated way, indicating the difficulty of the problems.

Following the above model, Wang and Meng (2012b) point out that it may be too costly, or even impossible, to enforce the vessel to be on time at all ports. Thus there is the need to allow certain delay at some ports with a delay penalty, while maintaining the strict on-time arrival at a few important ports. To address this issue they aim to minimize the total fuel and ship deployment cost, as well as the delay penalties, where the penalty cost is assumed to be a linear increasing function of delay when the delay is within a certain range, and becomes a constant after the delay exceeds the range. In the model, they propose a concept of target arrival time at each port, usually earlier than the announced arrival time, which is the ideal time for a vessel to arrive at a port in order to hedge against the uncertain port time. A mixed integer nonlinear stochastic programming formulation is presented and solved by sample average approximation and linearization techniques. One important finding from the model is that the planned sailing speed of a leg should usually be small if the leg has a short distance, which provides a theoretical explanation for the practices of shipping lines.

Qi and Song (2011) address the problem from a different perspective. Rather than design the entire shipping network, they consider the problem of re-optimizing a single route in a given network, specifically, to determine to planned arrival time at each port for the vessel without changing the routing sequence. This can be regarded as a local adjustment when the business environment changes only for some specific routes. In addition, focusing on a single route also enables them to analytically study the structural results of an optimal vessel schedule and derive useful managerial insights, which is hard to obtain from a complicated network model. For example, they quantitatively establish that the leg with longer distance should have a larger planned sailing time, but the allocation of the optimal time is disproportional to their distances, actually in favor of the shorter leg. This reveals that the same port uncertainty will have a more severe impact on a shorter leg than on a longer leg. Further, it identifies a set of sufficient conditions that guarantee the strict on time arrival at all ports, and formula to compute the maximum possible delays of the vessel at each port on the route when delays are allowed.

While the most studied uncertainty in liner shipping is vessel delay, there are other uncertainties that affect the business performance of a shipping line. For example, Meng et al. (2012) consider the uncertainty in container shipment volume and suggest that a shipping line can dynamically charter in or out vessels to respond the uncertain demand. They formulate the problem as a stochastic programming to maximize the expected profit. Wang et al. (2013b) extend the work by a joint chance constraint programming model to ensure a desired service level when facing uncertain container demand. Dong and Song (2012) study container fleet sizing problem in liner services with uncertain customer demands and stochastic inland transport times, with the aim of investigating the impact of uncertainties in inland times on the fleet sizing of the shipping lines.

In a larger picture, uncertainty and disruptions are regarded as a major source of maritime safety and security risks, categorized as risk associated with physical flow in a recent researches conducted by Chang et al. (2014). It is found that the risks associated with physical flows are generally more likely to have serious impacts on damages than others such as risks associated with information flow and payment flow. Also from the perspective of risk management, Lam (2012), Lam and Yip (2012) study the impact of port disruption on supply chains by a Petri net model. All these works highlight the importance of effective disruption management.

At the operational level, there has been much less published work on disruption management for liner shipping. In the next two sections, we will introduce the results from two known papers by Li et al. (2014a) and Brouer et al. (2013) respectively. There are a few other works that are related to our topic of disruption management but beyond our scope of vessel scheduling in liner shipping. We make a brief discussion here without discussing the details.

Paul and Maloni (2010) have studied a similar disruption management problem. They present a decision support system for rescheduling vessels after the occurrence of port disasters. However, the model is not developed for liner shipping in that no published vessel schedule is considered. Francesco et al. (2013) study the problem of repositioning of empty containers under certain port disruptions. Wang and Qi (2014) consider another special disruption where a vessel is chased by a pirate boat, and they study the problem of real-time changing the vessel sailing route and speed to evade the attack. Zhen (2014) addresses the yard operations problem of storage allocation in a transshipment hub, and proposes a real-time decision support system for coping with uncertainties caused by vessels. Finally, there is a large amount of work on berth allocation under uncertainty; see Zhen et al. (2011) and the references cited therein.

8.4 Disruption Recovery for a Single Vessel

We first discuss a relatively simple case where a single vessel is delayed at a port due to an unexpected event. Our focus is to study how the vessel can adjust its speed, together with other recovery options, in the following legs of the route so that it can catch up with the original schedule in the best way, with respect to the balance between the fuel cost and delay penalty in the following ports.

In practice, a major disruption such as a port closure may affect multiple vessels simultaneously, and it may be necessary to consider all affected vessels together. Despite of that, there is still a need of narrowing down the scope to the problem of a single vessel, due to a number of reasons. First, we do have problems with one vessel delayed by its own reason such as a mechanical malfunction. Second, it enables us to do an in-depth study so as to analytically understand the different behaviors of the optimal reschedule after a disruption, thus providing important managerial insights and guidance. Third, most properties from a single vessel will still hold, at least approximately, in multiple-vessel problems (but becoming hard to be analytically proven) because they are essentially the necessary conditions for the optimality of any individual vessel in a network.

8.4.1 Problem Description

We now introduce the problem with more details. There is a planned schedule for a vessel to depart from a port of call P_0 at time zero, to sequentially visit n ports of call $\{\mathrm{P}_1, \mathrm{P}_2, \ldots, \mathrm{P}_n\}$. We use segment j to refer to the trip from port P_{j-1} to port P_j. The given planned schedule can be described by the notation below.

d_j	The distance of segment j
$\bar{T}_j$	The arrival time at port P_j in the planned schedule
s_j	The port time at port P_j, i.e., the time for cargo loading and unloading

To model the impact of a disruption, suppose that the vessel is late when leaving port P_0 where the actual departure time is $\Delta > 0$. To reduce the delay penalty at the following ports, we need a recovery schedule so that the delay can be reduced. To this end, we introduce a delay penalty cost function $F_j(\delta)$ for each port P_j with delay δ, where the penalty is nondecreasing of δ for each P_j.

Note that the function $F_j(\delta)$ can measure a number of negative factors caused by the delay, for example, a real monetary cost paid to the shippers, a loss of goodwill or reputation, and an additional cost incurred by missing connections for some container transshipment. $F_j(\delta)$ may take different forms depending on the emphasis. When the monetary penalty or goodwill loss is dominant, we may reasonably assume $F_j(\delta)$ is convex increasing which includes linearly increasing as a special case. When there are a large number of container transshipment at a port P_j, then $F_j(\delta)$ may be more like a stepwise function because containers under transshipment usually belong to different groups and each container group has a different due date of being loaded to its next vessel.

In one simple case for the recovery schedule, the vessel has the option of speeding up at certain segments so that the delays can be reduced with more fuel consumption (than planned). We use the following notation to define the speed-up decision in the recovery schedule.

v_j	The speed of segment j in the recovery schedule
t_j	The travel time spent on segment j in the recovery schedule
T_j	The arrival time at port P_j in the recovery schedule

To measure the fuel consumption cost, we define the fuel consumption cost function $f_j(v_j)$ of segment j, per nautical mile at speed v_j. The following assumption is broadly used in vessel scheduling and consistent with the ship building technology.

Assumption 1 The fuel consumption cost function $f_j(v_j)$ is convex and increasing of v_j in its designed speed range $v_{\min} \le v_j \le v_{\max}$.

The decision of speeding up the vessel can be formulated by a nonlinear programming (NLP) as follows.

$$\min \sum_{j=1}^{n} d_j f_j(v_j) + \sum_{j=1}^{n} F_j(T_j - \bar{T}_j) \tag{8.1}$$

$$s.t. \qquad v_j t_j = d_j, j = 1, ..., n.$$

$$T_j = \Delta + \sum_{i=1}^{j} (s_{i-1} + t_i), j = 1, ..., n.$$

It can be shown that NLP (8.1) is a convex programming when the delay cost function $F_j(\delta)$ is convex. So it can be solved easily by any NLP solver for this special case. For the general case, it can be solved by dynamic programming that is to be presented below.

Besides speeding up, the vessel has another option, port skipping, under which the vessel goes from a port P_i to P_j, $i < j + 1$, directly without visiting the in-between ports $P_{i+1}, \ldots, P_{j-1}$ in the planned schedule. We need the following notation to define the option of port skipping, for $j = i + 1, i + 2, \ldots, n$.

D_{ij}	The direct distance from port P_i to P_j
$f_{ij}(v_{ij})$	The fuel consumption function of the direct sail from port P_i to port P_j at speed v_{ij}
S_{ij}^i	The extra port time incurred at port P_i for skipping ports $P_{i+1}, \ldots, P_{j-1}$
S_{ij}^j	The extra port time incurred at port P_j for skipping ports $P_{i+1}, \ldots, P_{j-1}$
C_{ij}	The cost of skipping ports $P_{i+1}, \ldots, P_{j-1}$

We make a few comments on the above defined S_{ij}^i, S_{ij}^j and C_{ij} that are jointly used to quantify the impact caused by port skipping. Depending on the level of consideration, there may be different issues to include.

For example, the cargo to be discharged at a skipped port P_{i+1} has to be discharged at either port P_i or P_j with additional transportation arrangement that moves the cargo from P_i or P_j to P_{i+1}. Hence the port times at P_i and P_j may be increased, which can be captured by S_{ij}^i and S_{ij}^j, respectively. Besides discharging cargo at different ports, we may also consider the opportunity cost of losing business at the ports skipped, and the possibility of loading more cargo at port P_j because of some unoccupied vessel capacity originally planned to be loaded at those skipped ports. All such costs are included in C_{ij}, an aggregated cost incurred by skipping ports $P_{i+1}, \ldots, P_{j-1}$.

Assumption 2 The impact by port skipping, S_{ij}^i, S_{ij}^j and C_{ij} can be evaluated by an external procedure. In addition, S_{ij}^i, S_{ij}^j and C_{ij} are increasing of j.

Finally, there is the third option, vessel swapping, where the vessel may change the sequence of visiting the ports. Although theoretically speaking any new sequence may be possible, which leads to a highly complicated combinatorial optimization problem, practically we only need to consider swapping the ports close to each other, which we will call a localized port swapping. A localized port swapping is good enough because of the large geographic area covered by a vessel route, making ports naturally clustered by their mutual distances.

Under localized port swapping, we assume that the entire route of n ports can be partitioned into m sequential subroutes, and port swapping may only incur within each subroute. If an optimal reroute by port swapping can be efficiently derived within a subroute, then the entire route can be optimized at the subroute level by relinking the subroutes sequentially.

We use the following example to illustrate the concept of localized port swapping. Consider an Asia-Europe round trip route operated by Maersk shown in Fig. 8.1. Note that in the route some ports are called twice in the route, and they are regarded as two ports of call, e.g., Yantian port is both ports P_4 and P_{17}.

In view of the above route, we may identify four subroutes as shown in the dashed boxes, (P_0, P_1), (P_2, P_3, P_4), (P_9, P_{10}, P_{11}) and (P_{17}, P_{18}). For example, with respect to subroute (P_2, P_3, P_4), it is possible to have a reroute as $P_1 \rightarrow P_3 \rightarrow P_2 \rightarrow P_4 \rightarrow P_5$, or even $P_1 \rightarrow P_3 \rightarrow P_4 \rightarrow P_2 \rightarrow P_5$, but practically hard to have $P_1 \rightarrow P_2 \rightarrow P_3 \rightarrow P_5 \rightarrow P_4 \rightarrow P_6$,

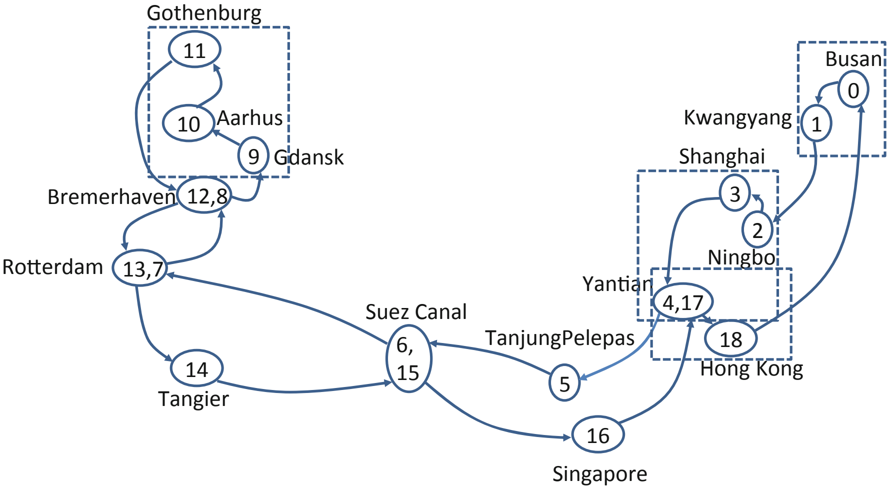


Fig. 8.1 An Asia-Europe route with examples of subroutes

due to the longer distance between P_4 (Yantian, China) and P_5 (Tanjung Pelepas, Malaysia).

8.4.2 Structural Results

We now introduce a few structural properties of an optimal recovery schedule when the vessel takes the option of speeding up. We only present the results with respect to the implications. The formal statement and detailed proof can be found in Li et al. (2014a). These results help us understand the insights of the problems; for example, why port skipping and swapping are needed on top of vessel speeding up. They may also be helpful to develop simple heuristic solutions when the mathematical optimization is unavailable or inappropriate.

Result 1. If two consecutive segments j and $j+1$ have the same fuel consumption function $f_j(v) = f_{j+1}(v)$, then we have the optimal v_j and v_{j+1} satisfying $v_j \geq v_{j+1}$ before the vessel catches up with the original schedule.

This result shows a trend of decreasing speed in the recovery schedule. The intuition behind is actually easy to understand. In words, if the journey is delayed, it is better to speed up more in early segments so that the delay will not be propagated too much. This is true regardless of the length of the segments, the degree of the delay, and the delay penalty function of the ports.

Result 2. If all delay penalty functions are convex, the total cost incurred by the initial delay Δ to the entire journey in an optimal recovery schedule is convex and increasing with Δ.

This result on convexity shows that the cost incurred by a delay Δ has an increasing margin when Δ increases. In other words, when Δ is small, it can be effectively handled by appropriately speeding up the vessel, but the marginal cost of speeding up increases when Δ becomes larger. This necessitates the need of considering other options such as port skipping and swapping to cope with larger disruptions.

Result 3. If all delay penalty functions are convex, at any port P_i, a larger delay will lead to a higher speed at the next segment.

This result shows a strategy that is consistent with intuition, which may seem straightforward or even trivial. However, this does not hold without any conditions, and may not be true for non-convex delay penalty functions. For example, under a stepwise function where the delay penalty is constant within certain intervals, it may lead to a phenomenon of "no need to hurry if too late" where, if the delay is just above one threshold, the vessel should speed up more in order to reduce delay penalty; but if the delay is too large, it is better to take a lower speed in order to save fuel consumption because the additional delay does not cause additional costs under pure stepwise delay cost functions.

The above Result 3 also raises the issue of defining a reasonable delay penalty function. In literature people indeed have proposed using a stepwise function to quantify delay cost, mostly motivated by the fact that a major cost source is due to the misconnection of the batch-by-batch container transshipment where each batch has a due date of transshipment. However, since the caused phenomenon of "no need to hurry if too late" may seem unreasonable in practice, we suggest that a more appropriate delay penalty function include both a convex increasing (or simply linear) term and a stepwise term, should the missed transshipment connection be considered.

Result 4. When all delay penalty functions are linear, there exists a threshold value Δ_0 such that for any initial delay $\Delta > \Delta_0$, the vessel never speeds up more than the case with delay Δ_0.

This result further underscores the importance of the options of port skipping and swapping when the delay is large. Specifically, when the delay is too large, the vessel will stop increasing its speed even if the speed has not reached the maximum speed limit. This may happen when the margin of delay penalty is lowered than the margin of fuel cost consumption. In such a case, only port skipping or swapping can help to further reduce the total cost.

8.4.3 Dynamic Programming Algorithm

The problem with the option of port skipping cannot be effectively handled by NLP techniques due to its combinatorial nature. We can develop a dynamic programming (DP) algorithm. The idea is outlined below.

Let $H(i, x)$ be the minimum total cost from the vessel leaving port P_i until arriving at P_n, given that the modified arrival time at port P_i is $\bar{T}_i + x$, where the modified arrival time refers to its physical arrival time plus a possible extra time needed to process the rerouted cargo due to the possible port skipping immediately before port P_i. Under this definition, the vessel shall depart from P_i at time $\bar{T}_i + x + s_i + S_{ij}^i$ assuming the next stop is at port P_j.

Then we have a DP recursion as

$$H(i,x) = \min_{j,y} \left\{ H(j,y) + D_{ij} f_{ij} \left(\frac{D_{ij}}{\left(\overline{T_j} + y - S_{ij}^j\right) - \left(\bar{T}_i + x + s_i + S_{ij}^i\right)} \right) \right.$$
$$\left. + C_{ij} + J_j \left(y - S_{ij}^j \right) \right\} \tag{8.2}$$

In the above DP (8.2), for each state (i, x), we enumerate all the possible combinations of the next arrival port P_j and y, implying ports $\mathrm{P}_{i+1}, \ldots, \mathrm{P}_{j-1}$ being skipped and the modified arrival time at P_j being $\bar{T}_j + y$. In particular, the term

$$D_{ij} f_{ij} \left(\frac{D_{ij}}{\left(\bar{T}_j + y - S_{ij}^j\right) - \left(\bar{T}_i + x + s_i + S_{ij}^i\right)} \right) \tag{8.3}$$

is the fuel cost consumed from P_i to P_j.

DP (8.2) starts from the initial conditions $H(n, x) = 0$, and the optimal recovery schedule can be obtained from H (0, Δ). To implement the DP computationally, we can discretize the time dimension x in the state space (i, x) into a finite number of time points for enumeration. Computational experiments show that the DP is time efficient even if the discrete time unit is as short as 10 min. This has achieved the necessary accuracy for practical needs.

The DP can be extended to deal with the option with localized port swapping, where the route is partitioned into a series of sequential subroutes and port swapping is allowed only within a subroute. The algorithm includes two layers: an inner layer that finds a best port swapping schedule for each subroute, and an outer layer that optimizes the schedule at the subroute level. For the inner layer, brute-force enumeration can be used to consider each possible sequence within the subroute because each subroute usually contains at most 5–6 ports in practice. Then DP (8.2) can be modified to link the subroutes to construct the entire route.

8.4.4 Insights from Computational Experiments

Li et al. (2014a) conduct a series of computational experiment that reveals useful managerial insights. For example, those structure results of an optimal recovery schedule, though they are analytically proven under certain assumptions, actually hold, at least approximately, when some assumptions are relaxed. In addition, some other properties are observed, some of which are highlighted here.

First, regarding the two options of port skipping and swapping, they need to be taken only when there is a sufficiently large delay, which means that the speed-up-only decision is sufficient enough to handle a small delay such as within 24 h.

The benefit of port skipping increases with the delay incurred, sometimes able to bring in up to 30 % cost saving compared with the speeding-up-only decision. As a comparison, the option of port swapping is less frequently taken in an optimal recovery schedule, and if taken, its extra cost reduction is often less than 10 %. This suggests that combining speeding up and port skipping would be effective enough for most cases.

Second, the option of port skipping is more valuable when the vessel has a planned speed close to its maximum speed limit or when the fuel price is high. Such an observation indicates one benefit of the practice of slow steaming, i.e., enhancing the route stability since port skipping causes a much higher deviation from the planned route.

Third, the value of port swapping is insensitive to factors such as the maximum vessel speed and fuel cost, partially due to the high cost of extra sailing distance caused by port swapping. This further confirms our earlier finding that port skipping is relatively more helpful over port swapping.

8.5 Disruption Recovery for Multiple Vessels

In many cases a major disruption may affect a number of vessels in a region, for example, multiple vessels being delayed at one port. When the vessels are independent of each other, each vessel can then be scheduled individually for their continuing journeys after leaving the port. In other cases, some vessels are coupled by container flows, for example, container transshipment from one vessel to another. In such cases, it is necessary to reschedule all effected vessels together. To this end, Brouer et al. (2013) have proposed an integer linear programming model based on a time-spaced network.

Taking the ideas of Brouer et al. (2013), we now develop a model that handles multiple vessels in a network. In the model, there are two different decisions, the vessel routing and speed decision, and the container flow routing. The model is an extension of Brouer et al. (2013) where they implicitly assume that containers will be transported by their original vessels even when the vessels are delayed, i.e., no container rerouting is considered.

A time-spaced network can be regarded as a multiple-layered expansion of the physical sailing network where each layer is a copy of the physical network associated with a specific time. Specifically, a node can be defined by a pair of (j, t) referring to port j and time t, and an arc $(j_1, t_1) \rightarrow (j_2, t_2)$ models a potential sailing segment from port j_1 with departure time t_1 to port j_2 with arrival time t_2. Therefore, for each technically feasible arc $e = (j_1, t_1) \rightarrow (j_2, t_2)$ the sailing distance and time are known based on which we can calculate the fuel cost. In addition, an arc $(j_1, t_1) \rightarrow (j_1, t_2)$ models the berth time at a port j_1 from time t_1 to time t_2.

For an arc $e = (j_1, t_1) \rightarrow (j_2, t_2)$, we let $J_1(e) = j_1$ denote its origin port, $J_2(e) = j_2$ its arrival port, $T_1(e) = t_1$ denote its departure time, and $T_2(e) = t_2$ its arrival time. In addition, let the bunker cost of vessel v on arc e be denoted by c_{ev}.

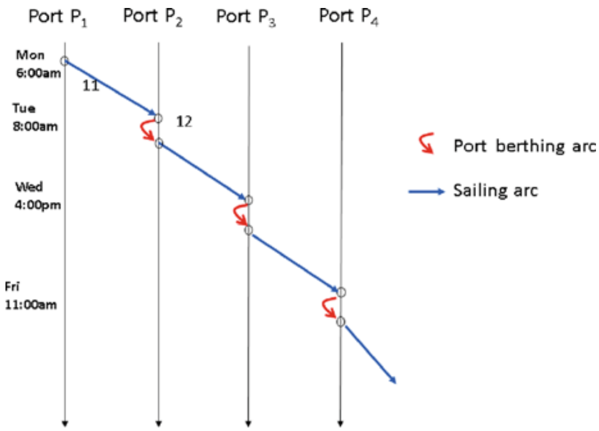


Fig. 8.2 A single vessel schedule represented in a time-space network

In Fig. 8.2, we give an example of a partial planned route of a single vessel. In the figure, all nodes associated with one port are placed in a vertical line, with upper nodes representing earlier times. The figure shows both sailing arcs and port berthing arcs. In the following figures, we will omit the port berthing arcs for better visibility.

Let V be the set of vessels where each v in V has a planned sailing schedule $H_v$that can be represented by a path on the time-space network. For each port h in H_v, we use L_{vh} to define the set of arcs of which the corresponding sailing is feasible for vessel v to take from port h to its next port in H_v. Note that the set L_{vh} can model all three possible decisions for a vessel, namely, speed adjustment, port skipping and port swapping, which is explained as follows.

First, form each port h to its next port of call, we can define multiple arcs in $L_{vh}$where each arc e in L_{vh} has different departure and arrival times, hence different sailing times between the two ports, thus modeling the different speed choices of the vessel. Note that the feasibility of a particular arc is determined by the speed range of the vessel.

Second, the end point of an arc in L_{vh} may link to a port different from the immediate next port of port h in the original route H_v. This gives the chance of modeling port skipping and port swapping. For example, suppose that ports h_1, h_2 and h_3 are three consecutive ports on the planned route of vessel v. Then an arc $(h_1,t_1)\rightarrow(h_3,t_3)$ models the schedule of skipping port h_2, and the path $(h_1,t_1)\rightarrow(h_3,t_3)\rightarrow(h_3,t_3+\beta_{v,\,h3})\rightarrow(h_2,t_2)$ represents a new route h_1, h_3 and h_2 after port swapping where $\beta_{v,\,h3}$ is the berth time of vessel v at port h_3.

Figures 8.3 and 8.4 demonstrate how different decisions can be modeled on a time-space network. Suppose the vessel is delayed when leaving from port P_0 (note the dashed arc represents the original schedule). The vessel can speed up at different speeds and arrive at P_2 at different times, or even skip port P_1, as demonstrated in Fig. 8.3; the vessel may also swap P_1 and P_2 as shown in Fig. 8.4. Note that in solving

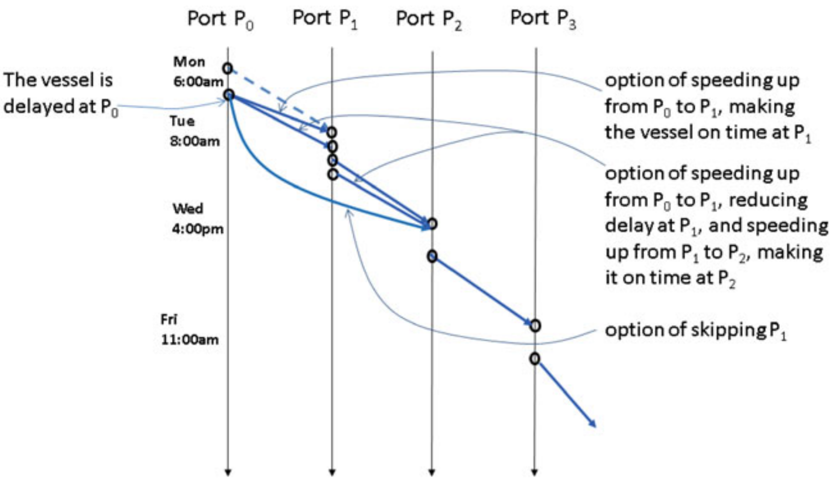


Fig. 8.3 The options of speeding up and port skipping

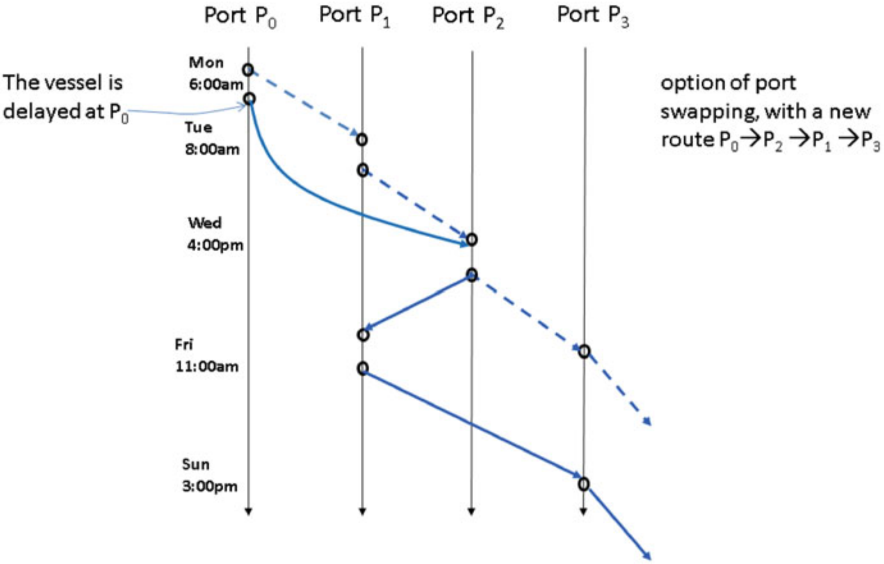


Fig. 8.4 The option of port swapping

the problem, we actually need to merge Figs. 8.3 and 8.4 into a single network and select the best path for each vessel.

We still need to consider the container flows to measure the impact of containers delay in the recovery schedule. To this end, we consider each container group g that is defined by (B_g, D_g, A_g), where B_g is the origin port, D_g is the final destination port, and A_g is the planned arrival time at final destination port D_g. Let function $F(.)$ measure the delay penalty for container group g.

We have the following decision variables

x_{ev} Binary variable defining whether an arc e will be selected by vessel v
y_{ge} Binary variable defining whether container group g is assigned to arc e

Then the total cost to be minimized is

$$\sum_{e,v} x_{ev}c_{ev} + \sum_{g} F_g \left(\sum_{e:J_2(e)\in D_g} T_2\,(e)\, y_{ge} - A_g \right). \tag{8.4}$$

We have the following constraints. For each vessel v, we need to specify a path, which can be enforced by flow balance at each port h on H_v, and time t. For h as an intermediate port on H_v, we need

$$\sum_{e:J_2(e)=h,T_2(e)=t} x_{ev} \quad = \quad \sum_{e:J_1(e)=h,T_1(e)=t+\beta_{vh}} x_{ev}. \tag{8.5}$$

The meaning of the above constraint is that if vessel v arrives at port h at time t, it leaves at time $t+\beta_{vh}$ where β_{vh} is the berth time. For h as the first port on H_v with vessel v departing at time Δ_v, we need

$$\sum_{e:J_1(e)=h,T_1(e)=\Delta_v} x_{ev} = 1. \tag{8.6}$$

For h as the last port on H_v, we need

$$\sum_{e:J_2(e)=h} x_{ev} = 1. \tag{8.7}$$

Considering the container flow routing, for each container group g and arc e, we need

$$y_{ge} \leq \sum\nolimits_{v} x_{ev} \tag{8.8}$$

and also at each port h and time t, if h is not B_g or D_g, we have

$$\sum_{e:J_2(e)=h,T_2(e)=t} y_{ge} \quad = \quad \sum_{e:J_1(e)=h,T_1(e)\geq t+\gamma_{gh}} y_{ge} \tag{8.9}$$

The former constraint enforces the feasibility of an arc e to group g, and the latter one is for flow balance where γ_{gh} is the minimum time needed for transshipment at port h. Finally, at B_g, the origin port of container group g, we have

$$\sum_{e:J_1(e)=B_g} y_{ge} = 1, \tag{8.10}$$

and at D_g, the destination port of container group g, we have

$$\sum_{e:J_2(e)=D_g} y_{ge} = 1. \tag{8.11}$$

These last constraints ensure all containers to be delivered from their respective origins to destinations.

The above formulation includes two inter-correlated multi-commodity network flow problems, which may be hard to solve for large-scale problems. Without considering the container flow rerouting, the problem has been shown to be directly solvable by an MIP solver in Brouer et al. (2013), for example, in a few seconds to solve real cases with up to 10 vessels and 33 ports. However, it is unknown whether this is still true when container flows are to be included. One possible approach may be to solve the vessel scheduling problem first, then do the container flow routing in the second stage.

The above formulation has ignored a few important factors that can actually be included in an extended model. For example, each vessel has a capacity limit, each port may have a fixed time window for access, and there is certain flexibility of serving or rejecting some container groups. Moreover, the port time is also controllable at certain cost in some ports, which gives additional room to reduce cost, and also makes the problem more complicated.

8.6 Conclusion and Future Directions

To summarize, uncertainties in liner shipping can be handled at two different levels, network design at tactical level and real-time recovery at the operational level. This chapter has focused on the latter case. Specifically, we consider the aftermath of a disruption that has already happened, aiming to mitigate the negative impact or loss caused by the disruption. We have introduced two major models, one for a single vessel and another one for multiple vessels in one network. While the first model enables us to dig into the decision making process, the second model has a broader applicability in practice.

As we have discussed, this is an emerging research field that has only a few initial results. There are much work to do for both theoretical values and practical relevance. We list a few directions below.

It is important to study operational proactive strategies for a known incoming disruption. For example, when a vessel is on the way of heading to a port, it is informed that a typhoon is forming around the port area. Then the vessel may face the decision of either speeding up so as to arrive at the port before the typhoon or purposely slowing down so as to arrive after the typhoon. The problem is challenging because of the randomness in the formation and evolution process of the typhoon; also, for any decision we have to consider its impact on the following ports of the route. This is a current on-going work of us (Li et al. 2014b).

Another worthy direction is to study the terminal operator's decision and its effects on shipping lines. For shipping lines it is assumed that port times are either fixed or sometimes controllable with cost. However, each terminal has its own decision of scheduling resources such as berths and cranes, which directly determines the port time of each vessel. Therefore, it may be necessary to include the terminal resource scheduling in vessel scheduling. Such an issue has been raised by Kontovas and Psaraftis (2011).

We would also like to emphasize the possible need of studying decentralized decision making where multiple shipping lines are involved. For example, a freight forwarder may arrange transshipment between vessels of two shipping lines. Ideally these two lines should consider synchronizing their vessel recovery schedules for the transshipment. However, because each shipping line has its own cost function to minimize, any centralized optimal scheduling may be hard to implement in practice. As such, it calls for models using game theoretic analysis besides pure optimization.

Last but not the least, we need to apply the models and results in the real applications. The success of disruption manage in the airlines is due to its proven value when being implemented by the industry (Yu et al. 2003). In addition more problems would be raised in the practice. For this aspect, we still have a long way to go in the liner shipping industry.

References

Agarwal, R., & Ergun, O. (2008). Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science, 42*(2), 175–196.

Brouer, B. D., Dirksen, J., Pisinger, D., Plum, C. E. M., & Vaaben, B. (2013). The vessel schedule recovery problem (VSRP)—A MIP model for handling disruptions in liner shipping. *European Journal of Operational Research, 224*(2), 362–374.

Cariou, P. (2011). Is slow steaming a sustainable means of reducing CO2 emissions from container shipping? *Transportation Research Part D, 16*(3), 260–264.

Chang, C. -H., Xu, J., & Song, D. -P. (2014). An analysis of safety and security risks in container shipping operations: A case study of Taiwan. *Safety Science, 63,* 168–178.

Christiansen, M., Fagerholt, K., Nygreen, B., & Ronen, D. (2013). Ship routing and scheduling in the new millennium. *European Journal of Operational Research, 228*(3), 467–478.

Clausen, J., Larsen, A., Larsen, J., & Rezanova, N.J. (2010). Disruption management in the airline industry-concepts, models and methods. *Computers & Operations Research, 37,* 809–821.

Corbett, J., Wang, H., & Winebrake, J. (2009). The effectiveness and costs of speed reductions on emissions from international shipping. *Transportation Research Part D, 14,* 539–598.

Dong, J.X., & Song, D. -P. (2012). Quantifying the impact of inland transport time on container fleet sizing in a liner shipping service with uncertainties. *OR Spectrum, 34*(1), 155–180.

Francesco, M.D., Lai, M., & Zuddas, P. (2013). Maritime repositioning of empty containers under uncertain port disruptions *Computers and Industrial Engineering, 64,* 827–837.

Fransoo, J.C., & Lee, C. -Y. (2013). The critical role of ocean container transport in global supply chain performance. *Production and Operations Management, 22*(2), 253–268.

Kontovas, C., & Psaraftis, H.N. (2011). Reduction of emissions along the maritime intermodal container chain: Operational models and policies. *Maritime Policy & Management, 38*(4), 451–469.

Lam, J. S. L. (2012). Risk management in maritime logistics and supply chains. In D. W Song & P. M. Panayides (Eds.), *Maritime logistics: Contemporary issues* (pp. 117–132). Bingley: Emerald.

Lam, J. S. L., & Yip, T. L. (2012). Impact of port disruption on supply chains: A petri net approach. *Lecture Notes in Computer Science, 7555,* 72–85.

Li, C., Qi, X., & Lee, C. -Y. (2014a). Disruption recovery for liner shipping. Working paper, The Hong Kong University of Science and Technology.

Li, C., Qi, X., & Song, D.P. (2014b). Real-time scheduling and rescheduling in liner shipping service with regular uncertainties and disruption events. Working paper, The Hong Kong University of Science and Technology.

Meng, Q., Wang, T., & Wang, S. (2012). Short-term liner ship fleet planning with container transshipment and uncertain container shipment demand. *European Journal of Operational Research, 223*(1), 96–105.

Meng, Q., Wang, S., Andersson, H., & Thun, K. (2014). Containership routing and scheduling in liner shipping: Overview and future research directions. *Transportation Science*. doi:10.1287/trsc.2013.0461.

Notteboom, T. E. (2006). The time factor in liner shipping services. *Maritime Economics & Logistics, 8,* 19–39.

Paul, J. A., & Maloni, M. J. (2010). Modeling the effects of port disasters. *Maritime Economics & Logistics, 12*(2), 127–146.

Psaraftis, H. N., & Kontovas, C. A. (2013). Speed models for energy-efficient maritime transportation: A taxonomy and survey. *Transportation Research Part C, 26,* 331–351.

Qi, X., & Song, D.P. (2012). Minimizing fuel emissions by optimizing vessel schedules in liner shipping with uncertain port times. *Transportation Research Part E, 48*(4), 863–880.

Thengvall, B., Bard, J. F., & Yu, G. (2000). Balancing user preferences for aircraft schedule recovery during irregular operations. *IIE Transactions, 32*(3), 181–193.

Tran, N.K., & Haasis, H. -D. (2013) Literature survey of network optimization in container liner shipping. *Flexible Services and Manufacturing Journal*. doi:10.1007/s10696-013-9179-2.

Vernimmen, B., Dullaert, W., & Engelen, S. (2007). Schedule unreliability in liner shipping: Origins and consequences for the hinterland supply chain. *Maritime Economics & Logistics, 9,* 193–213.

Wang, S., & Meng, Q. (2012a). Liner ship route schedule design with sea contingency time and port time uncertainty. *Transportation Research Part B, 46*(5), 615–633.

Wang, S., & Meng, Q. (2012b). Robust schedule design for liner shipping services. *Transportation Research Part E, 48*(6), 1093–1106.

Wang, Y., & Qi, X. (2014) Vessel Speeding Decisions Under The Chasing Of A Pirate Boat. Working paper, The Hong Kong University of Science and Technology.

Wang, S., Meng, Q. & Liu Z. (2013a). Bunker consumption optimization methods in shipping: A critical review and extensions. *Transportation Research Part E, 53,* 49–62.

Wang, T., Meng, Q., Wang, S., & Tan, Z. (2013b) Document risk management in liner ship fleet deployment: A joint chance constrained programming model. *Transportation Research Part E, 60,* 1–12.

Yu, G., & Qi, X. (2004). Disruption management: Framework, models and applications. New Jersey: World Scientific Publisher.

Yu, G., Arguello, M., Song, G., McCowan, S., & White, A. (2003). A new era for crew recovery at continental airlines. *Interfaces, 33,* 5–22.

Zhen, L. (2014). Storage allocation in transshipment hubs under uncertainties. *International Journal of Production Research, 52*(1), 72–88.

Zhen, L., Lee, L.H., & Chew, E.P. (2011). A decision model for berth allocation under uncertainty. *European Journal of Operational Research, 212*(1), 54–68.

Chapter 9
Bunker Purchasing in Liner Shipping

Christian E. M. Plum, David Pisinger and Peter N. Jensen

Abstract The cost for bunker fuel represents a major part of the daily running costs of liner shipping vessels. The vessels, sailing on a fixed roundtrip of ports, can lift bunker at these ports, but prices in each port may be differing and fluctuating. The stock of bunker on a vessel is subject to a number of operational constraints such as capacity limits, reserve requirements and sulphur content. Contracts are often used for bunker purchasing, ensuring supply and often giving a discounted price. A contract can supply any vessel in a period and port, and is thus a shared resource between vessels, which must be distributed optimally to reduce overall costs. An overview of formulations and solution methods is given, and computational results are reported for some representative models.

9.1 Introduction

Liner shipping companies are at the core of the major supply chains in the world, providing relatively cheap and reliable transport to and from any corner of the world. This industry has grown massively in the last decades, often with two digit percentage growth rates. Lately the supply of vessels have exceeded the demand for container transport, resulting in many liner carriers being loss giving. The profit margins in liner shipping are very slim, with marginal changes resulting in a company loss instead of profit.

C. E. M. Plum (✉) · D. Pisinger
DTU Management Engineering, Technical University of Denmark,
Produktionstorvet 424, 2800 Kgs. Lyngby, Denmark
e-mail: chrplum@gmail.com

C. E. M. Plum
Maersk Line, Esplanaden 50, 1098 København K., Denmark

D. Pisinger
e-mail: pisinger@man.dtu.dk

P. N. Jensen
Maersk Oil Trading, Esplanaden 50, 1098 København K., Denmark
e-mail: peter.neergaard.jensen@maersk.com

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_9

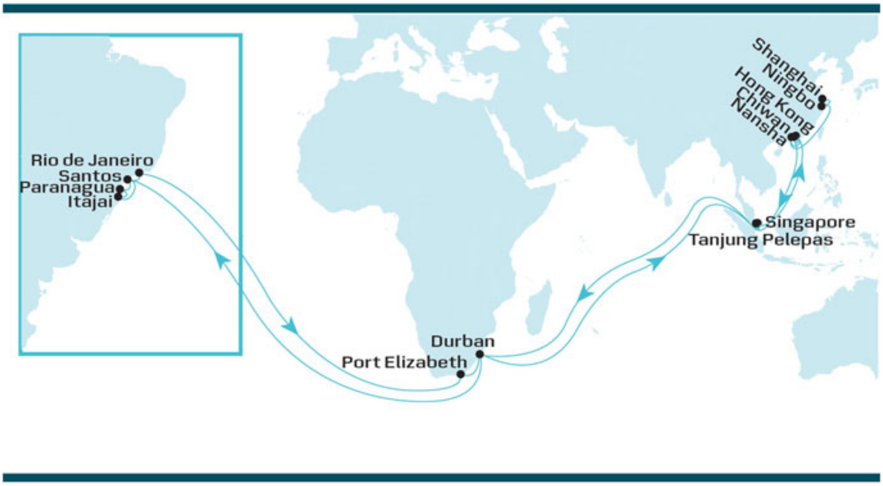


Fig. 9.1 The ASAS 2 Service, transporting containers between East Coast South America, South Africa and the Far East

This has shifted the shipping industry from a revenue optimizing focus, to use more resources on controlling and minimizing their costs. An example is the spend on bunker fuel, as this constitutes a very large part of the variable operating cost for the vessels. Also, the inventory holding costs of the bunker on board may constitute a significant expense to the liner shipping company.

For liner shipping companies in particular, the purchasing of bunkers can be planned some months ahead, as the vessels are sailing on a fixed schedule allowing for planning, as opposed to other types of shipping. An example of a liner shipping service can be seen in Fig. 9.1, where the vessels are sailing between East Coast South America, South Africa and the Far East. This service allows for bunkering in three distinct markets, making it attractive to plan with a long time horizon.

Bunker prices are fluctuating and generally correlated with the crude oil price, but there are significant price differences between ports of up to 100 \$/mt (of a $\approx$ 600 \$/mt price). The price differences between ports are not stable, and the cheapest port on a roundtrip today may not be the cheapest tomorrow. In Fig. 9.2 the prices for five important ports have been plotted for a time period of 18 months, illustrating how much the prices fluctuate. This creates the need for frequent (daily) reoptimization of the bunker plan for a vessel, to ensure the lowest bunker costs.

The bunker purchasing problem is to satisfy the vessels consumption by purchasing bunkers at the minimum overall cost, while considering reserve requirements, and other operational constraints. Bunker can be purchased on the spot market when arriving to a port, but normally it is purchased some weeks ahead of arrival. Long-term contracts between a liner shipping company and a port can result in reduced bunkering costs by committing the company to purchase a given amount of bunker. Bunkering contracts may cover several vessels sailing on different services, making the planning quite complex.

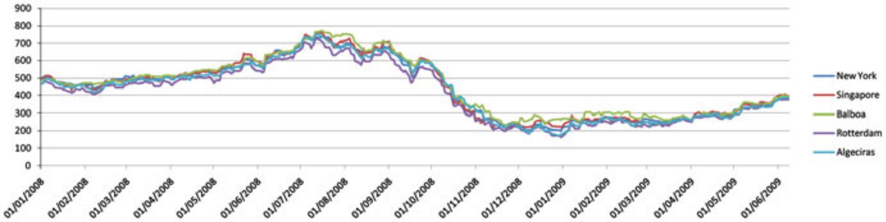


Fig. 9.2 Bunker prices in New York, Singapore, Balboa, Rotterdam and Algeciras plotted from January 2008 to June 2009. On January 1st 2009, Balboa was almost 30 % more expensive than Rotterdam and New York

An example of a bunkering plan can be seen in Table 9.1. *Departure stock* is the stock of bunker at departure of the port, as calculated by the model. *Consumption* is, the estimated consumption of bunker from this port to the next. *Purchase* is the quantity of bunker purchased at the port and *Spot Price* is the market price of bunker at the spot market. LSFO denotes low sulphur bunker, while HSFO denotes high sulphur bunker. Quantities are given in metric tonnes (mt). Possible bunker contracts are not shown. At the fourth port call 186 mt HSFO is bought at the spot market and 818 mt HSFO through a contract.

9.1.1 Literature

For a broad introduction to shipping and the importance of bunker spend refer to Stopford (2009) and for an introduction to operations research within the maritime industry Christiansen et al. (2004), Christiansen et al. (2007) and Christiansen et al. (2013) provide excellent overviews. A detailed description of Liner Shipping Network Design and the impact of bunker usage and other relevant factors appears in Brouer et al. (2013). The authors also introduce **LINER-LIB 2012**, a benchmark data suite, consisting of liner shipping relevant data and benchmarks specifically for liner shipping network design problems. The work of Plum et al. (2014) designs a liner shipping network taking bunker consumption into account. Details on the bunkering industry in relation to shipping can befound in Boutsikas (2004).

The effect of the bunker price on Liner Shipping Network Design has been studied in a number of recent papers, as Wang and Meng (2012) and Meng et al. (2012). The effect of bunker usage by the maritime industry in relation to the bunker price is investigated by Corbett et al. (2009) with the aim of reducing CO_2 emissions by imposing tax on bunkers. The work of Acosta et al. (2011) considers factors impacting the choice of bunker port. Fagerholt et al. (2009) considers the optimal speed and route for a ship with respect to bunker costs. Other work on bunker costs and its impact on maritime transportation includes Notteboom and Vernimmen (2009), who consider how slow steaming and the cost structure of liner shipping networks are affected by changes in bunker costs, and Ronen (2010), who considers the bunker

Table 9.1 An example of a bunker plan for a vessel operating a schedule with 29 ports

Port Id	Departure stock (mt)		Consumption (mt)		Purchase (mt)		Spot price ($/mt)	
	LSFO	HSFO	LSFO	HSFO	LSFO	HSFO	LSFO	HSFO
NLROT	462	648	5	0	0	0	481	455
DEBRV	457	648	100	0	0	0	500	490
GBFXS	357	648	97	648	0	0	1000	675
USNWK	260	1004	0	134	0	186 + *818*	491	465
USCHS	260	870	0	425	0	0	490	477
USSAV	260	445	0	74	0	0	493	457
USMIA	260	371	0	211	0	0	1000	1000
USHOU	260	1456	0	122	0	1296	1000	442
USMOB	260	1334	183	201	0	0	1000	1000
USNFK	77	1133	24	555	0	0	484	469
GBFXS	53	578	8	0	0	0	1000	641
NLROT	1053	4314	0	2340	1008	3737	442	421
DEBRV	1053	1974	0	447	0	0	457	447
USNWK	1053	1527	2	110	0	0	490	466
USCHS	1051	1417	0	25	0	0	495	482
USSAV	1051	1392	0	82	0	0	502	471
USMIA	1051	1310	0	211	0	0	1000	1000
USHOU	1051	1099	0	128	0	0	1000	451
USMOB	1051	971	0	365	0	0	1000	1000
USNFK	1051	606	221	606	0	0	510	495
NLROT	830	4021	21	0	0	4021	436	415
GBFXS	809	4021	65	0	0	0	1000	652
DEBRV	744	4021	98	511	0	0	467	456
USNWK	646	3510	0	161	0	0	485	464
USCHS	646	3349	0	19	0	0	496	483
USSAV	646	3330	0	98	0	0	499	468
USMIA	646	3232	0	183	0	0	1000	1000
USHOU	646	3049	0	135	0	0	1000	465
USMOB	646	2914	0	388	0	0	1000	1000

price's effect on speed and fleet size. The recent work of Wang et al. (2013) provides an overview of available bunker optimization methods in shipping.

Fuel Purchasing. The problem of reducing fuel costs by optimizing fuel purchase has been investigated in a number of papers, studying different transport modes. Vilhelmsen et al. (2014) investigated how tramp ships can be routed, while considering the impact on bunker costs. Oh and Karimi (2010) plan bunker purchases for multi parcel tankers considering a fixed route under uncertain prices. This problems resembles bunker purchasing for liner vessels except that the vessel must make route deviations for bunker purchasing, thus making the problem partly a route selection problem, giving a different problem structure.

Research investigating how to refuel a transportation fleet has also been done for other transportation areas as the airline industry (Stroup and Wollmer (1992), Abdelghany et al. (2005)), trucking industry (Suzuki 2008, 2011) and in more general (Lin 2008). These papers take offset in the specific operational reality of the transport mode and possibly generates routes for the transport vehicle at the same time. This gives somewhat different optimization problems, not directly applicable to liner shipping bunker purchasing problems.

Bunker Purchasing in liner shipping. For a vessel sailing on a given port to port voyage at a given speed, the bunker consumption can be fairly accurately predicted. This gives an advantage in bunker purchasing, when a vessel has a stable schedule known for some months ahead. The regularity in the vessel schedules in liner shipping allows for detailed planning of a specific vessel, as considered in the works of Plum and Jensen (2007), Besbes and Savin (2009), Kim et al. (2012), Kim (2014), Sheng et al. (2014) and Yao et al. (2012). These papers consider variants of a bunker optimization problem considering a single vessel.

Besbes and Savin (2009) consider different refueling policies for liner vessels and presents some interesting considerations on the modeling of stochastic bunker prices using Markov processes. This is used to show that the bunkering problem in liner shipping can be seen as a stochastic capacitated inventory management problem. Capacity is the only considered operational constraint.

The work of Plum and Jensen (2007) considers multiple tanks in the vessel and stochasticity of both prices and consumption, as well as a range of operational constraints. Yao et al. (2012) does not consider stochastic elements nor tanks, but has vessel speed as an variable of the model. The work of Kim et al. (2012) minimizes bunker costs as well as startup costs and inventory costs for a single liner shipping vessel. This is done by choosing bunker ports and bunker volumes but also having vessel roundtrip speed (and thus the number of vessels on the service) as an variable of the model; Kim (2014) presents a different algorithm for a similar problem scoping.

In Sheng et al. (2014) a model is developed which considers the uncertainty of bunker prices and bunker consumption, modelling their uncertainty by markov processes in a scenario tree. The work can be seen as an extension of Yao et al. (2012), as it considers vessel speed as a variable within the same time window bounds. Capacity and fixed bunkering costs is considered, as is the holding / tied capital cost of the bunkers. Improved solutions are found, as compared to Yao et al. (2012), credited to the uncertainty modelling. It only plans bunker purchases for a single roundtrip.

The studies described above do not consider bunker contracts, and all model the bunker purchasing for a single vessel. The work of Farina (2012) is an extension of Plum and Jensen (2007) with the additional consideration of bunker contracts, where a MIP model is presented capable of solving a 50 vessel instance for a 6 month period, falling short of solving real world instances of hundreds of vessels.

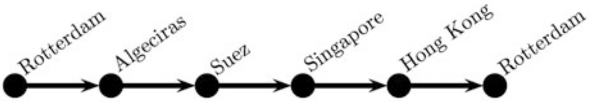


Fig. 9.3 A simple example of a *service* starting and ending in Rotterdam

Plum et al. (2014) presents a decomposition algorithm for bunker purchasing with contracts, and showed that the model is able to solve even very large real-life instances involving more than 500 vessels, 40,000 port calls, and 750 contracts.

In the following section we will define the basics of bunker purchasing in liner shipping, and discuss all relevant constraints. The next section presents a basic model for bunker purchasing of a single vessel, using spot prices, and discuss extensions and variations presented in the literature. Finally, the model is extended to handle several vessels having shared contracts for buying bunker at fixed prices. This problem is more complex, and a decomposition model is needed to solve the problem to optimality for large instances. Finally, the chapter is concluded with a summary of the most important results, and directions for future research.

9.2 The Bunkering Problem

In this section we define the bunkering problem more formally, and introduce all relevant terms and constraints: We first introduce different bunker *grades*, and then describe how the *prices* for bunker at given dates and ports are obtained. Next, we describe the rules dictating how bunker is requested, *ordered* and *delivered*. We then describe how the vessels' bunker *consumption* is calculated, the *tanks* used for storage, and bunker *reserve*. Finally we describe the testing of bunker and the *quarantine* periods this invokes, and various constraints for *mixing* different bunkers.

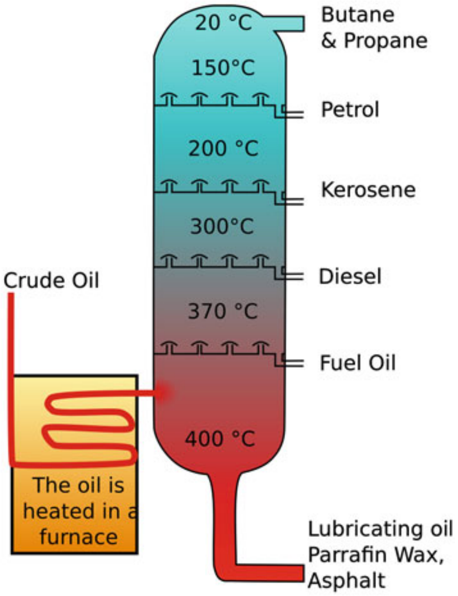
A vessel sails along a *path* of ports during a period of time. A *path* of length n can be divided into $n-1$ *legs*. A *service* is a *path* starting and ending in the same port, a *round trip*. An example can be seen in Fig. 9.3. For each port visit the vessel arrives and departs at specific dates and times, these are estimated time of arrival (ETA) and estimated time of departure (ETD).

All paths and the corresponding ETA and ETD are assumed to be fixed and can thus not be changed by a bunker purchasing model. The services of the vessels are constructed by shipping company considering other factors such as market demand for container transportation, etc.

It follows from this that interaction between vessels is impossible. Each vessel follows its own path and visits the ports at the specified times. Since no interaction is possible between the vessels we do only need toconsider exactly *one* vessel in the model.

We have a given set of bunkers available (described in Sect. 9.2.1) in each port, and since we know exactly when the vessel visited each port, we can apply a fixed price for each bunker, for each port. This implies that we do not need to include time in the model, since we for each port visited have a fixed mapping from a bunker to a price.

Fig. 9.4 Illustration of the distillates from crude oil. (source: Wikimedia)



9.2.1 *Bunker Grades*

Crude oil is refined into a variety of different products, ranging from jet fuel to gasoline, to bunker, to asphalt, see Fig. 9.4. Bunker fuel oil is one of the heavier distillates from the distillation process.

Bunker fuel is sold in different grades, mainly distinguished by their viscosity, but also characteristics such as density, sulphur content and others are relevant. Some characteristics of the main bunker grades can be seen in Table 9.2. We mention RMF 180, RMG 380 etc. simply as 180, 380 and 700. For the case of RMG 380 15 we say 380 low sulphur.

Table 9.2 Characteristics of typical bunker grades. The RMG 380 15 is a low sulphur bunker and the rest are high sulphur bunkers

Parameter	Unit	Limit	RMF 180	RMG 380	RMG 380 15	RMK 700
Density at 15 ° C	kg/m^3	Max	991.0	991.0	1010.0	1010.0
Viscosity at 50 ° C	mm^2/s	Max	180.0	380.0	380.0	700.0
Water	% V/V	Max	0.5	0.5	0.5	0.5
Sulphur	% (m/m)	Max	4.5	4.5	1.5	4.5
Alum. + Silicon	mg/kg	Max	80	80	80	80
Flash point	° C	Min	60	60	60	60
Pour point, Summer	° C	Max	30	30	30	30
Pour point, Winter	° C	Max	30	30	30	30

Two of the characteristics are relevant for the optimization problem: the viscosity of the bunker and its sulphur content. The rest of the factors do not affect the modelling or the solution.

We will assume that all bunker types show approximately the same fuel efficiency. If this is not the case, it is easy to modify the proposed models to take into account the efficiency. Since vessels generally only use one viscosity of bunker, the bunker consumption per nautical mile is easily adjusted to the designed bunker type.

Viscosity

Because of the high viscosity of bunker fuel oil, it is heated to 60–80 °C prior to injection into the engine. The vessel's engine is designed to handle bunker of a certain viscosity, thus an engine can always burn more viscous bunker than the minimum grade it can take, but not less viscous bunker. E. g. a vessel engine which can burn a 380 bunker as the lowest quality grade, can also burn any 180 bunker, but not a 700 bunker.

The prices of the bunkers are directly proportional with their viscosity, so a shipping company will always purchase the bunker with the lowest quality grade/highest viscosity available, which can be burned in the vessel's engine. Moreover, since bunker above the engine's viscosity limit cannot be burned, these bunkering options can be removed in a preprocessing phase.

Sulphur Contents

Apart from categorizing a bunker by its density, viscosity, etc. it can be categorized by its sulphur content. Bunkers containing sulphur below a given limit are classified as *low sulphur* bunker. In certain parts of the world, known as SECA-areas (SO_x Emission Control Area), vessels are only allowed to burn low sulphur bunkers to limit air pollution. The Baltic Sea has been a SECA-area since May 19th 2006 by the MARPOL Annex VI protocol The North Sea became a SECA-area by November 22nd 2007. Also an area of 24 miles of California coast is a SECA area. The sulphur limits for fuel in SECA are 1.0 % until January 2015, and 0.1 % after January 2015. The general sulphur limits in other sea areas are 3.5 % until January 2020, and 0.5 % after January 2020. The future dates may be postponed if political agreement cannot be reached.

9.2.2 Bunker Prices

In practice there are two types of orders done by bunker traders: spot and contract orders. Spot orders are handled by a trader requesting bunker prices from one or more suppliers of the day and then places an order based on the price quotes. Contract orders are done on the basis of a contract, where the shipping company is obliged to purchase a certain volume of bunker within a certain period of time.

The market for bunker trading is commoditized and liquid, the use of contracts for a specified amount, port and price (or discount to some price-index) is widespread.

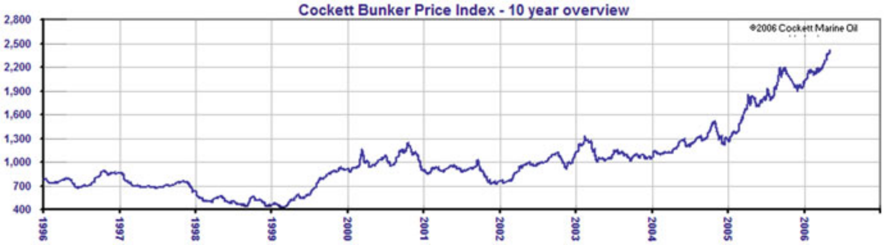


Fig. 9.5 Price development for bunker oil over the last 10 years. The index was started at 1000 on January 1st 1986, and represents the global price movement of bunkers

Table 9.3 Price quotes per metric tonne for 5 ports on April 24th 2006

Port	Rotterdam		Algeciras		Suez		Singapore		Hong Kong	
	Price	Price	Price	Price	Price	Price	Price	Price	Price	Price
Grade	Low	High	Low	High	Low	High	Low	High	Low	High
380	327	330	339	349	345	347	355	356	368	370
180	350	352	364	365	357	359	365	366	381	382
Marine diesel	590	595	672	677	647	650	645	650	645	650

This is done to reduce both delivery and price risk and to leverage the strength of being a large player on this market.

Liner shipping companies engage in contracts for the purchase of bunkers at ports where they have a large and regular demand. This is done both to gain a discount compared to the spot market, by leveraging on the large volumes involved, and to increase supply certainty. Bunker contracts will usually concern total lifted volumes within a calendar month, with specified minimum and maximal quantities.

The price can be agreed on in different manners, usually by using a fixed discount below the monthly average of a bunker index (Bunkerwire (2013)) of the port in question. A contract is for one or more bunker grades and one or more ports, which usually will be located geographically close and considered as the same market. Many contracts can be available in a port for a bunkering vessel, and it must then be chosen which, if any, to purchase bunker from. *Spot* bunker is assumed freely available at all ports with a given *price quotes*. For an example of the bunker price development see Fig. 9.5.

Ordering Time Window

An order for bunker must be placed before the vessel arrives at port. In order to prepare the suppliers and ensure that the bunker is available when the vessel enters the port, the order must be placed at some time interval before arrival (generally 72 h). The order can be placed up to 2 weeks before the vessel arrives at the port, but normally happens 3–7 days before.

Price Structure

The objective of all bunker purchasing models is to minimize the total cost of bunker purchases. When calculating the bunker purchase costs, a number of factors should be taken into account:

Barge, Startup Costs Generally the bunker pricing scheme differs between ports in America and ports in the rest of the world. European and other ports will price a fixed cost times the amount of bunker purchased. US ports will add a *barging* cost, for each of the barging vessels used to load bunker onto the vessel. (usually a vessel gets the bunker from barges holding around 2000 t of bunker, which sails up to the vessel). This could be modelled using a piecewise linear objective function.

Different Price Over/Under 500 t Some bunker retailers operate with different per ton price for purchases over and under 500 t of bunker. This is a different version of the non-linear pricing structure mentioned above, but less used. This structure can also be modeled with a piecewise linear objective function.

Cost of Capital Tie Up The amount of bunker in a fully loaded vessel can represent a value of several million dollars. This value is tied-up in the bunker and cannot be used to generate interest nor be invested elsewhere. This represents a loss of profit which can be modelled by including the capital tie up term in the objective function.

Cost of Bunker Carriage The heavier a vessel is, the more bunker it will consume.

Bunker Tests The price of a bunker is related to its specifications as described in Sect. 9.2.1. This gives the vendors of bunker an incentive to dilute their products with heavier components until it is just barely fulfilling the requirements of the bunker's ISO specifications. Hence it sometimes happens that the vendors sell products that does not satisfy the specifications they claim. These bunkers are *off-spec*. The consequences of burning off-spec bunker can be grave if the bunker already was the lowest possible burnable in the engine. The engine can malfunction and leave the vessel immobilized.

In many major ports it is possible to purchase bunker already tested by independent laboratories, so called *pre-tested* bunker. When this is available it can be trusted that the bunker is within the specification limits as claimed by the vendor, and the vessel can hence use the bunker immediately after purchasing.

For any purchased bunker a *post-test* is also carried out at an appointed analysis institute. This is also the case even for bunkers for which pre-test is available. Usually this process will take at most 5 days. The purchased bunker cannot be used in this time period.

Purchasing Limits

Other constraints apply to the bunkering in a port. A vessel will usually stay in a port between 12 and 24 h from arrival to departure. The bunker can be loaded onto the vessel in two different ways. Larger ports have piping facilities in the actual docks, which can be connected to the vessel directly, and through these pipes large volumes of bunker can quite quickly be loaded onto the vessel. But this only exists in larger ports and only at certain anchorage places. When the vessel is lying at other ports or anchorage places, the bunkering is done with barges loading bunker from large land-side tanks, sailing to the vessel and then loading onto the vessel. The barges will usually take up to 2000 mt. of bunker each.

At busy times the maneuverability of the barges can be hindered by traffic from other vessels and port congestion can prevent the barges from sailing to the vessels, and thus prevent bunker from being loaded onto the vessel. Other minor ports might not have enough barges available. Together, these factors may impose a limit on how much bunker can be loaded onto the vessel. Therefore, a model should be able to handle an upper limit on the bunkering capacity, although most ports in reality do not have an upper limit.

A proper model should also enforce that some ports have a minimum amount of bunker that can be purchased, since vendors will refuse selling too small quantities of bunker. Typically this lower limit is 200 mt.

9.2.3 Bunker Consumption

The consumption of the vessel can (roughly) be regarded as a function of bunker type, weight of bunker, weather conditions, ocean currents, container load and speed of the vessel. Typically, the bunker type, speed of vessel and approximate container load is known some time in advance, while weather conditions and currents can change the bunker consumption significantly. A deterministic model for bunker purchasing must use average historic values to estimate bunker purchasing. The bunker consumption can also be considered as a stochastic variable in a model, leading to a more correct formulation that, however, is more difficult to solve.

9.2.4 Bunker Tanks

The vessels have multiple tanks. Some of these, typically four to sixteen tanks, are larger tanks, holding up to some thousands of metric tonnes. A number of smaller special tanks in close connection to the vessel's engine exist as well, among these the day-tank which is directly connected to the engine, and holds enough bunker for 24 h usage.

Some tanks are placed symmetrically at each side of the vessel. In order to keep the vessel's balance stable during travel, consumption must take place more or less symmetrically.

Certain tanks can be dedicated for some types of bunkers. Specifically since we are purchasing more expensive low sulphur bunkers to use when the vessel is sailing in the SECA area, we need to ensure that this bunker is not contaminated with bunker with a higher sulphur content. Therefore vessels sailing in the SECA area will have at least one tank dedicated to low sulphur bunkers. We can still use low sulphur bunker outside the SECA area.

Tank Limits

The volume of bunker grows with increasing temperature, so in order to avoid overflowing tanks, the tanks may only be filled to around 98 % of their maximal capacity (at 25 °C). This can easily be handled in a preprocessing phase, where the volume of the tanks is decreased to compensate for the fill limit.

The suction systems of the tanks are relying on the weight of the bunker in the tanks to extract the bunker to the engine. This means that under normal conditions the tanks cannot be emptied completely.

It is possible to move bunker between tanks, and this could be an advantage for an optimization model in some cases in order to avoid quarantine or mixing situations.

Bunker Reserves

When the vessel is sailing on the ocean it is paramount for any solution that the vessel can overcome any eventualities on its own. The cost of recovering a stranded vessel far exceeds any cost to the actual bunker. Therefore a solution needs to ensure that the vessel carries adequate reserves of tested bunker at all times.

Shipping companies have a number of requirements to bunker reserves which a valid fueling strategy must satisfy. Generally a vessel must hold 8 days of bunker in reserve at all times. This requirement can be lowered to 4 days reserve, when the next port on the vessel's service offers pre-tested bunker, see Sect. 9.2.2.

The actual amounts needed to be kept in reserve should be calculated for each port in the preprocessing face, based on the estimated consumptions per day of the vessel.

9.2.5 Bunker Quarantine Time

It is needed to ensure, that the bunker is not off-spec before it is used as fuel. As described in Sect. 9.2.2 this is done by possibly pre-testing and post-testing, which can take up to 5 days.

Meanwhile we cannot use the bunker, and it is therefore quarantined for 5 days. If we have mixed the bunker it is also quarantined, see the next section.

We only detail our model to how much bunker we use at which leg of the vessel's service, hence the 5 days quarantine will be translated to how many legs on the service the bunker must be in quarantine before we can use it. This can be calculated effectively for each port in preprocessing.

9.2.6 Mixing of Bunkers

As mentioned in Sect. 9.2.4 each vessel is equipped with a number of tanks in different sizes. It is desirable to store the different bunker types separately, i.e. in different tanks in order to avoid the mixing of different bunker types. Two bunkers are considered as different types if at least one of the following properties holds:

- Differing fuel grades (density, viscosity, sulphur content, etc.)
- Same fuel grade bought at different ports
- Same fuel grade bought in the same port at different times
- Same fuel grade bought in the same port at different bunker vendors.

If two incompatible bunker types are mixed it will cause the creation of sediments, in this context asphalt, which when injected into the engine of the vessel can cause its malfunction. In this case the warranty from the manufacturer of the engine does not cover and shipping company will have to pay for a new engine plus they will loose money from lost earnings.

The risk of producing a bad mix is proportional to the density and viscosity of the bunkers and also depends on the ratio of the mixed bunkers. E. g. there is low risk of producing a bad mix when blending bunkers with low density, low viscosity and one bunker being added in a higher proportion than the other.

When a tank with a bunker volume less than 10 % of the tank capacity subsequently is filled with at least nine times as much bunker of a different type it is not considered a mixing of bunkers. The tank is then considered as containing the bunker of the highest volume. We call this a 9:1 mix.

It is only permissible to mix two bunkers at a time, so mixes with three or more constituents should not be considered. This also implies that two mixed bunkers, cannot be mixed again.

The amount of tanks where mixing takes place should be minimized. This takes precedence over minimizing the ratio of mixture. A scenario could be considered where the ratio of mixture is lowered by spreading the original bunker over a number of tanks, and mixing in each of these. This strategy should not be encouraged by an optimization model.

9.3 A Model for Single-Vessel Bunker Purchasing

We will now introduce a basic model for bunker purchasing considering only one vessel. This means that bunker is bought on the spot market, or through contracts only covering the studied vessel.

We use a deterministic model, based on the formulations from Besbes and Savin (2009) and Plum and Jensen (2007), hence assuming that the bunker consumption for every leg is known in advance. We will also assume that a sufficient amount of tanks are present for each vessel, so that the mixing constraints discussed in Sect. 9.2.6 does not need to be taken into account. This is the case for most large vessels. The basic model is given without special cost structures such as piece-wise linear bunker costs and capital start up cost. Also some operational constraints are omitted, such as quarantine requirements, several bunker tanks and possibility of mixing bunkers. These cost structures and operational constraints are described in Sect. 9.4 as extensions to this basic bunker purchasing model.

We introduce the mathematical notation used throughout this chapter. Let $v \in V$ be the set of vessels. Let $i \in I$ be an ordered set of *port calls*, the vessel's schedule. A port call i will be uniquely defined by a port, a vessel, $v(i)$ and the date of arrival. Let $init(v)$ and $term(v)$ be the first and last considered port call of vessel v. Let $b \in B = \{L, H\}$ be the two considered bunker types. The startup cost for bunkering at a port call i is $startcost_i$. Each vessel, v has a capacity $D_{v,b}$ for each bunker type, b. For each leg i of the schedule, the vessel consumes $F_{i,b}$ bunker, between port call i and $i + 1$.

Variables used in the model are as follows: $l_{i,b}$ is the purchase of bunker for each port i, bunker type b. The binary variable $\delta_{i,b}$ is set to one if a purchase of a bunker type b is made at a port call i. The volume $h_{i,b}$ of bunker after a vessel leaves port is a continuous variable, as is the consumption $f_{i,b}$ of each bunker type on vessel between port i and $i + 1$.

9.3.1 Model

A basic model for Bunker Purchasing for Liner Shipping can be formulated as the following Mixed Integer Program:

$$\min \sum_{i \in I} \sum_{b \in B} (\delta_{i,b} \cdot startcost_i) + \sum_{i \in I} \sum_{b \in B} (p_{i,b} \cdot l_{i,b})$$

Subject to

$$h_{i,b} = h_{i-1,b} + \sum_{i \in I} \sum_{b \in B} l_{i,b} - f_{i-1,b} \ \forall i, b \tag{9.1}$$

$$f_{i,b} \leq h_{i,b} \ \forall i, b \tag{9.2}$$

$$\sum_{b \in B} f_{i,b} = F_{i,H} + F_{i,L} \ \forall i \tag{9.3}$$

$$f_{i,L} \geq F_{i,L} \ \forall i \tag{9.4}$$

$$h_{i,b} \leq D_{v(i),b} \ \forall i, b \tag{9.5}$$

$$l_{i,b} \leq \delta_{i,b} \cdot D_{v(i),b} \ \forall i, b \tag{9.6}$$

The objective minimizes startup costs and bunker cost. The constraint (9.1) ensures flow conservation at each port, vessel and bunker type. Constraint (9.2) ensures that no more bunker than available is used between port i and $i + 1$. Constraints (9.3) and (9.4) maintains the consumption of bunker, allowing LSFO to substitute HSFO, but not the other way around. The bunker capacity of the vessels are enforced by constraint (9.5). The decision variables $\delta_{i,b}$, indicating if any bunker is purchased by vessel b at port i, are set by constraint (9.6).

Finally, initialization and termination criteria for start and end bunker volumes must also be set. Let $S_{v,b}$ and $T_{v,b}$ be the start and terminal volume of bunker b on vessel v. This leads to the following constraints:

$$h_{init(v),b} = S_{v,b} \quad \forall v, b \tag{9.7}$$

$$\sum_{b \in B} h_{term(v),b} \geq \sum_{b \in B} T_{v,b} \quad \forall v \tag{9.8}$$

$$h_{term(v),L} \geq T_{v,L} \quad \forall v \tag{9.9}$$

Constraint (9.7) sets the start volume for both bunker types. Constraint (9.9) is the terminal volume for low sulphur bunker, while constraint (9.8) is the terminal volume for high sulphur bunker, allowing substitution.

The domains of the variables are:

$$h_{i,b}, l_{i,b}, f_{i,b} \in \mathbb{R}^+ \quad \forall i, b \tag{9.10}$$

$$\delta_{i,b} \in \{0, 1\} \quad \forall i, b \tag{9.11}$$

A usual time horizon for a single vessel model is 3–6 months, where a vessel may call up to a 100 ports. The MIP resulting from such problem instances can be solved by state of the art commercial MIP solvers, usually in a matter of seconds.

9.4 Operational Constraints

In practice bunker purchasing in liner shipping is influenced by a wide range of operational, commercial and financial factors, which dictates the properties of a good bunker plan. Some of these factors are modeled as MIP constraints in the following. Refer to the earlier mentioned literature for an elaborate discussion of other factors.

Bunker Reserves and Startup Costs As the consumption of bunker on a leg is an uncertain parameter due to factors as changed schedule (and thus speed), wind, current, waves and hull roughness, a good bunker plan will allow for variation in the bunker consumption. A way to handle this is to enforce a minimum reserve requirement of bunker at port arrival. This can be modeled as in (9.12), where $F_{i,b}$ is the minimal reserve requirement at port arrival.

Besides the startup cost for bunkering, $startcost_i$, bunker suppliers will usually require a minimum quantity to be purchased at each bunkering, this can be handled with constraints (9.13), where $\underline{L}_{i,b}$ is the minimal quantity.

$$F_{i,b} \leq \sum_{b \in B} (h_{i,b} - l_{i,b}) \quad \forall i, b \tag{9.12}$$

$$\delta_{i,b} \cdot \underline{L}_{i,b} \leq l_{i,b} \quad \forall i, b \tag{9.13}$$

Capital and Carriage Cost The capital costs of bunker is extensive, due to the large volumes and high prices. A model could consider this by adding this cost (or lacking interest) to the objective, proportional to the average load of bunker on the vessels. Assuming an interest rate of α and a bunker cost of C \$/mt, we can estimate a daily capital cost per mt of bunker as:

$$(\sqrt[365]{1+\alpha} - 1) \cdot C \ \$/mt \tag{9.14}$$

A typical value for the interest rate is $\alpha = 10\%$ and a typical value of bunker cost is 600 \$/mt, leading to the daily capital cost per mt of bunker:

$$(\sqrt[365]{1.1} - 1) \cdot 600 \ \$/mt = 0.157 \ \$/(day \cdot mt) \tag{9.15}$$

Let d_i denote the number of days a vessel uses to travel from port i to port $i+1$. Then the cost of capital tie up can be included in the objective function as:

$$0.157 \cdot d_i \Big(\sum_{b \in B} h_{i,b} - \frac{F_{i,b}}{2} \Big) \tag{9.16}$$

Similarly a vessel carrying a large volume of bunker will, all things equal, have a larger draft. This will in general (but not always, due to specifics in the vessels design as the bulb) imply an increased bunker consumption proportional to an increased load. This term could be considered in the objective in the same manner as the capital costs.

Alternatively this can be modeled directly as an increased consumption proportional to increased bunker load. If we assume an increased consumption of $\gamma \cdot mt_{extra}/(day \cdot mt_{carried})$, the corrected bunker consumption per leg, $F^{corr}_{i,b}$ becomes:

$$F^{corr}_{i,b} = F_{i,b} + \gamma \ d_i \Big(\sum_{b \in B} h_{i,b} - \frac{F_{i,b}}{2} \Big) \tag{9.17}$$

We can then replace the term $F_{i,b}$ with $F^{corr}_{i,b}$ to model the increased cost of carrying bunker.

California Sales Tax The California bunker sales tax, as described by California Legislative Analyst's Office (2001), imposes a tax on bunker bought in California, which necessarily must be used en-journey to the first out of state port. I.e. if a vessel arrives with 1000 mt at an Californian port and requires 2000 mt to reach the first non-Californian port on its schedule, it must pay a tax for the first 1000 mt purchased. With additional decision variables this can be modelled and included in the objective.

Quarantine A sample is usually taken from purchased bunker, to be analyzed for its specific content of carbohydrates, sulphur, water, ashes, etc. The sample must be within the ISO specifications of the purchased bunker grade. Until the result of the laboratory test are received, the bunker may not be used. This test can take three to five days. This can be handled by increasing the reserve requirements at port calls with bunker purchased within the last five days, $i' \in Quar(i)$:

$$F_{i,b} \leq \sum_{b \in B} \left(h_{i,b} - l_{i,b} - \sum_{i' \in Quar(i)} l_{i',b} \right) \quad \forall i, b \tag{9.18}$$

Other Constraints Other more detailed operational constraints can be handled in preprocessing of data for the model or by adding new constraints. This includes:

- Vessels that cannot bunker at a port due to wielding works imposing fire hazards. In such ports the maxlift limit can be set to: $\overline{L}_{i,b} = 0$.
- Ports with limited quantity k_1 of bunkers available can be handled by setting the maxlift limit to: $\overline{L}_{i,b} = k_1$.
- Vessels with stability or air draft requirements requiring high drafts. This can be imposed by forcing k_2 mt of bunkers at arrival by increasing the reserve requirements $F_i = k_2$.
- Tank fill limits can be handled by lowering the tank limits as follows: $D^*_{v,b} = 0.98 \cdot D_{v,b}$.
- A maximal number N_{Bunker} of bunkerings can be enforced by the constraint: $\sum_{i \in I} \sum_{b \in B} \delta_{i,b} \leq N_{Bunker}$.

The above constraints can be added to the model without increasing its complexity significantly. All of them can be formulated linearly and only relate to a single vessel at a time, allowing them to be considered in a vessel specific subproblem.

9.4.1 Complexity

The problem (9.1) to (9.6) is NP-hard to solve, which can be seen by reduction from the knapsack problem. The knapsack problem in minimization form as described in Kellerer et al. (2004) can be formulated as follows: Given a set N of items having profit p_i and weight w_i and a knapsack of capacity c, the problem is to fill the knapsack at minimum overall profit, such that the overall weight is at least c. Given an instance of the knapsack problem, we construct an instance of the bunker purchasing problem by having one vessel, visiting N ports. The fuel consumption between each pair of ports is 0, except the leg after the last port visit, where the consumption is c. In each port, we have a contract of maximum w_i, and the minimum limit for lifting bunker is also w_i. The cost of buying the quantity w_i is p_i. It is easily seen that solving the bunker purchasing problem also solves the knapsack problem.

9.5 Bunker Purchasing with Contracts

The single-vessel model presented in the previous section does not take into account volume refueling discounts, which can only be fully exploited using more than one vessel. Thus we will now present a multi-vessel optimization model, taking into account various contracts covering multiple vessels.

The model considered in this chapter uses a *crystal ball* approach, i.e. using data not known at decision time, to benchmark the quality of already executed decisions. As the actual price of the contract is not known before a month has passed, the model will use after-the-fact prices for calculations.

Contract Bunker Contract bunker must be purchased according to details given by a number of contracts $c \in C$, minimal and maximal quantities are given by $\underline{q_c}$ and $\overline{q}_c$. The specified quantities are soft constraints, which can be violated by paying a high cost, $\underline{w}$, for violating the minimum volume and a lower cost for breaking the maximal constraint, $\overline{w}$. Contract c may cover several ports and multiple vessels can call at these ports in the duration of the contract. Each contract will give rise to a number of purchase options, $m \in M$, i.e. discrete events where a specific port call i, and thus vessel v, calls within a time period, allowing it to purchase bunker from a contract c. Purchases on a purchase option m will be done at a price p_m, specified by the contract c. To simplify modelling and to increase the density of the derived model, the sets of port calls, $i \in I$ and purchase options, $m \in M$ will be used instead of their underlying sets: ports, vessels and contracts, which could give an equivalent but much larger model.

The possibility of purchasing on the spot market, is considered as a special type of contract. The minimal and maximal volumes are relaxed as $\underline{q_c} = 0$ and $\overline{q}_c = \infty$. All port calls i have two spot purchase options m for LSFO and HSFO, with prices set at the corresponding spot price of the day and port. For ports where bunker prices are not published, we assume a high cost.

The model makes use of the following variables: l_m is the purchase of bunker for each purchase option m. The binary variable $\delta_{i,b}$ is set iff a purchase of a bunker type b is made at a port call i. The volume $h_{i,b}$ of bunker b after a vessel leaves port i is a continuous variable, as is the consumption $f_{i,b}$ of each bunker type b between port i and $i + 1$. The contract violation or slack variables are $\underline{s_c}$ and $\overline{s}_c$. We let $M(c)$ denote the set of purchase options specified by contract c, and $M(i, b)$ the set of purchase options for bunker b in port i.

9.5.1 Model

The Bunker Purchasing with Contracts Problem (BPCP) can be formulated as the following Mixed Integer Program:

$$\min \sum_{i \in I} \sum_{b \in B} (\delta_{i,b} \cdot startcost_i) + \sum_{m \in M} (p_m \cdot l_m) + \sum_{c \in C} (\underline{s_c} \cdot \underline{w} + \overline{s}_c \cdot \overline{w})$$

subject to

$$h_{i,b} = h_{i-1,b} + \sum_{m \in M(i,b)} l_m - f_{i-1,b} \quad \forall i, b \tag{9.19}$$

$$f_{i,b} \leq h_{i,b} \quad \forall i, b \tag{9.20}$$

$$\sum_{b \in B} f_{i,b} = F_{i,H} + F_{i,L} \quad \forall i \tag{9.21}$$

$$f_{i,L} \geq F_{i,L} \quad \forall i \tag{9.22}$$

$$h_{i,b} \leq D_{v(i),b} \quad \forall i, b \tag{9.23}$$

$$\underline{q}_c - \underline{s}_c \leq \sum_{m \in M(c)} l_m \leq \overline{q}_c + \overline{s}_c \quad \forall c \tag{9.24}$$

$$\sum_{m \in M(i,b)} l_m \leq \delta_{i,b} \cdot D_{v(i),b} \quad \forall i, b \tag{9.25}$$

The objective minimizes startup costs, bunker cost and contract violation penalties. Constraint (9.19) ensures flow conservation at each port, for the given vessel and bunker type. Constraint (9.20) ensures that between port i and $i + 1$ bunker can only be used if available. Constraints (9.21) and (9.22) maintains the consumption of bunker, allowing LSFO to substitute HSFO, but not opposite. The bunker capacity of the vessels are enforced by constraint (9.23). The minimal and maximal quantity required by the contracts are ensured by the double sided constraints (9.24), allowing for violations. The decision variables $\delta_{i,b}$ are set by constraint (9.25).

Initialization and termination criteria for start and end bunker volumes must also be set:

$$h_{init(v),b} = S_{v,b} \quad \forall v, b \tag{9.26}$$

$$\sum_{b \in B} h_{term(v),b} \geq \sum_{b \in B} T_{v,b} \quad \forall v \tag{9.27}$$

$$h_{term(v),L} \geq T_{v,L} \quad \forall v \tag{9.28}$$

The first constraint defines the start volume, while the last two constraints define the terminal volume, allowing low sulphur bunker to substitute high sulphur bunker.

Finally the domain of the variables is given as follows:

$$h_{i,b}, l_m, f_{i,b}, \underline{s}_c, \overline{s}_c \in \mathbb{R}^+ \quad \forall i, b, m, c \tag{9.29}$$

$$\delta_{i,b} \in \{0, 1\} \quad \forall i, b \tag{9.30}$$

9.5.2 Bunker Contracts - Operational Constraints

Contracts may have minimum and maximal volumes that must be lifted per purchase, $\underline{N}_{i,b,c}$ and $\overline{N}_{i,b,c}$. This can be modelled similarly to the minimum lift constraints. As can purchases at port calls have maximal lift restrictions, $\overline{L}_{i,b}$, due to short port stays or limited supply:

$$\delta_{i,b} \cdot \underline{N}_{i,b,c} \leq \sum_{m \in M(i,b,c)} l_m \quad \forall i, b, c \tag{9.31}$$

$$\sum_{m \in M(i,b,c)} l_m \leq \delta_{i,b} \cdot \overline{N}_{i,b,c} \quad \forall i, b, c \tag{9.32}$$

$$\sum_{m \in M(i,b)} l_m \leq \delta_{i,b} \cdot \overline{L}_{i,b} \quad \forall i, b \tag{9.33}$$

9.6 Advanced Model Extensions

In order to get a closer correspondence with reality, the model can be extended to handle uncertainty in bunker consumption and prices. Moreover modelling of multiple bunker tanks and the properties of mixing different batches of bunker can add precision to the model. However, not all shipping companies have sufficient data quality to justify these extensions. Vessel speed can also be considered a variable which can be adjusted by the model.

9.6.1 Uncertainty

The previously described models assumes deterministic problem instances. Like in all models of real world problems this is an approximation of the full problem. In particular the uncertainty applies for the consumption of bunker and the price of bunkers, neither of which are deterministic.

Consumption Uncertainty The bunker consumption for a given vessel, distance, speed and displacement can be fairly accurately predicted. Still the bunker consumption is affected by uncertainty due to factors as: unforseen changes in distance (changed schedule), speed (earlier / later arrival), displacement (more or less cargo), uneven speed (lowest consumption is attained at an even speed throughout), weather and many other factors.

The model can be translated into a multi stage stochastic program, working on a generated scenario tree taking offset in the consumption estimates. For more details on this please refer to Plum and Jensen (2007) or Sheng et al. (2014).

Price Uncertainty Bunker prices which are correlated with crude oil prices are also stochastic and much harder to predict than bunker consumption, though some interesting atempts are done in Sheng et al. (2014). A wealth of literature and experts devote their time in how to predict such commodity prices and this text will not add to this. Instead we use the current bunker price at a port as the basic estimator for the future price. One enhancement could be to use the direction of the bunker forward price for the region of the port to predict the direction of the price.

9.6.2 Bunker Tanks

The vessels have multiple tanks as described in Sect. 9.2.4. Handling the tanks in an optimization model imposes a number of extra constraints.

Tank Limits The volume of bunker grows with increasing temperature, so in order to avoid overflowing tanks, the tanks may only be filled to around 98 % of their maximal capacity (at 25 °C). As mentioned, this should be considered in a data preprocessing phase, where the volume of the tanks should be decreased accordingly.

The suction systems of the tanks cannot be empty the tanks completely, leaving about 1 % of the capacity. This can be implemented as a reserve limit specific for each tank.

Commingling of Bunkers In general commingling of bunkers of different supplier, grade or batch is inadvisable as the combined properties are hard to predict and may form sediments or have unpredictable properties which the engine system can not handle. The constant mL indicates the limit for when a mixing is hazardous, and $mL = 9$ indicates that if bunker is mixed at a ratio greater than 9 to 1, it is not considered a commingling.

Two approaches can be taken to ensurethat bunker commingling does not take place:

- Tanks are not modelled explicitly, but it is assumed that the number of tanks will allow vessel crew to easily find a feasible solution of a concrete bunker plan.
- Bunker tanks can be modelled explicitly, adding constraints to ensure that commingling cannot take place.

Modelling Bunker Tanks Bunker tanks can be modelled by adding an extra index $t \in T$ for each bunker tank t. This means that parameters and variables l_m, $\delta_{i,b}$, $h_{i,b}$, $f_{i,b}$, and $D_{v,b}$ are replaced by $l_{m,t}$, $\delta_{i,b,t}$, $h_{i,b,t}$, $f_{i,b,t}$, and $D_{v,b,t}$.

The original constraints will be replaced by a constraint for each $t \in T$ where applicable. The mixing constraint then becomes:

$$mL \cdot (h_{i-1,b,t} - f_{i-1,b,t}) - l_{m,t} \leq (1 - \delta_{i,b,t} + \gamma_{i,b,t}) \cdot mL \cdot D_{v,b,t} \tag{9.34}$$

The mixing constraint forces the volume $l_{m,t}$ of newly bought bunker in tank t, to be at least mL times as much as the current volume of bunker, $h_{i-1,b,t} - f_{i-1,b,t}$, unless a penalty indicated by $\gamma_{i,b,t}$ is paid in the objective. It is not practice to dilute 9 units of bunker with one unit of newly bought bunker, so this isnot modelled. The term

$$\sum_{i \in I} \sum_{b \in B} \sum_{t \in T} mixpen \cdot \gamma_{i,b,t} \tag{9.35}$$

should be added to the objective, where $mixpen$ is the penalty for mixing. To avoid mixing the penalty should be set as $mixpen = \infty$. It is possible to move bunker between tanks, which can be used to avoid quarantine or mixing situations.

9.6.3 Speed Adjustment

The work of Yao et al. (2012) extent the bunker purchasing problem by considering the speeds of the vessels on a fixed itinerary of ports, but with some flexibility on the departure and arrival times at the ports. The model of Kim et al. (2012) control the full roundtrip speed and thus the roundtrip time and number of vessels assigned. This results in very slow service speeds, which arguably can be uncompetitive commercially. The variable speed is considered by letting the constant $F_{i,b}$ be a variable $F^{var}_{i,b}$ dependent on vessel speed following the equation:

$$F^{var}_{i,b} = k_b(k_1 V_i^3 + k_2) \tag{9.36}$$

Where V_i is the speed from port i to port $i+1$, k_1 and k_2 are constants and $0 \leq k_b \leq 1$ is the fraction of bunker of bunker type b used on leg b. Additional constraints impact V_i :

$$v_{min} \leq V_i \leq v_{max} \qquad \forall i \in I \tag{9.37}$$

$$A_i + t_i + \frac{d_i}{V_i} = A_{i+1} \qquad \forall i \in I \tag{9.38}$$

$$e_i \leq A_i \leq l_i \qquad \forall i \in I \tag{9.39}$$

Where constraint (9.37) ensures that the vessel does not exceed its min speed v_{min} and its max speed v_{max} on any legs. Constraint (9.38) sets the arrival times A_i and A_{i+1} in relation to the speed V_i, distance d_i and port time t_i. Constraint (9.39) ensures that the arrival time A_i are within the time windows of the port call e_i and l_i.

This model is cubic, but can be linearized as it is a convex function being minimized. For details please refer to Yao et al. (2012). For single vessel instances this problem can be solved by commercial solvers.

It should be noted that in practice the choice of port arrival time is impacted by many other factors than minimizing bunker costs. This could be the amount of cargo that needs to be loaded / unloaded at the port; When vessels that containers must be transshipped to arrive; When the port berth is available, etc. Due to the cubic nature of the bunker consumption curve, the best sailing speed for a vessel on a given rotation, will be an even speed throughout the rotation. This dictates that when buffer time is available at a port, the leg with highest speed before/after the port, should use the buffer to lower the speed. This could easily be handled inpreprocessing for the problem.

9.7 Solving the Multi-vessel Model

The fleet of a global liner shipping company may consist of hundreds of vessels, with many of these having overlapping schedules visiting the same hub ports. This means that the full problem can be of a very large size, making the MIP model impossible to solve for large instances as observed by Plum et al. (2014). This makes it interesting to consider a decomposition of the MIP model, to solve these large problem instances.

The arc flow model given by (9.19) – (9.30) is Dantzig-Wolfe decomposed on the variables l_m. Let R_v be the set of all feasible bunkering patterns for a vessel v, satisfying constraints (9.19) – (9.30), except (9.24). This set has an exponential number of elements. Each pattern $r \in R_v$ is denoted as a set of bunkerings. Let $u_r = \sum_{m \in M} (p_m \cdot l_m) + \sum_{i \in I} \sum_{v \in V} \sum_{b \in B} (\delta_{i,b} \cdot startcost_i)$ be the cost for pattern $r \in R_v$. Let λ_r be a binary variable, set to 1 if the bunkering pattern r is used. Let $o_{r,c}$ be the quantity purchased of contract c by pattern r. The BPCP can then be formulated as:

$$\min \sum_{v \in V} \sum_{r \in R_v} \lambda_r \cdot u_r + \sum_{c \in C} (\underline{s}_c \cdot \underline{w} + \overline{s}_c \cdot \overline{w}) \tag{9.40}$$

Subject to

$$\underline{q}_c - \underline{s}_c \le \sum_{v \in V} \sum_{r \in R_v} \lambda_r \cdot o_{r,c} \le \overline{q}_c + \overline{s}_c \qquad \forall c \tag{9.41}$$

$$\sum_{r \in R_v} \lambda_r = 1 \qquad \forall v \tag{9.42}$$

$$\lambda_r \in \{0, 1\} \qquad \forall r \tag{9.43}$$

The objective minimizes the costs of purchased bunker, startup costs and slack costs. Constraints (9.41) ensures that all contracts are fulfilled. Convexity constraints (9.42) ensure that exactly one bunker pattern is chosen for each vessel.

9.7.1 Pricing Problem

Let $\underline{\pi_c} \le 0$ and $\overline{\pi_c} \le 0$ be the dual variables for the upper and lower contract constraints (9.41), due to the structure of these constraints at least one of these will be 0 for each contract c. Let $\theta_v \in \mathbb{R}$ be dual variables for the convexity constraints (9.42). Then the pricing problem becomes:

$$\text{Min: } u_r + \sum_{c \in C} (\underline{\pi_c} - \overline{\pi_c}) - \theta_v \tag{9.44}$$

Subject to constraints (9.19) – (9.30), except (9.24).

This pricing problem is a Mixed Integer Program, considering a single vessel. This size of problem can be solved in reasonable time by a standard MIP solver, as done in Plum and Jensen (2007). Columns λ_r with negative reduced cost will then be added to the master problem, also solved as a MIP.

9.7.2 Column Generation Algorithm

Due to the large number of columns in model (9.41) to (9.43) Plum et al. (2014) proposed to solve the LP relaxed model by *Column Generation*. Using the generated columns from the LP-solution, the resulting problem is then solved to integer optimality using a MIP solver, leading to a heuristic solution for the original problem.

Initially all dual variables are set to zero, a subproblem is constructed for each vessel and solved as a MIP problem. The first master problem is then constructed with one solution for each vessel as columns. This master is solved and the first dual values are found. The subproblems are resolved for all vessels (only the objective coefficients for the contracts needs updating) and new columns are generated for the master. This continues until no negative reduced cost columns can be generated, and the LP optimal solution is achieved.

In the next step, the problem is solved as a MIP, providing an integral solution. The subproblems only need to find *a* negative reduced costs column, to ensure progress of the algorithm. This means that initially they are allowed to return solutions with considerable subproblem gaps. As the algorithm progresses, the allowable subproblem gap is reduced, until it reaches the tolerance level.

9.7.3 Dual Stabilization

A simple form of dual stabilization has been used in the implementation by Plum et al. (2014) to speed up convergence. The Box-step method described in Marsten et al. (1975) imposes a box around the dual variables, which are limited from changing more than π_{max} per iteration. This has been motivated by the dual variables only taking on values $\{-\underline{w}, \overline{w}, 0\}$ in the first iteration, these then stabilize at smaller numerical values in subsequent iterations.

9.7.4 Interpretation of Dual Values

The dual variables $\underline{\pi_c}$ and $\overline{\pi_c}$ for the upper and lower contract constraints (9.41) can be used to evaluate the gain of a given contract.

Using best estimates for bunker consumption and prices (current prices for instance) together with known or expected contracts, a baseline bunker purchasing plan could be run. A new scenario could then be constructed with the addition of the considered contract and by analyzing the output, it could be seen whether the overall costs of the scenario increased or decreased as compared with the baseline.

Another investigation could be to solely consider the baseline's final dual variables, $\underline{\pi_c}$ and $\overline{\pi_c}$, and depending on the magnitude of these evaluate the contracts effect. As these dual values are the same for all subproblems, they can be interpreted as balancing out the price of the contract, increasing the price if it is a popular contract or decreasing it otherwise, converging when they are in balance. The magnitude of this will be proportional to the contracts gain.

Table 9.4 Instances of varying sizes for the BPCP. **Instance** is the name of the instance, **Size** is a grouping of the instances. V the number of vessels, P the number of port calls, C the number of Contracts

Instance	Size	V	P	C
RULED	Small	6	1048	29
FRFSM	Small	8	2128	10
ZADUR	Small	49	5973	35
US_WC	Small	32	6022	68
USNWK	Medium	49	9048	69
USSAV	Medium	50	9194	23
PABLB	Medium	65	9817	27
AEJAL	Medium	80	15442	9
09_H2	Large	408	16214	307
11_H2	Large	572	18426	254
10_H1	Large	469	18704	332
10_H2	Large	534	21907	424
11_H1	Large	609	23453	376
HKHKG	Large	158	29177	20
10_FY	Large	535	40611	756

9.8 Computational Results

According to Plum and Jensen (2007) the simple model proposed in Sect. 9.3.1 can be solved in a couple of seconds, since every vessel is considered independently.

The model with bunker contract presented in Sect. 9.5 is more difficult to solve as reported in Plum et al. (2014). To give an impression of the complexity, we will now present computational results for a number of real-life instances including up to 500 vessels, 40,000 ports, and 750 contracts. These instances are representative for what problems need to be solved in a major liner shipping company.

The **MIP** model described in Sect. 9.3.1 has been implemented in **CPLEX 12.2**, while the column generation algorithm outlined in Sect. 9.7.2 has been implemented in **ILOG OPL** as modelling language and **CPLEX 12.2** as LP/MIP solver. We will use **DW** to denote the column generation implementation.

Real life data for a large number of liner vessels describing their schedules, consumptions, tank capacities and other relevant data has been made available by **Maersk Oil Trading**, who have also supplied data on a large number of actual bunker contracts and spot prices available in a range of ports. Based on these data a number of instances have been constructed to test the scalability and performance of the implementations. Due to confidentiality reasons the price's have been distorted by $\pm 10\,\%$. This small amount of noise, however, does not affect the main structure of the problem. The penalty $\underline{w}$ for violating minimum volume is set at 200 \$/mt, and the penalty $\overline{w}$ for breaking the maximal constraint at 50 \$/mt. If a bunker price is

Table 9.5 Results and performance of **MIP** and **DW** implementation. **Instance** is the instance name. Obj_{MIP} is the best found solution for the **MIP** algorithm, and LB_{MIP} is the best found lower bound. Gap_{MIP} is the resulting gap between upper and lower bound and t_{MIP} is the time used in seconds. Obj_{DW} is the best found solution for the **DW** algorithm, and LB_{DW} is the best found lower bound. Gap_{DW} is the resulting gap between upper and lower bound and t_{DW} is the time used in seconds

Instance	Obj_{MIP}	LB_{MIP}	Gap_{MIP}	t_{MIP}	Obj_{DW}	LB_{DW}	Gap_{DW}	t_{DW}
RULED	5.404$e+7$	5.404$e+7$	0.00 %	1083	5.408 e+7	5.404 e+7	0.08 %	118
FRFSM	1.319$e+8$	1.319$e+8$	0.00 %	21	1.321 e+8	1.319 e+8	0.20 %	86
ZADUR	7.064$e+8$	7.063$e+8$	0.02 %	609	7.071 e+8	7.064 e+8	0.10 %	653
US_WC	6.628$e+8$	6.626$e+8$	0.03 %	481	6.654 e+8	6.627 e+8	0.41 %	1142
USNWK	9.067$e+8$	9.063 e+8	0.03 %	834	9.077 e+8	9.066 e+8	0.11 %	1114
USSAV	9, 830$e+8$	9.826 e+8	0.04 %	775	9.830 e+8	9.829 e+8	0.00 %	399
PABLB	1.108$e+9$	1.107 e+9	0.06 %	906	1.108 e+9	1.108 e+9	0.01 %	672
AEJAL	1.490$e+9$	1.489 e+9	0.03 %	686	1.490 e+9	1.490 e+9	0.00 %	415
09_H2	2.115$e+9$	2.113 e+9	0.10 %	1160	2.120 e+9	2.115 e+9	0.22 %	8642
11_H2	2.478$e+9$	2.475 e+9	0.09 %	1107	2.479 e+9	2.477 e+9	0.07 %	9411
10_H1	2.255$e+9$	2.253 e+9	0.09 %	1181	2.259 e+9	2.255 e+9	0.19 %	7267
10_H2	Out of Mem				2.529 e+9	2.526 e+9	0.12 %	10649
11_H1	Out of Mem				3.217 e+9	3.214 e+9	0.09 %	10075
HKHKG	Out of Mem				3.427 e+9	3.427 e+9	0.00 %	4344
10_FY	Out of Mem				4.835 e+9	4.807 e+9	0.59 %	28922

not available at a port, the price is set at 1000 \$/mt. Details about the instances can be seen in Table 9.4.

An overview of the performance and results can be found in Table 9.5. It can be seen that the **DW** model is able to solve the problem for all instances. For larger instances **MIP** runs out of memory and finds no solution, due to the size of the instances and their resulting MIPs. Both models find solutions with very small gaps, but still considerable absolute gap's to the optimal solution. **MIP** only finds optimal solutions for the smallest instances, for all medium and large instances the solver runs out of memory before it has closed the gap. **DW** is able to find solutions with relatively small gaps for even the largest problem instances covering all vessels and all contracts on a global level. In practice the resulting gaps of the algorithms, can be much smaller since we benchmark against a lower boundand not against the optimal solution.

9.9 Conclusion and Further Work

We have given an in-depth description of how to optimize bunker purchasing in liner shipping. First, a mathematical model for bunker bought on the spot market was presented, and various extensions from the literature were discussed. Next, a model for bunker purchasing with contracts was presented, and a novel solution approach based on decomposition was described.

Since bunker prices are stochastic of nature, future research should be focused on modeling the price fluctuation. However, the models tend to become quite complex and difficult to solve as observed by Plum and Jensen (2007), while only adding small extra improvements to the results. So a trade-off must be done between model complexity and gain in bunker costs. The work of Sheng et al. (2014) shows some promising developments in this important direction.

Also, instruments from finance could be used to control risk in bunker purchasing, and to increase the margins on oil trade. Bunker purchasing for liner ships constitutes such a big market that it deserves a professional trading approach.

Acknowledgements The authors gratefully acknowledge the support from The Danish Council for Strategic Research to the ENERPLAN project. This work would not have been possible without the adept guidance and support from experienced Maersk Oil Trading and Maersk Line staff, in particular we would like to thank Claus Anker and Mikkel M. Sigurd who have given invaluable support for this project.

References

Abdelghany, K., Abdelghany, A., & Raina, S. (2005). A model for the airlines fuel management strategies. Journal of Air Transport Management, 11(4), 199–206.

Acosta, M., Coronado, D., & Cerban, M. D. M. (2011). Bunkering competition and competitiveness at the ports of the gibraltar strait. *Journal of Transport Geography, 19*(4), 911–916.

Besbes, O., & Savin, S. (2009). Going bunkers: The joint route selection and refueling problem. *Manufacturing and Service Operations Management, 11*(4), 694–711.

Boutsikas, A. (2004). The bunkering industry and its effect on shipping tanker operations. Master's thesis, Massachusetts Institute of Technology. Dept. of Ocean Engineering.

Brouer, B. D., Alvarez, J. F., Plum, C. E. M., Pisinger, D., & Sigurd, M. M. (2013). A base integer programming model and benchmark suite for liner shipping network design. *Transportation Science*. Accepted by Transportation Science.

Bunkerwire, P. (2013). http://www.platts.com/IM.Platts.Content/ProductsServices/Products/bunkerwire.pdf.

California Legislative Analyst's Office. (2001). California bunker sales tax. http://www.lao.ca.gov/2001/bunker_fuel/012501_bunker_fuel.html

Christiansen, M., Fagerholt, K., & Ronen, D. (2004). Ship routing and scheduling: Status and perspectives. *Transportation Science, 38*(1), 1–18.

Christiansen, M., Ronen, D., Nygren, B., & Fagerholt, K. (2007). Chapter 4, Maritime transportation. In C. Barnhart & G. Laporte (Eds.), *Handbooks in operations research and management sciences* (Vol. 14, pp. 189–284). Elsevier B.V.

Christiansen, M., Fagerholt, K., Nygreen, B., & Ronen, D. (2013). Ship routing and scheduling in the new millennium. *European Journal of Operational Research,228*(3), 467–483.

Corbett, J. J., Wang, H., & Winebrake, J. J. (2009). The effectiveness and costs of speed reductions on emissions from international shipping. *Transportation Research Part D: Transport and Environment, 14*(8), 593–598.

Fagerholt, K., Laporte, G., & Norstad, I. (2009). Reducing fuel emissions by optimizing speed on shipping routes. *The Journal of the Operational Research Society, 61*(3), 523–529.

Farina, F. (2012). Bunker optimization with contracts. Master's thesis, Universita Di Bologna.

Kellerer, H., Pferschy, U., & Pisinger, D. (2004). *Knapsack problems*. Springer.

Kim, H. (2014). A lagrangian heuristic for determining the speed and bunkering port of a ship. *Journal of the Operational Research Society 65,* 747–754.

Kim, H. -J., Chang, Y. -T., Kim, K. -T., & Kim, H. -J. (2012). An epsilon-optimal algorithm considering greenhouse gas emissions for the management of a ship's bunker fuel. *Transportation Research Part D: Transport and Environment, 17*(2), 97–103.

Lin, S. -H. (2008). Finding optimal refueling policies in transportation networks. In R. Fleischer & J. Xu (Eds.), *Algorithmic aspects in information and management*, (volume 5034 of *Lecture Notes in Computer Science*), (pp. 280–291). Berlin: Springer.

Marsten, R. E., Hogan, W. W., & Blankenship, J. W. (1975). The boxstep method for large-scale optimization. *Operations Research, 23*(3), 389–405.

Meng, Q., Wang, S., & Liu, Z. (2012). Network design for shipping service of large-scale intermodal liners. *Transportation Research Record: Journal of the Transportation Research Board, 2269,* 42–50.

Notteboom, T. E., & Vernimmen, B. (2009). The effect of high fuel costs on liner service configuration in container shipping. *Journal of Transport Geography, 17*(5), 325–337.

Oh, H. C., & Karimi, I. A. (2010). Operation planning of multiparcel tankers under fuel price uncertainty. *Industrial & Engineering Chemistry Research, 49*(13), 6104–6114.

Plum, C. E. M., & Jensen, P. N. (2007). Minimization of bunker costs. Master's thesis.

Plum, C. E. M., Pisinger, D., & Sigurd, M. M. (2014). A service flow model for the liner shipping network design problem. *European Journal of Operational Research, 235*(2), 378–386.

Plum, C. E. M., Pisinger, D., & Jensen, P. N. (2014). Bunker purchasing with contracts. *Maritime Economics and Logistics*. To appear.

Ronen, D. (2010). The effect of oil price on containership speed and fleet size. *Journal of the Operational Research Society, 62,* 211–216.

Sheng, X., Lee, L., & Chew, E. (2014). Dynamic determination of vessel speed and selection of bunkering ports for liner shipping under stochastic environment. *OR Spectrum, 36*(2), 455–480.

Stopford, M. (2009). *Maritime economics* (3rd. ed., 840 p.). London: Routledge.

Stroup, J. S., & Wollmer, R. D. (1992). A fuel management model for the airline industry. *Operations Research, 40*(2), 229–237.

Suzuki, Y. (2008). A generic model of motor-carrier fuel optimization. *Naval Research Logistics (NRL), 55*(8), 737–746.

Suzuki, Y. (2011). A new truck-routing approach for reducing fuel consumption and pollutants emission. *Transportation Research Part D: Transport and Environment, 16*(1), 73–77.

Vilhelmsen, C., Lusby, R., & Larsen, J. (2014). Tramp ship routing and scheduling with integrated bunker optimization. *EURO Journal on Transportation and Logistics, 3*(2), 143–175.

Wang, S., & Meng, Q. (2012). Robust schedule design for liner shipping services. *Transportation Research Part E: Logistics and Transportation Review, 48*(6), 1093–1106.

Wang, S., Meng, Q., & Liu, Z. (2013). Bunker consumption optimization methods in shipping: A critical review and extensions. *Transportation Research Part E: Logistics and Transportation Review, 53*(0), 49–62.

Yao, Z., Ng, S. H., & Lee, L. H. (2012). A study on bunker fuel management for the shipping liner services. *Computers and Operations Research, 39*(5), 1160–1172.

Chapter 10
Ship Route Schedule Based Interactions Between Container Shipping Lines and Port Operators

Shuaian Wang, Abdurahim Alharbi and Pam Davy

Abstract This chapter examines a practical tactical liner ship route schedule design problem, which involves the interaction between container shipping lines and port operators. When designing the schedule, the availability of each port in a week, i.e., port time window, is incorporated. As a result, the designed schedule can be applied in practice without or with only minimum revisions. We assume that each port on a ship route is visited only once in a round-trip journey. This problem is formulated as a nonlinear non-convex optimization model that aims to minimize the sum of ship cost, bunker cost and inventory cost. In view of the problem structure, an efficient dynamic-programming based solution approach is proposed. First, a lower bound of the number of ships is determined, and then we enumerate all possible numbers of ships. Given the number of ships, we can construct a space-time network that discretizes the time and represents the design of schedule. The optimal schedule in such a space-time network can be obtained by dynamic programming. The algorithm stops when the lower bound is not smaller than the optimal total cost of the best solution obtained. The proposed solution method is tested on a trans-Pacific ship route.

S. Wang (✉) · A. Alharbi · P. Davy
School of Mathematics and Applied Statistics, University of Wollongong,
Wollongong, NSW 2522, Australia
e-mail: wangshuaian@gmail.com

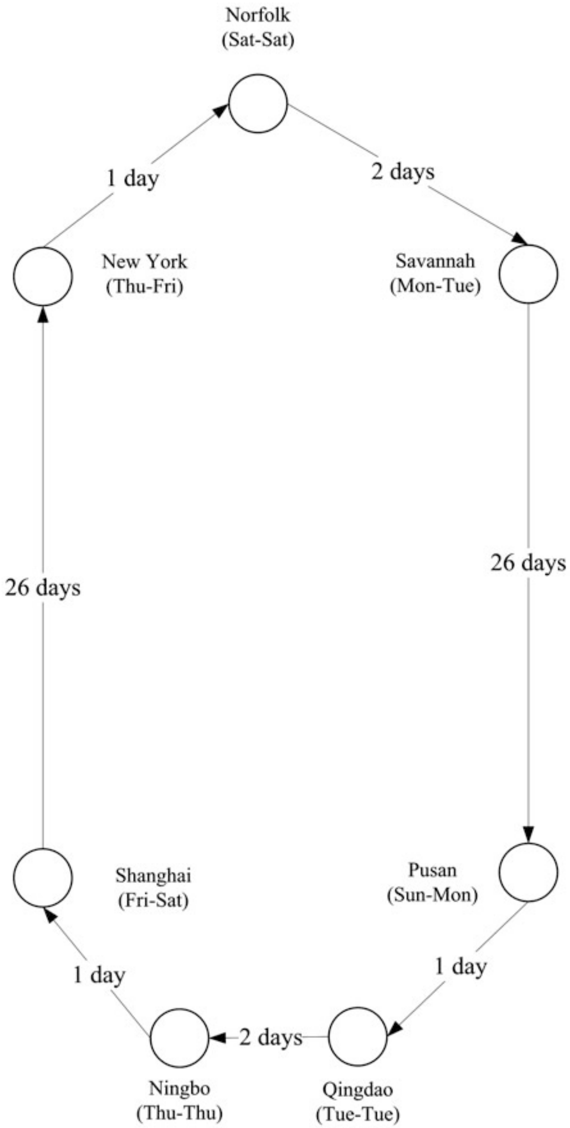
S. Wang
Strome College of Business, Old Dominion University,
Norfolk, Virginia 23529, USA

A. Alharbi
e-mail: afma987@uowmail.edu.au

P. Davy
e-mail: pjd@uow.edu.au

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics*,
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_10

Fig. 10.1 NCE service provided by OOCL (2013)



10.1 Introduction

Liner shipping mainly involves the transportation of containerized cargo (containers) such as manufactured products, food, and garment. Liner shipping services have fixed sequences of ports of call and fixed schedules, i.e., arrival and departure times at each port of call, similar to public transport operations. Liner services are announced in advance to attract potential customers. For example, Fig. 10.1 shows a liner service named North & Central China East Coast Express (NCE) provided by Orient Overseas Container Line (OOCL 2013). The ports of call and schedule are published in the website of OOCL. Customers can arrange the delivery of their cargo based on the available date of the cargo at the origin port and the expected arrival date at the destination port. For instance, a customer that has 20 containers to be transported from Pusan to New York may contact OOCL to transport the containers. As ships visit Pusan on Sunday, the customer has to make sure that the containers are stacked in the container yard of Pusan before Saturday, so that containers can be loaded to a ship when the ship arrives. The ship will not directly transport the containers from Pusan to New York. By contrast, the ship will transport the container via Qingdao, Ningbo, Shanghai, and finally to New York. At the port of New York, the containers will be unloaded from the ship.

In the process of container transportation, the main role of container port is to load and unload containers. First, a containership informs a port operator the estimated arrival time, and then the port operator makes a plan for servicing the ship. When the ship arrives, tug boats will tow the ship to the berth. Then the ship will be moored, and quay cranes will start to load and unload containers for the ship. At the same time, yard trucks will transport containers from and to the quay side. The container handling operation may take up to two days. After that, the ship is unmoored, and tug boats tow the ship out of the port.

10.1.1 Interactions Between Shipping Lines and Port Operators

Shipping lines and port operators interact with each other to fulfill container transport services. Their interactions occur in several aspects. Shipping lines need to choose which ports to visit on the shipping services and determine which ports serve as the hub ports. Port operators need to prioritize shipping lines when the ports have limited capacity. Ports provide services to ships such as bunkering, pilotage and towage, berthing, container handling and temporarily storing containers (including empty containers) in container yards. In the sequel, we take a closer look at two interactions: ship storage planning and berth allocation.

Ship Storage Planning

Ship storage plan determines how to stow a set of containers of different types into a set of available locations within a ship at a particular port, subject to some structural, stability and operational constraints related to both the containers and the ship, while minimizing the total container re-handling cost or time caused by unloading a container below other containers. A good ship storage plan is helpful for port operators to efficiently load and unload containers, and thus shortens the unproductive port time of ships.

As proved by Aslidis (1989), the ship storage planning (SSP) problem is NP-hard and several heuristic methods or computer simulation approaches have been proposed in the past two decades.

Aslidis (1989), Imai and Miki (1989) and Aslidis (1990) have contributed three pioneering works on the SSP problem. Aslidis (1989, 1990) have examined the SSP problem with the objective of minimizing the total container overstorage cost and proposed heuristic solution methods. Imai and Miki (1989) considered the minimization of container loading-related rearrangements. Avriel and Penn (1993) formulated the SSP problem as a binary integer linear programming model and argued that exact algorithms solving the integer programming model were too slow even after some preprocessing. Ambrosino and Sciomachen (1998) derived some rules for determining a better ship stowage configuration using a constraint satisfaction approach. Avriel et al. (1998) dealt with the SSP problem without taking into account ship stability and several other constraints. They presented a binary integer linear programming formulation and found that the optimal solution was difficult to obtain because of the large number of binary variables and constraints involved in the model. Consequently, they developed a heuristic procedure called the suspensory heuristic procedure. However, they assumed that the ship only had a large cargo bay without considering the hatch covers and stability. Wilson and Roach (1999, 2000) have proposed a methodology for generating computerised container stowage plans, which embodied a two-stage process: In the first stage, a branch-and-bound algorithm was utilized for solving the problem of assigning containers to a bay's block in a ship, and in the second stage a tabu search algorithm was employed to assign specific locations for specific containers. Wilson et al. (2001) applied a genetic algorithm approach for solving a ship stowage pre-planning problem. Dubrovsky et al. (2002) used a genetic algorithm for minimizing the number of container movements in the stowage planning. Ambrosino et al. (2004) presented an integer linear programming model for the SSP. Also, they proposed a tangible approach comprising heuristic pre-processing and pre-stowing procedures that allowed relaxation of some constraints of the exact model. Ambrosino et al. (2006) developed a three-phase algorithm for solving the SSP problem based on a partitioning procedure that split the ship into different portions and assigned containers on the basis of their destination.

Moreover, there are also a few software packages used by shipping lines for ship stowage planning, for example, the PowerStow Container Stowage System (www.navis.com), LOADMASTER X5 (http://www.kockumsonics.com) and CargoMax (www.herbertsoftware.com).

Berth Allocation

Another interaction between shipping lines and port operators is berth allocation. In fact, the berth allocation problem (BAP), which is the assignment of quay space and service time to vessels for container loading and unloading, is one of the essential quay-side decision problems faced by port operators. A good BAP can shorten the unproductive port time of ships, enabling liner shipping companies to make a higher profit.

The BAP can be classified according to different criteria. First, there are discrete BAP (DBAP) where each berth can serve one ship at a time, and continuous BAP (CBAP) with a long straight quay and how many ships can be accommodated at the same time depends on the sizes of the ships. Second, BAP can be classified as being either static (SBAP) or dynamic (DynBAP). In SBAP, all ships are already in the port when the berth allocation is planned, whereas in DynBAP some ships are still on the voyage to the port when the port operator allocates berths. The SBAP is applicable when the port is highly congested. Third, BAP can occur at the operational level (OBAP), or tactical level (TBAP). The OBAP covers a planning horizon of usually at most one week and the TBAP aims to support port operators to negotiate with shipping lines. If TBAP accounts for the periodicity of vessel schedules, e.g., weekly arrival patterns of containerships, then if a vessel is serviced at a berth on day 7 and day 8, other vessels cannot use the berth on day 1, because day 8 and day 1 correspond to the same day in a week. The time horizon of this type of TBAP is a cylinder whose circumference equals 1 week. Hence, the resulting models (Moorthy and Teo 2006, Zhen et al. 2011b) are significantly different from OBAP models. If in the TBAP vessels do not arrive periodically, the time horizon is simply a rectangle with an open end and the models are very similar to OBAP models.

Besides determining the berthing time and location, some studies on DynBAP (either DBAP or CBAP and either TBAP or OBAP) also integrate other decision issues such as quay crane assignment, quay crane scheduling, container storage planning at yard, and yard truck scheduling. The models on DynBAP all aim at providing berthing and other related services at minimum cost (cost associated with quay cranes and yard trucks). However, different models have different definitions for service. Most studies assume that each ship has a preferred arrival time. Giallombardo et al. (2010) is an exception in that it examined a TBAP and assumed that there was no difference for shipping lines when their ships were scheduled to arrive. The objective was to minimize the container handling time of ships by choosing quay crane assignment profiles.

The studies considering the preference of ship arrival times can be classified into four different lines, which are briefly summarized as follows. The first line aims to minimize the total service time (turnaround time) of all ships, including waiting time for berths and container handling time, or total weighted service time where different ships have different weights, for example, Imai et al. (2001, 2003, 2005, 2008a), Cordeau et al. (2005), Moorthy and Teo (2006), Golias et al. (2009b, 2010a), Lee et al. (2010). Note that if the handling time is constant, minimizing the service time is equivalent to minimizing the waiting time. Similarly, Imai et al. (2008b) required that if a ship's waiting time exceeded a certain limit, the ship must be served at an external terminal, and the target is to minimize the total service time of ships at the external terminal. Golias et al. (2009a) considered two objectives: minimizing the total service time of preferential customers, and minimizing the total service time of all vessels. The second line minimizes the total tardiness cost, which is the finish operation time (real departure time) minus the expected departure time if the former is larger, and 0 otherwise, for instance, Kim and Moon (2003), Chang et al. (2010), Zhen et al. (2011b). In addition, Han et al. (2010) proposed a proactive approach for a BAP with quay crane scheduling and stochastic arrival and handling time. They took into account the expected value and standard deviation of the total service time and weighted tardiness of all ships. Chen et al. (2012) minimized the maximum relative tardiness of all ships. The third line formulates the penalty for earliness and tardiness in greater details. Meisel and Bierwirth (2009) investigated a CBAP with quay crane allocation. They assumed that each ship has an expected arrival time, an earliest start operation time, expected finish operation time, and latest allowed finish operation time. All of these time components were included in the objective function. Zhen et al. (2011a) developed an integrated model for the TBAP with yard operations planning. The model minimized the weighted sum of deviation from vessels' expected turnaround time intervals and the operations cost associated with transshipment containers. The fourth line incorporates the bunker cost of the vessels in the models. Golias et al. (2010b) considered the following elements in the objective function: (i) the total service time, (ii) the tardiness, and (iii) the emissions and fuel cost for all vessels while in transit to their next port of call. By contrast, Du et al. (2011) incorporated the tardiness and the fuel cost for all vessels while in transit from their current positions to the focal port of the BAP. These berth allocation studies are all from the points of view of port operators.

10.1.2 Liner Ship Route Schedule Design with Port Time Windows

This book chapter examines the interaction between shipping lines and port operators on schedule design from the viewpoint of shipping lines. Schedule design for a liner service (ship route) is a tactical-level planning decision that is made every 3–6 months. To design the schedule of a ship route, the first factor to be considered is the availability of the ports. Since a port needs to provide services for many shipping lines and many ships, it cannot guarantee the availability of services whenever a ship arrives. For instance, a port may be able to provide services on Monday, Tuesday, and Friday, and is fully occupied on Wednesday, Thursday, Saturday, and Sunday. We use the term "port time window" to refer to the time in a week that a port can provide services to ships. Hence, schedule design is subject to the constraint of port time windows. Moreover, because of the fast growth of container trade and the long time required for the construction/expansion of port capacity, ports tend to be more congested. As a result, it is important to consider the availability of ports in schedule design. Otherwise the designed schedule may be infeasible in reality.

It should be noted that "port time window" here is different from the "time window" in other problems, e.g., the vehicle routing problems (VPRs), as shown in Fig. 10.2. In fact, in most other problems, time window is an interval that defines a convex set. However, in liner ship route schedule design, port time window defines a set of available times in a week that the port can provide berthing services, and more often than not, the set is disconnected and non-convex. Moreover, because of the weekly frequency of liner shipping services, the port time window should be considered from the viewpoint of a loop rather than a line or line segment. Take

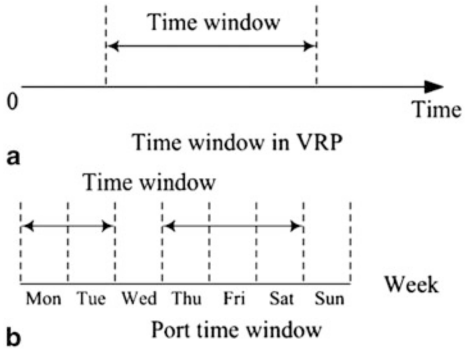


Fig. 10.2 Difference of time windows

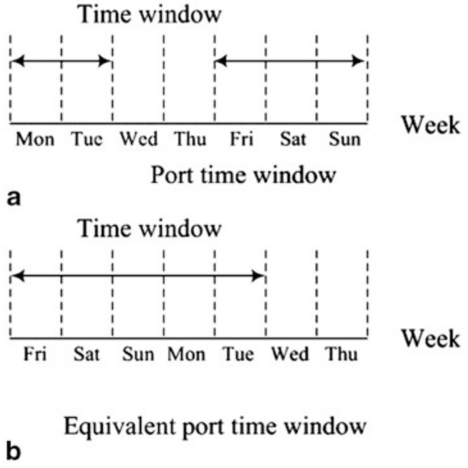


Fig. 10.3 Weekly property of port time window

Fig. 10.3 as an example. The port time window in Fig. 10.3a is equivalent to that in Fig. 10.3b.

Besides the availability of port time windows, the design of schedule is also influenced by the ship costs, bunker costs, and inventory costs. Liner services are usually weekly, which means that the round-trip journey time (weeks) of a ship route is equal to the number of ships deployed on it. As a result, sailing at a higher speed will reduce the round-trip journey time, thereby the number of ships required and the ship cost. However, a higher speed implies a higher bunker cost: the daily fuel consumption of ships increases approximately proportional to the sailing speed cubed (Ronen 2011). At the same time, a higher speed leads to a shorter transit time of containers from origin to destination, and therefore a lower inventory cost (OOCL 2006). Consequently, in schedule design a liner shipping company must balance the trade-off between ship cost, bunker cost, and inventory cost, while considering port time windows.

10.1.3 Objectives

The objective of this chapter is to address the liner ship route schedule design problem with port time windows (SDPTW). We assume that each port on the ship route is visited only once in a round-trip journey. We design the arrival time at each port of call on the ship route that satisfies the port time window constraint while minimizing the sum of ship cost, bunker cost, and inventory cost. The designed schedule is feasible in that it takes into account port time windows. The designed schedule is also optimal because the total cost of ships, bunker, and inventory is minimized. Therefore, this problem is of significant value for liner shipping companies.

The rest of the chapter is organized as follows. Section 10.2 reviews relevant studies on schedule design. Section 10.3 describes the problem. Section 10.4 formulates a mathematical model for the problem. Section 10.5 proposes a dynamic programming based holistic solution approach to address the problem. Section 10.6 reports a case study based on the NCE service of OOCL. Section 10.7 concludes and points out future research directions.

10.2 Literature Review

According to reviews of Christiansen et al. (2004), Christiansen et al. (2013) and Meng et al. (2014), most studies on liner shipping operations focus on network design, ship deployment, and container routing with fixed schedules. For example, Shintani et al. (2007), Agarwal and Ergun (2008), Alvarez (2009), Meng and Wang (2011a), Meng et al. (2012a), Reinhardt and Pisinger (2012), Wang and Meng (2013), Brouer et al. (2013a), and Wang and Meng (2014) have examined ship route design/network design/network alteration problems; Fagerholt (1999), Gelareh and Meng (2010), Meng and Wang (2011b), Wang et al. (2011), Meng and Wang (2012), Meng et al. (2012b), Wang and Meng (2012a), and Wang (2013) have investigated fleet deployment problems; Bell et al. (2011), Dong and Song (2012), and Wang et al. (2013) have studied container routing problems.

Liner shipping schedule design occurs both at the tactical level and the operational level. At the tactical level, the designed schedule will be announced in the website of the shipping line, so that customers can book ship slots for transporting containers. Nevertheless, in reality the announced schedule may not be strictly adhered to, due to adverse weather conditions, port congestion, mechanical breakdown, etc. Therefore, at the operational level, the shipping line may adjust the arrival times at each port of call based on real-time information. This is the operational-level schedule design problem.

10.2.1 Tactical-Level Schedule Design

There are only a few studies on ship route schedule design. Mourão et al. (2001) analyzed a small hub-and-spoke network at the tactical level. The network consisted of two routes—a feed route and a main route—and one pair of ports. It assumed that all containers must be transshipped at the hub port in the feeder route. The main route had two possible schedules: Monday roster and Thursday roster. Two integer programming models were developed. The decision variables in the first model include: number of mainline ships in each type assigned to the Monday roster, number of mainline ships in each type assigned to the Thursday roster, and number of feeder ships assigned to the feeder route. In the second model, the decision variables were: number of voyages per year of the mainline ships in each type assigned to the Monday roster, number of voyages per year of the mainline ships in each type assigned to the Thursday roster, and number of voyages per year of feeder ships. The inventory costs of the containers to be shipped were considered in the objective function. These two models were solved by Excel.

Wang and Meng (2011) investigated the schedule design and container routing problem in liner shipping. They considered a general liner shipping network with many ports, many ship routes, and many origin-destination (OD) pairs. Containers in each OD pair had more than one path to be transported from origin to destination, and these paths were assumed to be given a priori. Containers in each OD pair had a market-level transit time to ensure that the container delivery service was competitive. In particular, if the real transit time was longer than the market-level transit time, a penalty was incurred; if the real transit time was shorter than the market-level transit time, a bonus was given. It was assumed that the sailing speed of ships and the time spent at each port of call were all fixed. Hence, the main decision variables were when to arrive at the first port of call on each ship route. At the same time, container routing with transshipment were incorporated. In fact, how containers were transported affected the schedule design. Hence, schedule design and container routing were studied in a holistic manner. The formulation for the schedule design and container routing problem was nonlinear, non-continuous and non-convex. An efficient genetic local search heuristic was developed. Computational results showed that the genetic local search heuristic could efficiently find good quality solutions. Moreover, the model for the container routing sub-problem could be separately used to optimize the day-to-day container routing decisions for the realized container shipment demand after the schedules have been designed.

Qi and Song (2012) designed an optimal containership schedule for a liner ship route to minimize the total expected fuel consumption. They considered uncertain port time and weekly frequency. They defined the level of service as the probability that the containership would arrive at a port no later than the published arrival time. They analytically studied the special case of 100 % service levels. By proving the convexity and differentiability of the objective function, it was shown that the optimal schedule could be obtained by solving a nonlinear programming problem. With further assumption of identical distribution of the uncertain parts of port times, they analytically derived some properties of an optimal schedule, which led to useful managerial insights. For example, the shortest leg was the most problematic leg when designing the optimal schedule to achieve 100 % service level and to minimize the emissions within the speed constraints, and therefore a liner shipping company should plan relatively longer time for a short leg. A general optimal ship scheduling problem was formulated, and the formulation was solved by simulation-based stochastic approximation methods. They validated the model and the properties by numerical studies. Based on a real liner case study with various scenarios analysis, they found significant fuel savings could be achieved from their model compared to the company's original schedule or to the schedule based on deterministic data, especially for the cases with larger degree of uncertainties. They also found that the total fuel consumption could be reduced by sacrificing the service levels starting from the shortest legs; whereas as the vessel lateness penalty increased, higher service levels tended to be maintained and they became evener among all port-of-calls. This would help liner companies better understand the tradeoff between the fuel consumption and the service level.

Wang and Meng (2012c) examined the design of liner ship route schedules that could hedge against the uncertainties in port operations, which included the uncertain wait time due to port congestion and uncertain container handling time. They assumed that if a ship arrived at a port later than planned, then the penalty cost first increased linearly with the delay, and when the delay exceeded a particular threshold, the penalty cost did not change any more because the customers already resorted to other approaches to handle the delay. They further assumed that if a ship was delayed, it would try to catch up with the planned schedule as early as possible by sailing at the fastest speed. The designed schedule was robust in that uncertainties in port operations and schedule recovery by fast steaming were captured endogenously. The number of ships required to maintain a weekly frequency was considered as a decision variable. The objective function minimized the ship operating cost, expected bunker cost, and the penalty cost for delay. This problem was formulated as a mixed-integer nonlinear stochastic programming model. A solution algorithm which incorporated a sample average approximation method, linearization techniques, and a decomposition scheme, was proposed. Numerical experiments based on a long-haul ship route of Maersk Line were carried out. The ship route covered two trade lanes: trans-Pacific and trans-Atlantic, and three regions: Asia, America, and Europe, and had the sequence of ports of call as follows: Tokyo (1) → Kobe (2) → Chiwan (3) → Hong Kong (4) → Kaohsiung (5) → Busan (6) → Kobe (7) → Tokyo (8) → Balboa (9) → Manzanillo (10) → Miami (11) → Jacksonville (12) → Savannah (13) → Charleston (14) → New York (15) → Antwerp (16) → Felixstowe (17) → Bremerhaven (18) → Rotterdam (19) → Le Havre (20) → New York (21) → Norfolk (22) → Charleston (23) → Manzanillo (24) → Balboa (25) → San Pedro (26) → Oakland (27) → Tokyo (1). The numerical experiments demonstrated that the algorithm obtained near-optimal solutions with the stochastic optimality gap less than 1.5 % within reasonable time.

Wang and Meng (2012b) extended the work of Wang and Meng (2011). Both works have studied a liner shipping network, which contrasted Qi and Song (2012) and Wang and Meng (2012c), and both works have required a certain level of service in terms of OD transit time. Wang and Meng (2012b) was the first attempt to examine the optimal sailing speed function in view of sea contingency to minimize bunker consumption. The optimality condition for the sailing speed and the optimal sailing speed function with time were derived. They also contributed to the line of literature on optimization of sailing speed to control bunker consumption by providing an efficient and exact cutting-plane based solution algorithm. Moreover, they addressed the practical schedule design problem arising in liner shipping industry while considering port-to-port transit time with transshipment and sea contingency and uncertain port time. The port-to-port transit time with transshipment issue was solved with a mixed-integer programming model; sea contingency was investigated in the optimality condition of sailing speed; and the uncertain port time was addressed by proving the convexity of the expected bunker cost on each voyage leg with regard to the inter-arrival time between the two consecutive portcalls of the leg. The novel holistic solution algorithm exploited the special structure of the decision problem and integrated several techniques in a nice manner. The proposed model and algorithm were applied to an Asia-Europe-Oceania shipping network provided by a global liner shipping company. The network had a total of 46 ports in Asia, Europe, and Oceania. These 46 ports were served by 11 ship routes with three types of ships. There were a total of 100 container routes in the shipping network. The computational results demonstrated that the proposed model provided a useful planning tool for liner shipping companies.

10.2.2 Operational-Level Schedule Design

At the operational level, Yan et al. (2009) developed a container routing model from the perspective of a liner shipping company with the objective of maximizing operating profit while considering the arrival time of ships at ports. They performed a case study utilizing operating data from a major Taiwanese marine shipping company. Brouer et al. (2013b) proposed a vessel schedule recovery problem to evaluate a given disruption scenario and to select a recovery action balancing the tradeoff between increased bunker consumption and the impact on cargo in the remaining network and the customer service level. The model was applied to four real-life cases from Maersk Line and cost savings of up to 58 % were achieved by the suggested solutions compared to realized recoveries of the real life cases.

10.2.3 Gaps of Existing Studies

None of the above studies have taken into consideration the port time windows. In other words, they have all assumed that a port is always ready for service whenever a ship comes. This may not be consistent with the practice. Hence, the above literature review clearly shows that liner ship route schedule design with port time windows is a new research topic. It incorporates both shipping operations and port operations in the planning decision and hence has practical significance for liner shipping companies. This study thus focuses on the tactical-level schedule design problem.

10.3 Problem Description

Consider a ship route such as the NCE service in Fig. 10.1. The ship route has a weekly service frequency which means each port of call is visited on the same day every week. The port rotation of the ship route has a total of N ports. Define $I := \{1, 2, \cdots, N\}$, which is a set representing all the ports of call for simplifying the notation. Since the ports of call on a ship route form a loop, we can arbitrarily choose one port as the first port of call. For instance, if we let New York be the first port of call, the NCE service can be coded as follows: 1 (New York) → 2 (Norfolk) → 3 (Savannah) → 4 (Pusan) → 5 (Qingdao) → 6 (Ningbo) → 7 (Shanghai) → 1 (New York). If we let Norfolk be the first port of call, the NCE service can be coded as follows: 1 (Norfolk) → 2 (Savannah) → 3 (Pusan) → 4 (Qingdao) → 5 (Ningbo) → 6 (Shanghai) → 7 (New York) → 1 (Norfolk). We let p_i represent the physical port of the ith port of call, $i \in I$. We further define the voyage from the ith port to the $(i + 1)$th as leg i; leg N is the voyage from the Nth port of call to the first one. For instance, if we define New York to be the first port of call, then the first leg is the journey from New York to Norfolk, the second leg is the journey from Norfolk to Savannah, the third leg is the journey from Savannah to Pusan, the fourth leg is the journey from Pusan to Qingdao, the fifth leg is the journey from Qingdao to Ningbo, the sixth leg is the journey from Ningbo to Shanghai, the seventh leg is the journey from Shanghai to New York.

We assume that $p_i \neq p_j, i \neq j$. In other words, we assume that each physical port is visited only once during a round-trip journey. It should be noted that in reality there are many ship routes that visit a port twice in a round-trip journey, and in extreme cases, three times. The methods proposed in the chapter could be used for designing schedules for these ship routes, too, but need considerable modification, as to be discussed in Sect. 10.7. For better readability, we only consider the case that each physical port is visited only once during a round-trip journey.

10.3.1 Ship Cost, Bunker Cost and Inventory Cost

We assume that a string of m homogeneous containerships are deployed on the ship route to maintain a weekly service frequency. Ships are homogeneous means that they have the same capacity, age, designed speed, and other ship specific characteristics. In reality, two ships cannot be the same because even if they were the same when constructed, different past operating conditions would make them different (e.g., fuel efficiency). However, in mathematical modeling, it is convenient to model ships with similar characteristics as identical without losing much precision. That is why we also adopt such an approach. The highest possible sailing speed of the ships is denoted by $V^{\max}$ (knot). Represent by t_i^{port} the time (h) a ship spends at port i, and L_i (n mile) the distance of leg i. The maximum speed of containerships is usually higher than that of bulk cargo ships and tankers. This is mainly because containerships transport containerized cargos with higher unit value, and hence faster delivery is more desirable. The time a ship spends at a port consists of the time for towage, mooring and unmooring, possible wait time, and container handling. The most significant time component is container handling. For instance, if the average container handling efficiency is 100 containers/h, and a total of 2000 containers are loaded or unloaded, then the container handling time is 20 h. We assume that the container handling time is fixed. In reality, this time cannot be exactly predicted, and hence here we can consider t_i^{port} as already including some buffer time.

Ship Cost

Let v_i be the sailing speed (knot) of ships on leg i. To maintain a weekly service frequency, we have the relation:

$$\sum_{j \in I} L_j / v_j + \sum_{j \in I} t_j^{\text{port}} = 168m \tag{10.1}$$

In Eq. (10.1), the left-hand side is the round-trip journey time (h), and the right-hand side is the number of ships times 168 h/week. Equation (10.1) is the fundamental equation defining the number of ships required to maintain a weekly frequency. For instance, if the round-trip journey time is 336 h (2 weeks), two ships are needed to maintain a weekly frequency. If the round-trip journey time is 8 weeks, eight ships must be deployed to maintain a weekly frequency. If we can reduce the round-trip journey time from 8 to 7 weeks by sailing faster, skipping ports, or shortening port time, we can save one ship. Denote by C^{ship} (USD/week) the fixed operating cost of a ship, which is the ship chartering cost but does not include bunker fuel cost. Hence, the weekly operating cost of the ships deployed on the ship route is $C^{\text{ship}}m$.

Bunker Cost

As aforementioned, Eq. (10.1) implies that when the speed is higher, fewer ships need to be deployed to maintain the same weekly service frequency. However, a higher speed implies a larger amount of bunker consumed. To take into consideration the bunker cost, we let $g_i(v_i)$ (t/n mile) be the bunker consumption function at the speed v_i on leg i. Based on the results in existing studies (Psaraftis and Kontovas 2010; Kontovas and Psaraftis 2011; Ronen 2011; Wang and Meng 2012d; Psaraftis and Kontovas 2013), we assume that $g_i(v_i)$ is a power function of the form:

$$g_i(v_i) = a_i (v_i)^{b_i}, i \in I \tag{10.2}$$

where a_i and b_i are two coefficients calibrated from historical operating data and satisfy $a_i > 0$ and $b_i > 1$. Denote by α (USD/t) the bunker fuel price. The weekly bunker cost is $\alpha \sum_{i \in I} L_i g_i(v_i) = \alpha \sum_{i \in I} L_i a_i (v_i)^{b_i}$. It should be noted that although we assume that the bunker consumption function has the form of Eq. (10.2), the solution method that will be elaborated later is applicable to other forms of bunker consumption functions, too.

Inventory Cost

Besides the ship cost and bunker cost, the inventory cost of containers should also be incorporated. In fact, a lower speed (slow-steaming) would increase the transit time of containers, and thereby the inventory cost. We let $\bar{V}_i$ be the number of containers (twenty-foot equivalent units, or TEUs) on leg i, and β be the unit inventory cost (USD per TEU per h). Since the time spent at each port is constant, we only consider the inventory cost associated with sailing time at sea (sea time). Therefore, the total inventory cost is $\sum_{i \in I} \beta \bar{V}_i L_i / v_i$. It should be noted that $\bar{V}_i$ is actually a predicted value based on historical data. The inventory cost is included to reflect the quality of the liner shipping company's transport services[1]. Note further that β is also predicted and our model allows β to vary with different voyage legs.

10.3.2 Liner Ship Route Schedule

We define the time 00:00 of a certain Sunday as time 0 (h), and hence 10:00 on Monday is time $24 + 10 = 34$, and 10:00 next Tuesday is time $168 + 24 * 2 + 10 = 226$. Since we assume that the port time (t_i^{port}) is fixed, the time of departure (t_i^{dep}) at port i is determined by the time of arrival (t_i^{arr}) and the port time (t_i^{port}), that is:

$$t_i^{\text{dep}} = t_i^{\text{arr}} + t_i^{\text{port}}, i \in I \tag{10.3}$$

Because of the weekly service frequency, without loss of generality, we let

$$0 \leq t_1^{\text{arr}} < 168 \tag{10.4}$$

Note that the above equation is important to eliminate symmetric solutions. Because of the weekly frequency, there is no difference whether the first port of call is visited at time 20 (i.e., $t_1^{\text{arr}} = 20$) or $20 + 168$ (i.e., $t_1^{\text{arr}} = 188$). Hence, we only need to consider the case where the arrival at the first port of call is between time 0 and 168.

We define the time when the ship returns to the first port of call as t_{N+1}^{arr}, that is:

$$t_{N+1}^{\text{arr}} := t_1^{\text{arr}} + 168m \tag{10.5}$$

This equation implies that a ship needs $t_{N+1}^{\text{arr}} - t_1^{\text{arr}} = 168m$ hours to complete a round-trip journey. This is consistent with the weekly frequency.

The *schedule* of a liner ship route is the vector defined below:

$$(t_i^{\text{arr}}, i \in I; m) \tag{10.6}$$

We stress that the schedule of a liner ship route cannot be represented by $(t_i^{\text{arr}}, i \in I)$. This is because, given $(t_i^{\text{arr}}, i \in I)$, we do not know the inter-arrival time from the last port of call to the first. The number of ships m together with $(t_i^{\text{arr}}, i \in I)$ can define the inter-arrival time from the last port of call to the first. Of course, the schedule can also be uniquely determined by $(t_i^{\text{arr}}, i \in I; t_{N+1}^{\text{arr}})$.

---

[1] In reality the liner shipping company will not pay the customers for their inventory cost.

10.3.3 Port Time Window

A ship cannot arrive at a port at any time because the port may be busy during some periods of a week. Hence, we let $\Omega_i \subseteq [0, 168)$ be the time in a week during which port i is available for serving ships on the ship route, i.e., port time window. For example, $\Omega_i = [10, 20] \cup [96, 120]$ means that port i is available from 10:00 Sunday to 20:00 Sunday, and 00:00 Thursday to 00:00 Friday. $\Omega_j = [0, 24] \cup [144, 168)$ means that port j is available from 00:00 Sunday to 00:00 Monday, and 00:00 Saturday to 00:00 Sunday. In other words, the port is available from 00:00 Saturday to 00:00 Monday next week.

We assume that the port time window at each port (which corresponds to each port of call because we assume that each port is visited once in a round-trip journey) is known. In reality, a liner shipping company can obtain this port time window from port operators, because port operators have to tell it whether it is possible to arrive at a particular time.

A ship needs to stay at port i for t_i^{port} hours. Therefore, we could easily compute the feasible arrival times at port i based on Ω_i. For instance, $\Omega_i = [10, 20] \cup [96, 120]$ and $t_i^{\text{port}} = 5$ imply that t_i^{arr} could be any value in $[10, 15] \cup [96, 115]$. We let $\hat{\Omega}_i \subseteq [0, 168)$ be the set of feasible arrival times at port i in a week. It should be mentioned that because of the weekly service frequency, when $\hat{\Omega}_i = [10, 15] \cup [96, 115]$, the arrival time $t_i^{\text{arr}} = 180$ (which corresponds to time 12 of the next week) is also feasible. In fact, an arrival time is feasible if and only if $\left(t_i^{\text{arr}} \bmod 168\right) \in \hat{\Omega}_i$, where the "mod" operator obtains the modulus of two integer numbers.

Therefore, the ship route schedule design problem with port time window aims to determine the optimal arrival time at each port of call on a ship route that satisfies the port time window to minimize the total cost including ship cost, bunker cost, and inventory cost.

10.4 Mathematical Model

10.4.1 Notation

Before presenting the model, we list the notation below.

Variables

m	Number of ships deployed on the ship route
t_i^{arr}	Arrival time (h) at the ith port of call
t_{N+1}^{arr}	The time (h) when the ship returns to the 1st port of call
t_i^{dep}	Departure time (h) from the ith port of call
v_i	Sailing speed (knot) on leg i

Parameters

10.4.2 Model

The SDPTW can be formulated as:
[SDPTW]

$$\min C^{\text{ship}} m + \alpha \sum_{i \in I} L_i g_i(v_i) + \sum_{i \in I} \beta \bar{V}_i \frac{L_i}{v_i} \tag{10.7}$$

α	The bunker fuel price (USD/ton)
β	The unit inventory cost (USD per TEU per h)
$\hat{\Omega}_i$	The set of feasible arrival times at the ith port of call
C^{ship}	The weekly operating cost of a ship (USD/week)
$g_i(v_i)$	Bunker consumption per nautical mile at the speed v_i on leg i (tons/n mile)
I	Set of legs, $I = \{1, 2, \cdots, N\}$
L_i	Oceanic distance (n mile) of the leg i
N	Number of ports on the ship route
p_i	The port i on the ship route
t_i^{port}	Time (h) a ship spends at port i
$\bar{V}_i$	Number of containers (TEUs) on leg i
$V^{\max}$	Maximum speed of the ships (knot)

subject to:

$$\sum_{j \in I} L_j / v_j + \sum_{j \in I} t_j^{\text{port}} = 168m \tag{10.8}$$

$$t_i^{\text{dep}} = t_i^{\text{arr}} + t_i^{\text{port}}, i \in I \tag{10.9}$$

$$0 \leq t_1^{\text{arr}} < 168 \tag{10.10}$$

$$t_{N+1}^{\text{arr}} = t_1^{\text{arr}} + 168m \tag{10.11}$$

$$v_i = \frac{L_i}{t_{i+1}^{\text{arr}} - t_i^{\text{dep}}}, i \in I \tag{10.12}$$

$$\left(t_i^{\text{arr}} \mod 168\right) \in \hat{\Omega}_i, i \in I \tag{10.13}$$

$$0 \leq v_i \leq V^{\max}, i \in I \tag{10.14}$$

$$m \in \{1, 2, 3, \cdots\} \tag{10.15}$$

The objective function (10.7) minimizes the sum of ship cost, bunker cost, and inventory cost. The first term is the ship cost, which is proportional to the number of ships deployed. The second term is the bunker cost, which varies nonlinearly with speed. The third term is inventory cost, which is summed over all legs. Constraint (10.8) requires that the service on this ship route is weekly. Constraint (10.9) defines the departure time from each port of call. Constraint (10.10) eliminates symmetric solutions. Constraint (10.11) defines the time when the ship returns to the first port of call after one round trip. Constraint (10.12) calculates the sailing speed on each leg. Constraint (10.13) imposes the port time window restrictions. Constraint (10.14) enforces the lower and upper limits on the sailing speed. Constraint (10.15) indicates that the number of ships is a positive integer.

10.5 Solution Method

The model [SDPTW] is a mixed-integer nonlinear non-convex optimization problem. It is difficult to solve because (i) it has both continuous (sailing speed) and discrete variables (number of ships); (ii) it has nonlinear objective function (10.7) and nonlinear constraints (10.8) and (10.12); (iii) the set $\hat{\Omega}_i$ in Eq. (10.13) may consist of disjoint intervals, as shown in Fig. 10.2. This will lead to a non-convex domain even without considering the discrete decision variables; moreover, even if $\hat{\Omega}_i$ is convex, the "mod" operator still leads to a non-convex domain. These difficulties make the model challenging and hard to be solved by existing commercial solvers. To address the model, we have to develop our own solution algorithm.

We notice that the optimal arrival time at a port of call almost only depends on the arrival time at its previous port of call, and has little to do with the arrival times at even further previous ports of call. Such a property enlightens us to develop a dynamic programming based solution method that solves the problem to optimality.

10.5.1 Space-Time Network for a Given Number of Ships

Given the number of ships m, say, $\bar{m}$, the ship cost $C^{\text{ship}}\bar{m}$ is fixed. Moreover, the round-trip journey time is also fixed, that is, $168\bar{m}$ hours. The model can be reformulated as
[SDPTW-$\bar{m}$]

$$\min \alpha \sum_{i \in I} L_i g_i(v_i) + \sum_{i \in I} \beta \bar{V}_i \frac{L_i}{v_i} \tag{10.16}$$

subject to:

$$\sum_{j \in I} L_j / v_j + \sum_{j \in I} t_j^{\text{port}} = 168\bar{m} \tag{10.17}$$

$$t_i^{\text{dep}} = t_i^{\text{arr}} + t_i^{\text{port}}, i \in I \tag{10.18}$$

$$0 \le t_1^{\text{arr}} < 168 \tag{10.19}$$

$$t_{N+1}^{\text{arr}} = t_1^{\text{arr}} + 168\bar{m} \tag{10.20}$$

$$v_i = \frac{L_i}{t_{i+1}^{\text{arr}} - t_i^{\text{dep}}}, i \in I \tag{10.21}$$

$$\left(t_i^{\mathrm{arr}} \mod 168\right) \in \hat{\Omega}_i, i \in I \tag{10.22}$$

$$0 \le v_i \le V^{\max}, i \in I \tag{10.23}$$

Note that model [SDPTW-$\bar{m}$] no longer has discrete variables.

Property of the Problem

As t_1^{arr} is between 0 and 168, we can discretize it and enumerate all possible discretized values. Given m and t_1^{arr} (say, $\bar{m}$ and $\bar{t}_1^{\mathrm{arr}}$), if the arrival time at a particular port of call is known, then the bunker cost and inventory cost associated with the voyage legs after the port of call depend only on its arrival time and are independent of the arrival times at ports of call prior to it. For instance, if we know that the arrival time at the $\bar{i}$th port of call is $\bar{t}_{\bar{i}}^{\mathrm{arr}}$, the problem can be split into two subproblem: subproblem 1 determines the arrival time at each port of call 2, 3, $\cdots$, $\bar{i} - 1$; subproblem two determines the arrival time at each port of call $\bar{i} + 1$, $\bar{i} + 2$, $\cdots$, N. To be clear, we formulate the two subproblems below:

[SDPTW-$\bar{m}$-subproblem 1]

$$\min \alpha \sum_{i=1}^{\bar{i}-1} L_i g_i(v_i) + \sum_{i=1}^{\bar{i}-1} \beta \bar{V}_i \frac{L_i}{v_i} \tag{10.24}$$

subject to:

$$t_i^{\mathrm{dep}} = t_i^{\mathrm{arr}} + t_i^{\mathrm{port}}, i = 1, 2, 3, \cdots, \bar{i} - 1 \tag{10.25}$$

$$v_i = \frac{L_i}{t_{i+1}^{\mathrm{arr}} - t_i^{\mathrm{dep}}}, i = 1, 2, 3, \cdots, \bar{i} - 1 \tag{10.26}$$

$$\left(t_i^{\mathrm{arr}} \mod 168\right) \in \hat{\Omega}_i, i = 2, 3, \cdots, \bar{i} - 1 \tag{10.27}$$

$$0 \le v_i \le V^{\max}, i = 1, 2, 3, \cdots, \bar{i} - 1 \tag{10.28}$$

$$t_1^{\mathrm{arr}} = \bar{t}_1^{\mathrm{arr}} \tag{10.29}$$

$$t_{\bar{i}}^{\mathrm{arr}} = \bar{t}_{\bar{i}}^{\mathrm{arr}} \tag{10.30}$$

[SDPTW-$\bar{m}$-subproblem 2]

$$\min \alpha \sum_{i=\bar{i}}^{N} L_i g_i(v_i) + \sum_{i=\bar{i}}^{N} \beta \bar{V}_i \frac{L_i}{v_i} \tag{10.31}$$

subject to:

$$t_i^{\mathrm{dep}} = t_i^{\mathrm{arr}} + t_i^{\mathrm{port}}, i = \bar{i}, \bar{i} + 1, \cdots, N \tag{10.32}$$

$$v_i = \frac{L_i}{t_{i+1}^{\mathrm{arr}} - t_i^{\mathrm{dep}}}, i = \bar{i}, \bar{i} + 1, \cdots, N \tag{10.33}$$

$$\left(t_i^{\text{arr}} \mod 168\right) \in \hat{\Omega}_i, i = \bar{i}+1, \bar{i}+2, \cdots, N \tag{10.34}$$

$$0 \le v_i \le V^{\max}, i = \bar{i}, \bar{i}+1, \cdots, N \tag{10.35}$$

$$t_{\bar{i}}^{\text{arr}} = \bar{t}_{\bar{i}}^{\text{arr}} \tag{10.36}$$

$$t_{N+1}^{\text{arr}} = \bar{t}_1^{\text{arr}} + 168\bar{m} \tag{10.37}$$

Hence, the decisions about the arrival time at each port of call could be made in a sequential manner, that is, the optimal arrival time at the next port only depends on the arrival time at the current port (and of course $\bar{m}$ and $\bar{t}_1^{\text{arr}}$). Exploiting this property, we construct a space-time network and thereby develop a dynamic programming based solution approach.

Space-Time Network Construction Method

To construct a space-time network, in view of Eq. (10.10), we only need to consider a time horizon of $168(m+1)$ hours [2]. In other words, the time horizon is $m+1$ weeks. We discretize the time horizon into intervals, the length of each interval being 1 h [3]. To take into account the voyage from the Nth port of call to the first one, we consider $N+1$ ports in the space-axis, where the $(N+1)$th port corresponds to the returning to the first one. Each of the $N+1$ ports is copied $168(m+1)$ times. Hence, each node (t, i) in the space-time network corresponds to a port i at a particular time t. We define the time t as the arrival time at the port i. Therefore, node (t, i) in the space-time network means that port i is visited at time t. For each port i, if $(t \mod 168) \notin \hat{\Omega}_i$, then the port is busy at the time t. Hence, it is impossible to visit port i at time t. Consequently, for each port i, if $(t \mod 168) \notin \hat{\Omega}_i$, then we remove the node (or mark it as inactive as it will not be visited).

Moreover, from each active node (t, i), the ship may visit any active node $(t', i+1)$ satisfying

$$t' \ge t + t_i^{\text{port}} + \frac{L_i}{V^{\max}} \tag{10.38}$$

In words, from port i, a ship can only visit port $i+1$ and the sailing speed cannot exceed $V^{\max}$. Of course, the port time window at port $i+1$ is already implicitly considered by removing the nodes that cannot be visited.

We formally state the method for constructing the space-time network below:

Algorithm 1: Construction of space-time network $G(m)$[4] for a given ship number m

Step1. (Construct nodes): Construct a space-time network with the horizontal axis being the time (hours, starting from 0 which represents 00:00 of a particular

---

[2] If, for example, $t_1^{\text{arr}} = 167$, then the ship will return to the first port of call at time $168m + 167$. Therefore, the time horizon is $168(m+1)$ hours rather than $168m$ hours.

[3] The precision of 1 h is more than sufficient for liner shipping applications.

[4] G means "graph".

Sunday), and the vertical axis being the space (ports). The length of the time axis is $168(m+1)$ with the discrete time points being $0, 1, 2, \cdots, 168(m+1)-1$. The vertical axis has $N + 1$ ports, that is, the 1st port of call, the 2nd port of call, $\cdots$, the Nth port of call, and the $(N + 1)$th port of call. Note that the $(N + 1)$th port of call actually represents that the ship returns to the first port of call after a round-trip journey of $168m$ hours. Each of the $N + 1$ ports is copied $168(m + 1)$ times. Now, in the space time network, there are $168(m + 1)(N + 1)$ nodes. A node can be represented by an ordered pair (time unit, port ID), or (t, i), which means that port i is visited at time t.

Step 2. (Deactivate nodes):

Step 2.1 (Deactivate nodes that violate port time windows) For each node (t, i) in the space-time network, if $(t \bmod 168) \notin \hat{\Omega}_i$, the ship cannot visit the node and hence we mark it as inactive;

Step 2.2 (Deactivate nodes that violate Eq. (10.10)) For each node $(t, 1)$ that corresponds to the first port of call in the space-time network, if $t \geq 168$, the ship cannot visit the node and hence we mark it as inactive;

Step 2.3 (Deactivate nodes that violate Eq. (10.11)) For each node $(t, N + 1)$ that corresponds to the return to the first port of call in the space-time network, if $t \leq 168m - 1$, the ship cannot visit the node and hence we mark it as inactive (note that here the number of ships m is given).

Step 3. (Construct arcs):

Step 3.0 Set $i = 0$;

Step 3.1 Set $i := i+1$. For each active node (t, i), $t \in \{0, 1, 2, \cdots, 168(m+1)-1\}$, construct an arc from it to any of the active nodes $(t', i + 1)$ satisfying $t' \geq t+t_i^{\text{port}} + \frac{L_i}{V^{\max}}$. Hence, the sailing time of the arc is $t'-t-t_i^{\text{port}}$. Moreover, the sailing speed is also determined, which is $v_i = L_i/(t' - t - t_i^{\text{port}})$. Therefore, the corresponding bunker cost is $\alpha L_i g_i(v_i)$ and the inventory cost of the containers is $(t' - t - t_i^{\text{port}})\beta \bar{V}_i$ (as aforementioned, the inventory cost associated with port time is constant, and hence is not modeled). The cost (sum of bunker and inventory cost) of the arc is $\alpha L_i g_i(L_i/(t' - t - t_i^{\text{port}})) + (t' - t - t_i^{\text{port}})\beta \bar{V}_i$.

Step 3.2 If $i = N$, Stop. Otherwise, go to Step 3.1. □

An Example of Space-Time Network Construction

We use an example to demonstrate the space-time network construction method. For the ease of presentation, we use "day" rather than "hour" in the discretization. That is, 0 represents Sunday, 1 represents Monday, etc. If we do not use days but use hours, there would be too many nodes in the space-time network and it would be difficult to understand it. Suppose that there are three ports of call on the ship route. The feasible arrival days are $\hat{\Omega}_1 = \{2, 3, 6\}$, $\hat{\Omega}_2 = \{0, 1, 5, 6\}$, and $\hat{\Omega}_3 = \{4, 5\}$. In addition, suppose that $t_1^{\text{port}} + \frac{L_1}{V^{\max}} = 4$ days, $t_2^{\text{port}} + \frac{L_2}{V^{\max}} = 5$ days, and $t_3^{\text{port}} + \frac{L_3}{V^{\max}} = 1$ day.

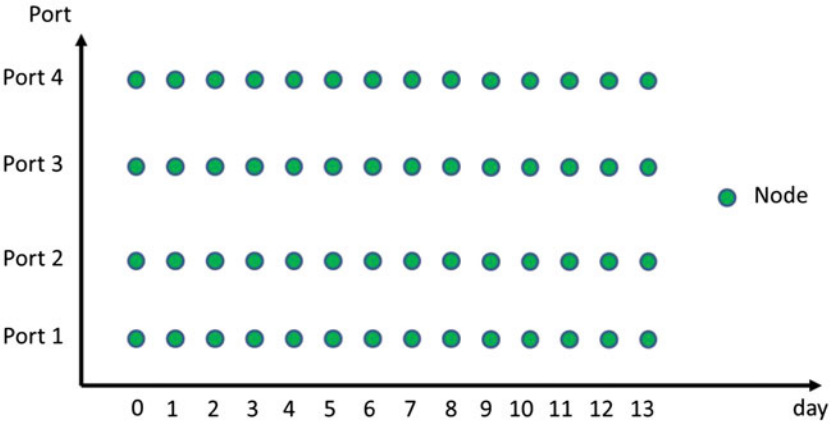


Fig. 10.4 Construct nodes

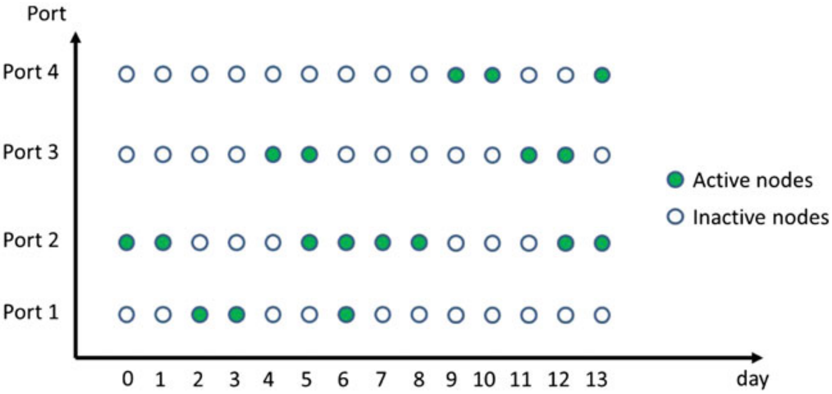


Fig. 10.5 Deactivate nodes

The number of ships $m = 2$. The three steps in Algorithm 1 are shown in Fig. 10.4, 10.5 and 10.6, respectively.

Let us look at Fig. 10.4 first. As $m = 2$ and we use "day" to discretize the time, there are a total of 14 days in the time axis. Hence, each port should be copied 14 times. Since there are three ports of call on the ship route, considering the loop property of the ship route, we need to consider 4 ports, where the fourth port is actually the return to the first port. As a result, there are a total of $4 \times 14 = 56$ nodes in the space-time network.

In Fig. 10.5, we deactivate nodes. In step 2.1, since $\hat{\Omega}_2 = \{0, 1, 5, 6\}$, nodes corresponding to port 2 are active only if $t = 0, 1, 5, 6, 7, 8, 12, 13$. In other words, only 8 nodes corresponding to port 2 are active. Since $\hat{\Omega}_3 = \{4, 5\}$, nodes corresponding to port 3 are active only if $t = 4, 5, 11, 12$. In other words, only 4 nodes corresponding to port 3 are active. Since $\hat{\Omega}_1 = \{2, 3, 6\}$ and port 1 must be visited in the first week,

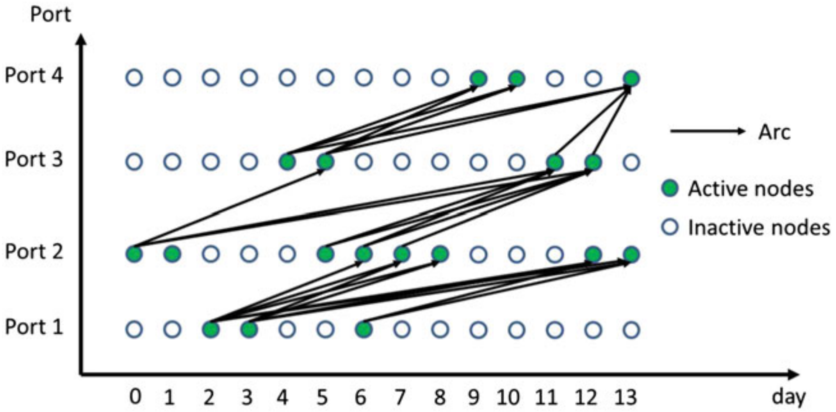


Fig. 10.6 Construct arcs

nodes corresponding to port 3 are active only if $t = 2, 3, 6$. In other words, only 3 nodes corresponding to port 1 are active. Since $\hat{\Omega}_1 = \{2, 3, 6\}$ and port of call 4 (which is the same port as port of call 1) must be visited in the second week (as $m = 2$), nodes corresponding to port 4 are active only if $t = 9, 10, 13$. In other words, only 3 nodes corresponding to port 4 are active.

In Fig. 10.6, we add arcs connecting the nodes. Note that the arcs connect only active nodes, and must respect the maximum speed of ships. For instance, both nodes $(2, 1)$ and $(5, 2)$ are active. However, their time difference $5 - 2 = 3$ is smaller than $t_1^{\text{port}} + \frac{L_1}{V^{\max}} = 4$. Hence, ships cannot visit node $(5, 2)$ from node $(2, 1)$.

Loop Property of Ship Route in the Space-Time Network

It should be noted that in the space-time network, if a ship visits port 1 at time t_1^{arr}, it must return to port 1 at time $t_1^{\text{arr}} + 168m$. This constraint poses difficulties for finding the schedule with the minimum cost. Nevertheless, we identify that the total number of possible t_1^{arr} is at most 168. Therefore, we could enumerate all possible t_1^{arr}. For each fixed t_1^{arr}, we can apply the dynamic programming approach to find the shortest path (minimum-cost path) from node $(t_1^{\text{arr}}, 1)$ to node $(t_1^{\text{arr}} + 168m, N + 1)$, denoted by $c(m, t_1^{\text{arr}})$[5]. Hence, the minimum total cost with given m is $C^{\text{ship}}m + \min_{t_1^{\text{arr}} \in \{0,1,2,\cdots,167\}} c(m, t_1^{\text{arr}})$.

---

[5] $c(m, t_1^{\text{arr}}) = \infty$ if $(t_1^{\text{arr}}, 1)$ is inactive or if there is no path from node $(t_1^{\text{arr}}, 1)$ to node $(t_1^{\text{arr}} + 168m, N + 1)$.

10.5.2 Lower Bound of the Number of Ships

The previous sub-section provides an approach for finding the optimal schedule with a given m. However, as m is a positive integer, we cannot enumerate all possible values of m. To overcome this difficult, we investigate how to confine the range of possible values of m.

According to Eq. (10.1), the minimum number of ships can be computed by:

$$m^{\min} = \left\lceil \left(\sum_{j \in I} L_j / V^{\max} + \sum_{j \in I} t_j^{\text{port}} \right) / 168 \right\rceil \tag{10.39}$$

where $\lceil x \rceil$ is the smallest integer greater than or equal to x.

10.5.3 Lower Bound of the Total Cost With Given Number of Ships

When the number of ships is m, a lower bound on the total cost, denoted by $LB(m)$, can be computed as follows. As the ship cost in Eq. (10.7) is fixed, we minimize the sum of bunker cost and inventory cost by optimizing the speed. To facilitate the computation of the lower bound, we relax relevant constraints and only require that the speed is nonnegative. Using the bunker consumption function (10.2), we have:

$$\min_{v_i} \sum_{i \in I} \alpha L_i a_i (v_i)^{b_i} + \sum_{i \in I} \beta \bar{V}_i \frac{L_i}{v_i} \tag{10.40}$$

subject to:

$$-v_i \le 0, i \in I \tag{10.41}$$

It is easy to see that the speed on different legs can be optimized independently. Let $\lambda_i \ge 0$ be the Lagrangian multiplier associated with constraint $-v_i \le 0$. The Karush-Kuhn-Tucker (KKT) condition of the above optimization problem is:

$$\alpha L_i a_i b_i (v_i)^{b_i - 1} - \beta \bar{V}_i L_i \frac{1}{(v_i)^2} - \lambda_i = 0 \tag{10.42}$$

$$\lambda_i(-v_i) = 0 \tag{10.43}$$

$$-v_i \le 0 \tag{10.44}$$

$$\lambda_i \ge 0 \tag{10.45}$$

Apparently $-v_i < 0$, and therefore $\lambda_i = 0$. Hence, we can compute the optimal speed in the model, denoted by $\tilde{v}_i$:

$$\tilde{v}_i = \left(\frac{\beta \bar{V}_i}{\alpha a_i b_i} \right)^{\frac{1}{b_i + 1}} \tag{10.46}$$

Consequently, a lower bound of the total cost with m ships is:

$$LB(m) = C^{\text{ship}} m + \alpha \sum_{i \in I} L_i g_i(\tilde{v}_i) + \sum_{i \in I} \beta \bar{V}_i \frac{L_i}{\tilde{v}_i} \tag{10.47}$$

10.5.4 Overall Algorithm

Sub-section 10.5.1 develops a space-time network model that can find the optimal schedule for a given number of ships using dynamic programming approach. Sub-section 10.5.2 obtains a lower bound on the number of ships that are needed. Sub-section 10.5.3 proposes a lower bound on the total cost for a given number of ships, and this lower bound increases with m as shown in Eq. (10.47). Based on these results, we now present the overall solution algorithm:

Algorithm 2: Solution method for the SDPTW

Step 0. Set $m = m^{\text{min}} - 1$. Denoted by $C^* := \infty$ the minimum total cost obtained (upper bound).

Step 1. Set $m := m + 1$. If $LB(m) \geq C^*$, we have obtained the optimal solution and hence stop. Otherwise, construct the space-time network $G(m)$.

Step 2. For each $t_1^{\text{arr}} \in \{0, 1, 2, \cdots, 167\}$, find the shortest path from node $(t_1^{\text{arr}}, 1)$ to node $(t_1^{\text{arr}} + 168m, N + 1)$ and its cost $c(m, t_1^{\text{arr}})$. If $C^{\text{ship}} m + c(m, t_1^{\text{arr}}) < C^*$, set $C^* := C^{\text{ship}} m + c(m, t_1^{\text{arr}})$ and record the current solution. When all the t_1^{arr} have been examined, go to Step 1. □

Algorithm 2 terminates in a finite number of iterations. This is because once a finite upper bound C^* is found, the algorithm will stop before or when $m = \lceil C^* / C^{\text{ship}} \rceil$.

10.6 Case Study

We choose a case study of the NCE ship route in Fig. 10.1 to evaluate the proposed model and solution method. We assume that 5000-TEU ships are deployed on it. We choose 5000-TEU ships because larger ships cannot transit the Panama Canal. The operating cost C^{ship} =500,000 USD/week, the maximum speed $V^{\max}$=30 knots, the bunker price $\alpha = 400$ USD/t and the unit inventory cost $\beta = 1$ USD per TEU per hour. The port time (h), distance (n mile), bunker consumption function $g_i(v_i)$, and volume of containers on each leg (TEUs) are shown in Table 10.1. In Table 10.1 we assume that the port time is either 1 day or 1.5 days, the bunker consumption functions may be different for different legs, and the number of containers on each leg implies that the ship load factor is between $\frac{2200}{5000} = 44\%$ and $\frac{4500}{5000} = 90\%$. The port time window at each port, i.e., Ω_i, is shown in Table 10.2, which indicates that no port is available seven days a week.

Table 10.1 Parameters in the case study

ID	Port	Port time	Distance	Bunker function	# containers
1	New York	36	261	$0.001(v_1)^2$	2200
2	Norfolk	24	436	$0.001(v_2)^{2.1}$	3000
3	Savannah	24	9678	$0.001(v_3)^{2.3}$	3500
4	Pusan	24	467	$0.001(v_4)^2$	4200
5	Qingdao	24	386	$0.001(v_5)^2$	4000
6	Ningbo	24	101	$0.001(v_3)^2$	4300
7	Shanghai	36	10553	$0.001(v_7)^2$	4500

Table 10.2 Port time windows

ID	Port	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
1	New York	Free	Free	Busy	Busy	Free	Free	Busy
2	Norfolk	Busy	Busy	Free	Busy	Busy	Busy	Busy
3	Savannah	Busy	Busy	Busy	Free	Busy	Busy	Free
4	Pusan	Free	Busy	Free	Busy	Busy	Busy	Busy
5	Qingdao	Busy	Busy	Busy	Busy	Busy	Busy	Free
6	Ningbo	Busy	Free	Busy	Busy	Free	Busy	Busy
7	Shanghai	Free	Free	Busy	Free	Free	Busy	Busy

10.6.1 Impact of Port Time Windows

Firstly, we examine the effect of port time windows on the total cost and the optimal schedule. We assume that currently the port of Norfolk is only available on Tuesday, as shown in Table 10.2. Both Norfolk and the liner shipping company are interested in looking at the result if more available time is provided at Norfolk. We hence examine the cases of 1 day available for service each week (Tuesday), 2 days (plus Friday), 3 days (plus Monday), 4 days (plus Saturday), 5 days (plus Thursday), 6 days (plus Sunday), and 7 days (which means that Norfolk is ready to serve ships at any time). The results of the total cost and the optimal number of ships deployed are shown in Fig. 10.7.

It can be seen that more available days at Norfolk leads to a lower total cost: when the number of available days is increased from 2 to 6, the total cost is reduced by 214,639 USD per week. Figure 10.7 also demonstrates that the number of available days at a port may affect the optimal number of ships deployed. The optimal ship schedule, i.e., arrival time at each port of call, is shown in Table 10.3, where e.g. “Cases 1, 2” means that Norfolk is available only 1 or 2 days in a week. We observe that when the availability of Norfolk is changed, the optimal arrival times at it and its neighboring ports may also change. However, there is no impact on the optimal arrival times at ports that are a few voyage legs away from Norfolk.

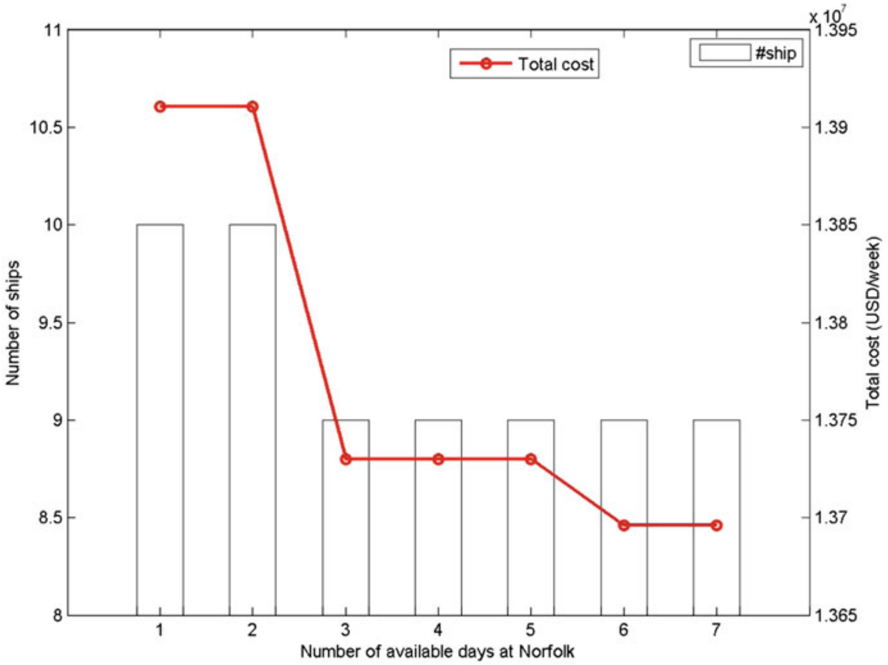


Fig. 10.7 Impact of port time windows on the total cost and the number of ships

Table 10.3 Impact of port time windows on the optimal schedule

ID	Port	Cases 1, 2	Cases 3, 4, 5	Cases 6, 7
1	New York	0	108	108
2	Norfolk	48	192	175
3	Savannah	144	240	240
4	Pusan	888	888	888
5	Qingdao	984	984	984
6	Ningbo	1032	1032	1032
7	Shanghai	1080	1080	1080
1	New York	1680	1620	1620

10.6.2 Consequence of Port Efficiency

The port time t_i^{port} to a large extent depends on the container handling efficiency. Therefore, port operators seek to improve efficiency by optimizing quay-side and yard-side operations. To investigate the effect of port handling efficiency, we change the port time at Shanghai from 12 h, 18 h, 24 h, 30 h, to 36 h, and compute the optimal solution. We find that the optimal number of ships is always 10. The total

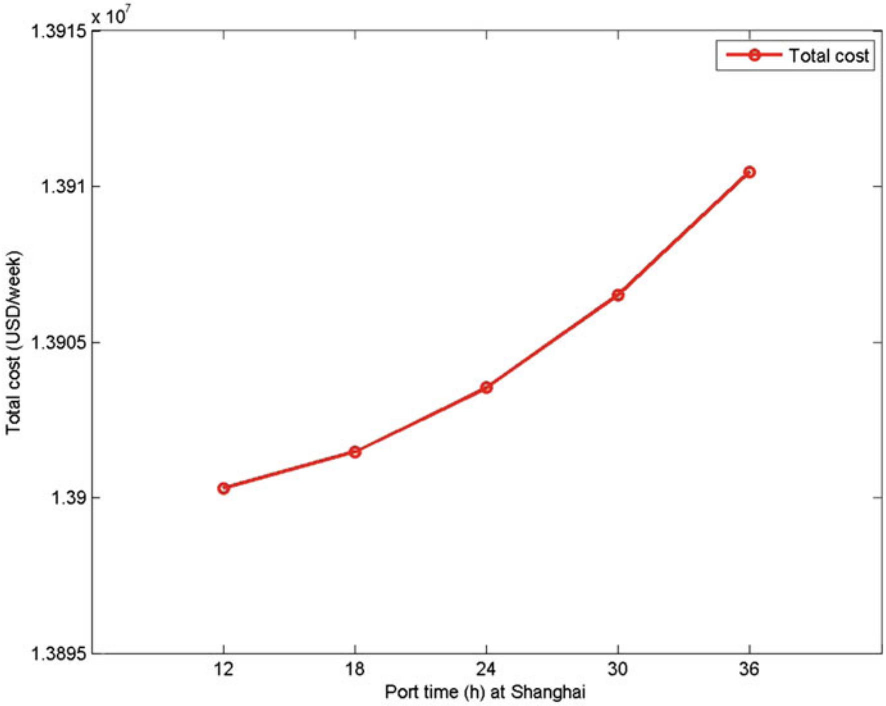


Fig. 10.8 Impact of port time at Shanghai on the total cost

cost increases with the time spent at Shanghai, as shown in Fig. 10.8. In fact, a ship creates value when it is moving cargo, whereas standing still at ports does not create value. Moreover, when the number of ships is given, a longer port time means a shorter sailing time, which leads to higher bunker consumption. Therefore, improving port efficiency will reduce the total cost for liner shipping companies.

We then fix the port time at Shanghai at 36 h, and change the port time at New York from 12 h, 18 h, 24 h, 30 h, to 36 h, and compute the optimal solution. The result is shown in Fig. 10.9. It clearly shows that when the port time is increased, not only the total cost increases, but also the optimal number of ships to deploy may increase.

10.6.3 Result of Bunker Prices

The bunker price is volatile and hence we examine the sensitivity of the solution with different bunker prices from 300, 400, 500, 600, 700, to 800 USD/t. The result is shown in Fig. 10.10. We observe that the total cost increases almost linearly (not strictly linearly) with the bunker price. Consequently, a higher bunker price

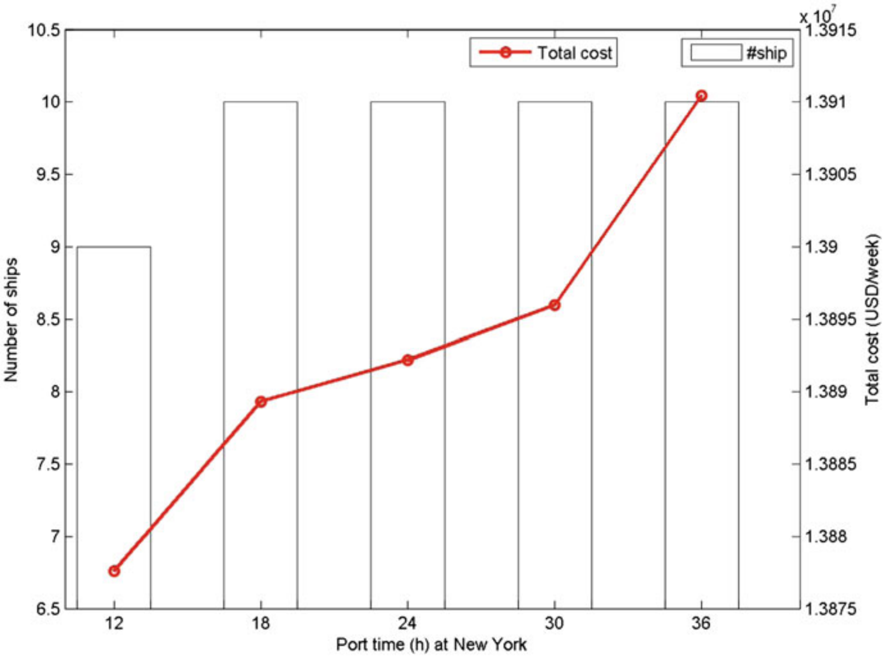


Fig. 10.9 Impact of port time at New York on the total cost and the number of ships

always leads to a higher cost for liner shipping companies. In addition, Fig. 10.10 clearly shows that there is a rise in the number of vessels used when the bunker price becomes higher. This is because when more ships are deployed, the sailing speed can be reduced, resulting in a lower bunker consumption. A reduction in bunker consumption is more significant when the bunker price is higher.

10.6.4 Effect of Inventory Cost

Finally, we investigate the effect of the unit inventory cost β on the total cost and the optimal number of ships to deploy by changing β from 1, 1.25 through to 2. The result is shown in Fig. 10.11, which indicates that the rise of unit inventory cost leads to a decreasing in the number of ships and an increasing of the total cost. This is because when the unit inventory cost is higher, containerships have to sail at a higher speed to shorten the transit time. Therefore, the number of ships is reduced. At the same time, the total cost inevitably becomes higher.

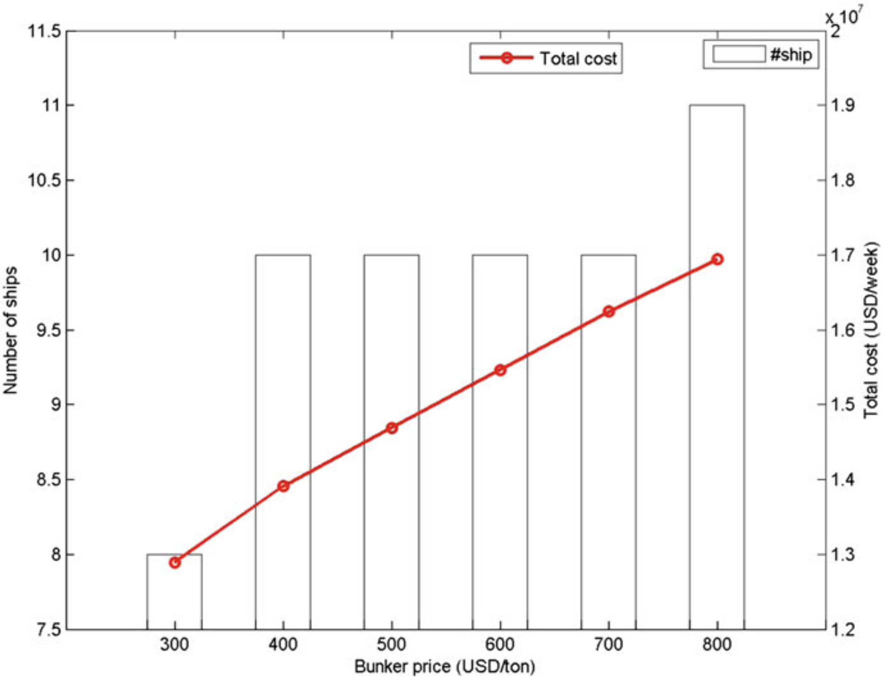


Fig. 10.10 Result of bunker prices on the total cost and the number of ships

10.7 Conclusions and Future Research

10.7.1 Summary of the Works

This research has studied the practical liner ship route schedule design problem with port time windows. This is a significant tactical planning decision problem because it considers the availability of ports when planning liner shipping services. As a result, the designed schedule can be applied in practice without or with only minimum revisions. This problem is formulated as a nonlinear non-convex optimization model. In view of the problem structure, we have developed an efficient dynamic-programming based holistic solution approach, which includes a space-time network model and a bounding technique for the total cost with give number of ships.

The proposed solution method is applied to the NCE service provided by OOCL. The results demonstrate that the port time windows, port handling efficiency, bunker price and unit inventory cost all affect the total cost, the optimal number of ships to deploy, and also the optimal schedule. A higher availability at ports, shorter port time, lower bunker price and larger unit inventory cost result in a lower total cost. Moreover, shorter port time, lower bunker price and smaller unit inventory cost lead to a smaller number of ships to deploy. Therefore, port operators can apply the proposed

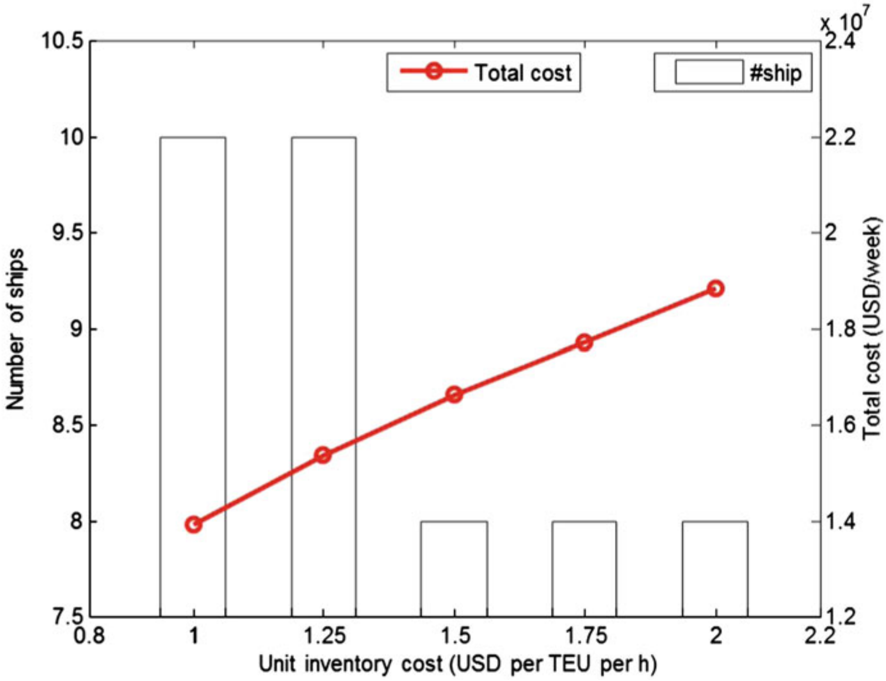


Fig. 10.11 Effect of unit inventory cost on the total cost and the number of ships

method to quantify the benefits to their customers, i.e., liner shipping companies, gained by expanding the ports' capacity and improving the ports' efficiency. Liner shipping companies may need to charter in more ships if they predict that the future bunker price will increase, or if they predict that a particular season is coming during which the value of the cargo is generally low.

10.7.2 Future Research Directions

Ship Route with Ports Visited More than Once

In this study we have assumed that each port of call is visited only once in a round-trip journey time. If some ports are visited twice in a round-trip journey, i.e., some ports of call correspond to the same physical port, then the port time window should be dealt with more carefully. First, in the dynamic programming approach, when we analyze the second arrival time at a port, we have to take into account the first arrival time at the port. As a result, in each step of the dynamic programming method, i.e., at port of call $\bar{i}$, we have to record information on the arrival time at all the ports that have been visited and are to be visited again. This, in theory, may lead to the "curse-of-dimensionality", because if there are n ports that are visited twice, in the worst-case we have to record the arrival times at n ports of call, and if the possible arrival times is e.g. 168, then the state space is 168^n (without even considering the possible arrival times at $\bar{i}$), which increases exponentially with n.

In reality, this problem may not be that serious. This is because on one side, the number of ports of call in a round-trip journey is not very large. On the other side, the number of ports that are visited twice in a round-trip journey is even smaller. In addition, some ports may be always available, especially those major transshipment hubs such as Singapore and Hong Kong that attract transshipment containers based on their quality of service.

Another minor issue that is worth mentioning is that when a ship route has ports that are visited twice in a round-trip journey, the definition of port time window at these ports should be changed. For instance, if a port is visited only once, we only need to record the possible arrival times in a week at the port, i.e., $\hat{\Omega}_i$, with regard to all berths at the port. We consider a port with two berths, assuming that both berths are available on Sunday and Monday, and the port time is 24 h, then $\hat{\Omega}_i = [0, 24)$. However, if the port is visited twice, we have to record the the time window of each *berth* at the port, because different arrivals may use different berths. Of course, this is only a minor issue in the dynamic programming algorithm, because we actually do not need to record which berth the first arrival has used.

Schedule Design for a Liner Shipping Network

Another issue that we will investigate is the schedule design problem with port time windows for a liner shipping network. In a liner shipping network, quite often than not, a port may be visited more than once in a week. Some major transshipment hubs, such as Singapore and Hong Kong, may be visited more than 20 times. Therefore, the berth time windows at each port have to be dealt with with special efforts. Apparently, dynamic programming is no longer applicable due to the "curse-of-dimensionality".

When designing schedules for a liner shipping network, there is further the problem of container transshipment. In particular, we hope that containers could stay at transshipment ports for as short time as possible, by matching the arrival times of different ships. As a result, the schedules of different ship routes interact with each other, because the arrival time at a port of call on a ship route affects the arrival time at the port by ships on other ship routes. Hence, the resulting problem is more challenging and interesting.

Schedule Design with Container Routing

The transportation of containers in a liner shipping network is mainly determined by the transshipment cost, i.e., the liner shipping company aims to transport containers at minimum transshipment cost. However, the transit time or inventory cost of containers should also be incorporated, because it affects the competitiveness of the liner shipping company. The transit time is determined by the schedules of ship routes in the liner shipping network. Therefore, the schedule design problem for a liner shipping network and container routing should be examined in a holistic approach.

The challenge with such a problem lies in that the joint planning of schedule design and container routing is a highly nonlinear problem. In fact, Wang and Meng (2011) examined such a problem with some simplifications of schedule design. Due to the highly nonlinear property, they developed a hybrid genetic local search heuristic. The framework of the heuristic, i.e., iteratively optimizing schedule and container routing, may be further explored.

Schedule Design Under Uncertainty

In our study the port time and sea time are assumed to be deterministic, and possible uncertainty is incorporated by adding some "buffer" time. Such an engineering-based approach may not lead to optimal decisions. A worthwhile avenue is to capture port time and sea time uncertainty endogenously. This problem is complex because a natural problem that cannot be circumvented is: what should a ship do if it is delayed? In reality, the ship may speed up (Qi and Song 2012; Wang and Meng 2012b, c), may skip ports of call (Chang et al. 2013b), may swap ports of call (Chang et al. 2013b) and may leave a port early without loading all containers. The handling of delay itself is a challenging topic, even when the planned schedule is given. There is a long way to go to address this problem satisfactorily.

References

Agarwal, R., & Ergun, Ö. (2008). Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science, 42*(2), 175–196.

Alvarez, J. (2009). Joint routing and deployment of a fleet of container vessels. *Maritime Economics and Logistics, 11*(2), 186–208.

Ambrosino, D., & Sciomachen, A. (1998). A constraint satisfaction approach for master bay plans. In G. Sciutto & C. Brebbia (eds.), *Maritime engineering and ports*. WIT Press, Boston, pp. 155–164.

Ambrosino, D., Sciomachen, A., & Tanfani, E. (2004). Stowing a containership: The master bay plan problem. *Transportation Research Part A, 38*(2), 81–99.

Ambrosino, D., Sciomachen, A., & Tanfani, E. (2006). A decomposition heuristics for the container ship stowage problem. *Journal of Heuristics, 12*(3), 211–233.

Aslidis, T. (1989). Combinatorial algorithms for stacking problems. PhD Thesis, MIT.

Aslidis, T. (1990). Minimizing of overstowage in container ship operations. *Operations Research, 90,* 457–471.

Avriel, M., & Penn, M. (1993). Exact and approximate solutions of the container ship stowage problem. *Computers and Industrial Engineering, 25,* 271–274.

Avriel, M., Penn, M., Shpirer, N., & Witteboon, S. (1998). Stowage planning for container ships to reduce the number of shifts. *Annals of Operations Research, 76,* 55–71.

Bell, M., Liu, X., Angeloudis, P., Fonzone, A., & Hosseinloo, S. (2011). A frequency-based maritime container assignment model. *Transportation Research Part B, 45*(8), 1152–1161.

Brouer, B., Alvarez, J. F., Plum, C., Pisinger, D., & Sigurd, M. (2013a). A base integer programming model and benchmark suite for linear shipping network design. *Transportation Science*. *48*(2), 281–312.

Brouer, B., Dirksen, J., Pisinger, D., Plum, C., & Vaaben, B. (2013b). The vessel schedule recovery problem (VSRP)—A MIP model for handling disruptions in liner shipping. *European Journal of Operational Research, 224*(2), 362–374.

Chang, D., Jiang, Z., Yan, W., & He, J. (2010). Integrating berth allocation and quay crane assignments. *Transportation Research Part E, 46*(6), 975–990.

Chen, J., Lee, D.-H., & Cao, J. (2012). A combinatorial benders cuts algorithm for the quayside operation problem at container terminals. *Transportation Research Part E, 48*(1), 266–275.

Christiansen, M., Fagerholt, K., & Ronen, D. (2004). Ship routing and scheduling: Status and perspectives. *Transportation Science, 38*(1), 1–18.

Christiansen, M., Fagerholt, K., Nygreen, B., & Ronen, D. (2013). Ship routing and scheduling in the new millennium. *European Journal of Operational Research, 228*(3), 467–478.

Cordeau, J.-F., Laporte, G., Legato, P., & Moccia, L. (2005). Models and tabu search heuristics for the berth-allocation problem. *Transportation Science, 39*(4), 526–538.

Dong, J., & Song, D. (2012). Cargo routing and empty container repositioning in multiple shipping service routes. *Transportation Research Part B, 46*(10), 1556–1575.

Du, Y., Chen, Q., Quan, X., Long, L., & Fung, R. (2011). Berth allocation considering fuel consumption and vessel emissions. *Transportation Research Part E, 47*(6), 1021–1037.

Dubrovsky, O., Levitin, G., & Penn, M. (2002). A genetic algorithm with compact solution encoding for the container ship stowage problem. *Journal of Heuristics*, *8,* 585–599.

Fagerholt, K. (1999). Optimal fleet design in a ship routing problem. *International Transactions in Operational Research, 6*(5), 453–464.

Gelareh, S., & Meng, Q. (2010). A novel modeling approach for the fleet deployment problem within a short-term planning horizon. *Transportation Research Part E, 46*(1), 76–89.

Giallombardo, G., Moccia, L., Salani, M., & Vacca, I. (2010). Modeling and solving the tactical berth allocation problem. *Transportation Research Part B, 44*(2), 232–245.

Golias, S., Boile, M., & Theofanis, S. (2009a). Berth scheduling by customer service differentiation: A multi-objective approach. *Transportation Research Part E, 45*(6), 878–892.

Golias, S., Saharidis, G., Boile, M., Theofanis, S., & Ierapetritou, M. (2009b). The berth allocation problem: Optimizing vessel arrival time. *Maritime Economics and Logistics, 11*(4), 358–377.

Golias, S., Boile, M., & Theofanis, S. (2010a). A lamda-optimal based heuristic for the berth scheduling problem. *Transportation Research Part C, 18*(5), 794–806.

Golias, S., Boile, M., Theofanis, S., & Efstathiou, C. (2010b). The berth-scheduling problem: Maximizing berth productivity and minimizing fuel consumption and emissions production. *Transportation Research Record, 2166,* 20–27.

Han, X.-L., Lu, Z.-Q., & Xi, L.-F. (2010). A proactive approach for simultaneous berth and quay crane scheduling problem with stochastic arrival and handling time. *European Journal of Operational Research, 207*(3), 1327–1340.

Imai, A., & Miki, T. (1989). A heuristic algorithm with expected utility for an optimal sequence of loading containers into a containerized ship. *Journal of Japan Institute of Navigation, 80,* 117–124.

Imai, A., Nishimura, E., & Papadimitriou, S. (2001). The dynamic berth allocation problem for a container port. *Transportation Research Part B, 35*(4), 401–417.

Imai, A., Nishimura, E., & Papadimitriou, S. (2003). Berth allocation with service priority. *Transportation Research Part B, 37*(5), 437–457.

Imai, A., Sun, X., Nishimura, E., & Papadimitriou, S. (2005). Berth allocation in a container port: Using a continuous location space approach. *Transportation Research Part B, 39*(3), 199–221.

Imai, A., Chen, H., Nishimura, E., & Papadimitriou, S. (2008a). The simultaneous berth and quay crane allocation problem. *Transportation Research Part E, 44*(5), 900–920.

Imai, A., Nishimura, E., & Papadimitriou, S. (2008b). Berthing ships at a multi-user container terminal with a limited quay capacity. *Transportation Research Part E, 44*(1), 136–151.

Kim, K., & Moon, K. (2003). Berth scheduling by simulated annealing. *Transportation Research Part B, 37*(6), 541–560.

Kontovas, C., & Psaraftis, H. (2011). Reduction of emissions along the maritime intermodal container chain: Operational models and policies. *Maritime Policy and Management, 38*(4), 451–469.

Lee, D.-H., Chen, J., & Cao, J. (2010). The continuous berth allocation problem: A greedy randomized adaptive search solution. *Transportation Research Part E, 46*(6), 1017–1029.

Meisel, F., & Bierwirth, C. (2009). Heuristics for the integration of crane productivity in the berth allocation problem. *Transportation Research Part E, 45*(1), 196–209.

Meng, Q., & Wang, S. (2011a). Liner shipping service network design with empty container repositioning. *Transportation Research Part E, 47*(5), 695–708.

Meng, Q., & Wang, T. (2011b). A scenario-based dynamic programming model for multi-period liner ship fleet planning. *Transportation Research Part E, 47*(4), 401–413.

Meng, Q., & Wang, S. (2012). Liner ship fleet deployment with week-dependent container shipment demand. *European Journal of Operational Research, 222*(2), 241–252.

Meng, Q., Wang, S., & Liu, Z. (2012a). Network design for shipping service of large-scale intermodal liners. *Transportation Research Record, 2269,* 42–50.

Meng, Q., Wang, T., & Wang, S. (2012b). Short-term liner ship fleet planning with container transshipment and uncertain demand. *European Journal of Operational Research, 223*(1), 96–105.

Meng, Q., Wang, S., Andersson, H., & Thun, K. (2014). Containership routing and scheduling in liner shipping: Overview and future research directions. *Transportation Science*. *48*(2), 265–280.

Moorthy, R., & Teo, C.-P. (2006). Berth management in container terminal: The template design problem. *OR Spectrum, 28*(4), 495–518.

Mourão, M., Pato, M., & Paixão, A. (2001). Ship assignment with hub and spoke constraints. *Maritime Policy and Management, 29*(2), 135–150.

Notteboom, T. (2006). The time factor in liner shipping services. *Maritime Economics and Logistics, 8*(1), 19–39.

OOCL. (2013). The North & Central China East coast express. Orient overseas container lineWebsite. http://www.oocl.com/eng/ourservices/serviceroutes/tpt/Pages/default.aspx. Accessed on 5 Feb 2013.

Psaraftis, H., & Kontovas, C. (2010). Balancing the economic and environmental performance of maritime transportation. *Transportation Research Part D, 15*(8), 458–462.

Psaraftis, H., & Kontovas, C. (2013). Speed models for energy-efficient maritime transportation: A taxonomy and survey. *Transportation Research Part C, 26,* 331–351.

Qi, X., & Song, D. (2012). Minimizing fuel emissions by optimizing vessel schedules in liner shipping with uncertain port times. *Transportation Research Part E, 48*(4), 863–880.

Reinhardt, L., & Pisinger, D. (2012). A branch and cut algorithm for the container shipping network design problem. *Flexible Services and Manufacturing Journal, 24*(3), 349–374.

Ronen, D. (2011). The effect of oil price on containership speed and fleet size. *Journal of the Operational Research Society, 62*(1), 211–216.

Shintani, K., Imai, A., Nishimura, E., & Papadimitriou, S. (2007). The container shipping network design problem with empty container repositioning. *Transportation Research Part E, 43*(1), 39–59.

Wang, S. (2013). Essential elements in tactical planning models for container liner shipping. *Transportation Research Part B, 54,* 84–99.

Wang, S., & Meng, Q. (2011). Schedule design and container routing in liner shipping. *Transportation Research Record, 2222,* 25–33.

Wang, S., & Meng, Q. (2012a). Liner ship fleet deployment with container transshipment operations. *Transportation Research Part E, 48*(2), 470–484.

Wang, S., & Meng, Q. (2012b). Liner ship route schedule design with sea contingency time and port time uncertainty. *Transportation Research Part B, 46*(5), 615–633.

Wang, S., & Meng, Q. (2012c). Robust schedule design for liner shipping services. *Transportation Research Part E, 48*(6), 1093–1106.

Wang, S., & Meng, Q. (2012d). Sailing speed optimization for container ships in a liner shipping network. *Transportation Research Part E, 48*(3), 701–714.

Wang, S., & Meng, Q. (2013). Reversing port rotation directions in a container liner shipping network. *Transportation Research Part B, 50,* 61–73.

Wang, S., & Meng, Q. (2014). Liner shipping network design with deadlines. *Computers & Operations Research, 41,* 140–149.

Wang, S., Wang, T., & Meng, Q. (2011). A note on liner ship fleet deployment. *Flexible Services and Manufacturing Journal, 23*(4), 422–430.

Wang, S., Meng, Q., & Sun, Z. (2013). Container routing in liner shipping. *Transportation Research Part E, 49*(1), 1–7.

Wilson, I., & Roach, P. (1999). Principles of combinatorial optimization applied to container-ship stowage planning. *Journal of Heuristics, 5,* 403–418.

Wilson, I., & Roach, P. (2000). Container stowage planning: A methodology for generating computerized solutions. *Journal of Operational Research Society, 51,* 1248–1255.

Wilson, I., Roach, P., & Ware, J. (2001). Container stowage pre-planning: Using search to generate solutions, a case study. *Knowledge-Based Systems, 14,* 137–145.

Yan, S., Chen, C.-Y., & Lin, S.-C. (2009). Ship scheduling and container shipment planning for liners in short-term operations. *Journal of Marine Science and Technology*, *14*(4), 417–435.

Zhen, L., Chew, E., & Lee, L. (2011a). An integrated model for berth template and yard template planning in transshipment hubs. *Transportation Science*, *45*(4), 483–504.

Zhen, L., Lee, L., & Chew, E. (2011b). A decision model for berth allocation under uncertainty. *European Journal of Operational Research*, *212*(1), 54–68.

Chapter 11
Slow Steaming in Maritime Transportation: Fundamentals, Trade-offs, and Decision Models

Harilaos N. Psaraftis and Christos A. Kontovas

Abstract Slow steaming is being practised in many sectors of the shipping industry. It is induced principally by depressed shipping markets and/or high fuel prices. In recent years the environmental dimension of slow steaming has also become important, as ship emissions are directly proportional to fuel burned. The purpose of this chapter is to examine the practice of slow steaming from various angles. In that context, a taxonomy of models is presented, some fundamentals are outlined, the main trade-offs are analysed, and some decision models are presented. Some examples are finally presented so as to highlight the main issues that are at play.

11.1 Introduction

In recent times, increasing fuel prices and depressed market conditions have brought a new perspective to ship speed. For a variety of reasons, economic but also environmental, sailing at full speed may not necessarily be the best choice. In that sense, optimizing ship speed is receiving increased emphasis these days and is likely to do so in the years ahead.

Ships travel slower than the other transportation modes, but a basic premise has always been that there is value in ship speed. As long-distance trips may typically last one to two months, the benefits of a higher speed may be significant: they mainly entail the economic added value of faster delivery of goods, lower inventory costs and increased trade throughput per unit time. The need for higher speeds in shipping was mainly spurred by strong growth in world trade and development, and in turn was made possible by significant technological advances in maritime transportation in a broad spectrum of areas, including hull design, hydrodynamic performance of vessels, engine and propulsion efficiency, to name just a few. By extension, developments in cargo handling systems and supply chain management and operation have also contributed significantly to fast door-to-door transportation. However, this basic premise is being challenged whenever shipping markets are

---

H. N. Psaraftis (✉) · C. A. Kontovas
Department of Transport, Technical University of Denmark,
Produktionstorvet 424, 2800 Kgs. Lyngby, Denmark
e-mail: hnpsar@transport.dtu.dk

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_11

not very high and whenever fuel prices are not low. In addition, perhaps the most significant factor that is making a difference in recent years is the environmental one: a ship has to be environmentally friendly as regards air emissions. Because of the non-linear relationship between speed and fuel consumption, it is obvious that a ship that goes slower will emit much less than the same ship going faster.

Even for the simple objective to reduce fuel costs (and by extension emissions) by reducing speed, this can be done at two levels. The first level is technological (strategic), that is, build future ships with reduced installed horsepower so that they cannot sail faster than a prescribed speed. However, the first cellular containerships that went up to 33 knots in the late 1960s when fuel was cheap are gone forever. Maersk's new 18,000 TEU 'Triple-E'[1] containerships have a design speed of 17.8 knots, down from the 22–25 knots range that has been the industry's norm, and will emit 20 % less CO_2 per container moved as compared to the Emma Maersk, previously the world's largest container vessel, and 50 % less than the industry average on the Asia-Europe trade lane (Maersk 2013).

The second level is logistics-based (tactical/operational), that is, have an existing ship go slower than its design speed. In shipping parlance this is known as "slow steaming" and may involve just slowing down or even 'derating' a ship's engine, that is, reconfiguring the engine so that a lower power output is achieved, so that even slower speeds can be attained[2]. Depending on engine technology, 'slow steaming kits' are provided by engine manufacturers so that ships can smoothly reduce speed at any desired level. In case speed is drastically reduced, the practice is known as "super slow steaming".

In practice, super slow steaming has been pioneered by Maersk Line after it initiated trials involving 110 vessels beginning in 2007. Maersk Line North Asia Region CEO Tim Smith said that the trials showed it was safe to reduce the engine load to as low as 10 %, compared with the traditional policy of reducing the load to no less than 40–60 % (TradeWinds 2009). Given the non-linear relationship between speed and power, for a containership a 10 % engine load means sailing at about half of the design speed. Furthermore, China Ocean Shipping (Group) and its partners in the CKYH alliance (K Line, Yang Ming Marine and Hanjin Shipping) were also reported to introduce super-slow steaming on certain routes (Lloyd's List 2009).

Slow steaming is not only practiced in the container market, although it may seem to make more sense there due to the higher speeds of containerships. Slow steaming is reported in every market. In December 2010, Maersk Tankers was reported to have their Very Large Crude Carriers (VLCCs) sailing at half their speed. The design speed of 16 knots was reduced to speeds less than 10 knots on almost one third of its ballast legs and between 11 and 13 knots on over one third of its operating days. For example, a typical voyage from the Persian Gulf to Asia normally takes 42 days (at 15 knots laden and 16 knots in ballast). Maersk Tankers decreased speed to 8.5

---

[1] Triple-E stands for Economy of scale, Energy efficiency and Environmentally improved performance.

[2] Such a reconfiguration may involve dropping a cylinder from the main engine or other measures.

knots on the ballast leg, thus increasing roundtrip time to 55 days and saving nearly $ 400,000 off the voyage's bunker bill (TradeWinds 2010).

Slow steaming has also an important role on absorbing fleet overcapacity. Since early 2009, the total containership capacity absorbed due to the longer duration of total roundtrip time for long haul services has reached 1.27 MTEU in October 2013 (taking early 2009 as a starting point), based on Alphaliner's latest estimates (Alphaliner 2013). The average duration of Far East-North Europe strings had increased from 8 weeks in 2006 to 9 weeks in 2009 when slow steaming was first adopted. The application of even lower speeds has pushed the figure to 11 weeks currently as carriers continue to seek further cost reductions by adopting slower sailing speeds. The same phenomenon has been observed on Far East-Med strings, where the average duration has risen to 10 weeks, compared to only 7 weeks in 2006. As a record number of deliveries of new vessels is continuing to hamper the supply and demand momentum, analysts expect that slow steaming is here to stay. As a record number of vessels were scrapped in 2013; the idle fleet averaged 595,000 TEUs in 2013 compared to 651,000 TEUs in 2012. The lay-up of surplus box ships has been the worst and has lasted for the longest period since early 2009. The twin impact of extra slow steaming and longer port stays has helped to absorb much of capacity but it seems that sailing at even slower speeds is not an option. A similar situation pertains to bulk carriers and tankers. Thus, slow steaming is here to stay for the foreseeable future.

The purpose of this chapter is to examine the practice of slow steaming from various angles. In that context, some fundamentals are outlined, the main trade-offs are analysed, and some decision models are presented. Some examples are finally presented so as to highlight the main issues that are at play. Material in this chapter is mainly taken from various papers and other documents by the authors and their colleagues, including Gkonis and Psaraftis (2012), and Psaraftis and Kontovas (2013, 2014).

The rest of this chapter is organized as follows. Section 2 discusses a taxonomy of speed models. Section 3 presents the fundamentals of slow steaming. Section 4 discusses the impact of inventory costs. Section 5 summarizes results for VLCCs. Section 6 discusses the case of multiple optimal speeds. Section 7 discusses slow steaming vis a vis ports. Section 8 analyses combined speed and route models. Section 9 presents a case in which sailing the minimum distance route at minimum speed may not minimize fuel costs. Section 10 discusses policy implications. Last but not least, sect. 11 presents the chapter's conclusions.

11.2 Taxonomy of Speed Models

11.2.1 General

Determining the optimal speed of ships is not new in the literature. A first observation is that most of the models (at least implicitly) assume that fuel costs are being borne by the ship owner. In the tramp shipping market (served by tankers, dry bulk carriers, product carriers, and gas carriers) this is the case if the ship is on spot charter. It is known that the predominance of charter party contracts are time charters, in which fuel costs are borne by the charterer. Even though most models assume the ship owner as the party that bears the costs, including fuel, the related optimization problem is typically cost minimization rather than profit maximization. This is tantamount to assuming that revenue for the service is fixed. This is not the case however in most instances and thus some of the models that optimize speed do not capture the trade-off between a higher speed to make more profit-earning trips per unit time and the impact of such higher speed on costs (mainly on fuel).

For problems in the liner market (served by containerships and ro/ro ships), a similar situation pertains. Ship owners who run liner services using their own ships want to maximize profits and the same trade-offs are at play. But also using chartered ships to provide liner services is not uncommon, as a liner company typically employs a mix of owned and chartered vessels.

For the so-called ‘industrial’ types of problems, in which a company (for instance, an oil company) uses its own ships to move its own cargoes, again the usual objective is cost minimization. What is not often mentioned however is that these companies always have the option to enter the market, by either offering excess capacity at the prevailing spot rate, or hiring extra capacity at such rate. Thus, in a boom period it may make sense for an oil company to run its ships at a higher speed, so as to offer the excess capacity obtained to the spot market. Of course, that this opportunity exists does not necessarily mean it will be used.

The other general observation is the scarcity of ‘dynamic’ speed models in the literature, even though a model that assumes no fixed cargo throughput within a certain time interval, but a rolling horizon in which costs or profits are optimized per unit time might want to consider ship speed as a key variable. However this is not necessarily the case: a recent paper that develops heuristics on dynamic ship routing and scheduling probems (Tirado et al. 2012) does not incorporate speed aspects. An exception concerns weather routing models, in which typically ship speed is dynamically updated.

There are many ways of classifying related papers and models in the literature. A first-order classification involves grouping references into two major categories:

a. Those that present models in which emissions are not considered, and
b. Those that present models in which emissions are considered (together with other considerations).

11.2.2 Non-emissions Speed Models

As expected, these are papers or other publications that chronologically constitute the oldest set of reviewed papers, in terms of average publication date, although some of them are relatively recent.

Alderton (1981) presents a variety of criteria to determine the speed that maximizes profit and discusses how sensitive these speeds are to such inputs as port time, voyage distance, freight rates and bunker costs. The influence of cargo inventory costs is also taken into account. He differentiates between what he terms "Least Cost Speed", the one that maximizes profit per tonne carried, and "Maximum Profit Speed", the one that maximizes profit per day.

Benford (1981) proposes a simple procedure to select the mix of available ships from a fleet and their speeds in order to achieve the best solution for a fleet owner. The approach is confined to non-liner trades (in fact his examples are from the coal trades in the Great Lakes). He assumes that the owner has only one contract, meaning that total revenues revenues are fixed, hence the objective is minimum cost.

Perakis (1985) relaxes some of Benford's fleet deployment model assumptions and arrives at an optimal solution that reduces by 15 % the operating costs vis-a-vis those of Benford's.

Ronen (1982) investigates the effect of oil prices on the optimal speed of ships and presents three models, namely for the ballast (or positioning) leg, the income generating (or laden) leg and a variant of the laden case for which a penalty/bonus is given for late/early arrival. He analyzes the tradeoff between fuel savings through slow steaming and loss of revenues due to the increase of voyage time.

Perakis and Papadakis (1987a, b) deal with fleet deployment and the optimal speed for ships operating between a single loading port and a single discharging port. In that problem laden and ballast speeds for each ship are treated as the decision variables. In their second paper a sensitivity analysis of the optimal solution is performed and for the longer term problem, a time-dependent cost function and probabilistic analysis is used.

An expanded model to address a set of loading and unloading ports is presented in Papadakis and Perakis (1989) under the same assumption as in their previous work that each ship returns to the loading port in ballast. The same authors address a weather routing problem in which the objective is to minimize transit time and in which using control theory it is proven that the maximum permissible speed is the optimal speed (Perakis and Papadakis 1989).

Another weather routing problem is examined in Lo and McCord (1998), who present a fuel consumption minimization approach that addresses the uncertainty that results from the time lags between the time to collect and process raw data on ocean currents and the delivery of the estimation. They formulate the routing problem as an adaptive, probabilistic dynamic program.

Brown et al. (1987) study a crude oil tanker routing and scheduling problem that takes into account cost components and generate feasible ship schedules with different speeds and alternate routes of the ballast legs. For the laden condition, speed is not a decision variable since it is implicitly determined by the given loading and discharging dates.

Perakis and Jaramillo (1991) develop a more complex fleet deployment model for the liner trades. The objective is to minimize costs. A linear programming formulation is developed and the speed problem is decomposed from the deployment problem.

The option to charter in additional vessels is also considered and the model includes port, canal and lay up costs.

Bausch et al. (1998) develop a spreadsheet-based interface for scheduling the fleet of tankers and barges. Embedded is a set partitioning model that optimizes cost. The model is used by dispatchers whose native language is not English but who communicate with one another via the model interface.

Fagerholt (2001) considers a flexible situation in determining the optimal speeds on the various legs of the schedule. This is the so-called 'soft- time window' case, in which penalties are imposed if the vessel arrives at a port outside a specified time window. It is motivated by the fact that allowing some customers to have controlled time violations for both loading and unloading of cargo it may be possible to obtain better schedules and high reductions in shipping costs.

An extensive discussion of the various aspects of speed in maritime transportation from various angles is in Stopford (2004), which is the well known seminal book on maritime economics. The basic model assumes a cubic speed function, although it is stated that other exponents may be applicable.

Alvarez (2009) presents a mixed integer programming (MIP) formulation for the joint optimization of routing and fleet deployment of container vessels. Speed is considered as a variable so that the sailing time between any two ports is assumed to be deterministic and the time in port is fixed for each port-ship combination. The model minimizes the operating expenses of a liner company over a tactical planning horizon and the algorithm includes the possibility of rejecting transportation demand on a selective basis, with lost revenue and some monetary penalty.

Notteboom and Vernimmen (2010) deal with the impact of high fuel costs on the design of liner services on the Europe–Far East trade and discuss the way that shipping lines have adapted their schedules in terms of speed and number of vessels deployed for each loop. Furthermore, a cost model is developed to estimate the impact of the additional bunker cost on the operational costs and cost comparisons for different vessel sizes and vessels speeds are presented.

Lang and Veenstra (2010) study the problem of container vessel arrival planning and in that context assume a linearized speed model in which fuel cost is to be minimized. Linearization takes place for computational purposes. Even though fuel price is not an explicit input, results are presented under high and low fuel price regimes, the price difference between these two scenarios being 35 %.

None of the surveyed models attempts to estimate the equilibrium spot rate that would be established as a function of the fuel price and the optimal speed that ships would choose as a result. Devanney (2010) presents such an approach for VLCCs, by looking at the interaction of the VLCC fleet supply and demand curves.

A more general model is presented in Devanney (2007), which models the world's petroleum transportation network as a linear program, and simultaneously determines tanker optimal speeds in the laden and ballast legs, FOB and CIF prices of crude oil at origin and destination points, and the market equilibrium spot rates in various routes. The related software (termed Martinet) is only commercially available.

Norstad et al. (2011) present the tramp ship routing and scheduling problem with speed optimization, where speed is introduced as a decision variable. Although the main objective is to maximize profit by allowing the option of picking up spot cargoes, for the speed optimization subproblem the objective is to minimize costs on a certain leg of the route. The paper presents search heuristics to solve this problem and propose alternative algorithms. Various comparisons are also provided.

Ronen (2011) studies the effect of oil price on the trade-off between reducing sailing speed and increasing the fleet size for container ships and develops a procedure to indentify the sailing speed and number of vessels that minimize annual operating costs.

Meng and Wang (2011) proposes an optimal operating strategy problem arising in liner shipping industry that aims to determine service frequency, containership fleet deployment plan, and sailing speed for a long-haul liner service route. The problem is formulated as a mixed-integer nonlinear programming model and solved using an efficient and exact branch-and-bound based e-optimal algorithm. A case study based on an existing long-haul liner service route with fixed service frequency and fixed ship type is presented and the results for the optimisation in ship number and sailing speed are compared with Ronen (2011) and Gelareh and Meng (2010).

Wang and Meng (2012a) investigate the optimal speed of a fleet of container ships on each leg of each ship route in a liner network using a mixed-integer nonlinear programming model while considering transshipment and container routing. Their model uses a power bunker consumption function which is calibrated using historical operating data from a global liner shipping company.

Wang and Meng (2012b) develop and solve a model for a proposed liner ship route schedule design problem with sea contingency and uncertain port time in order to minimize the ship cost and bunker cost, while fulfilling the port-to-port transit time constraints. For each leg of each ship route they solve the optimal sailing speed problem in order to identify the optimal bunker consumption function as a function of the available sailing time t. Then they solve the schedule design problem by determining the arrival time and the number of vessels for each route by minimizing the sum of ship cost and the expected total bunker cost while satisfying the transit time constraints. However, late arrival at a port is not allowed in their model.

Thus, Wang and Meng (2012c) present a robust schedule design problem which takes into account the penalty for late arrival or late container handling as a result from uncertain port time. The problem is formulated using a mixed-integer nonlinear stochastic programming model and solved using an algorithm that incorporates a sample average approximation method, linearization techniques, and a decomposition scheme. In addition, numerical results based on an Asia–America–Europe ship route are presented to demonstrate that the algorithm obtains near-optimal solutions.

Yao et al. (2012) perform a study on bunker fuel management for container trades in which ship speed and fuel purchase location are the main decision variables. Minimization of total bunker costs is the objective function.

Even though the above models do not consider emissions, possible extensions could examine what would happen if the social cost of emissions (and essentially CO_2) is incorporated into the cost functions assumed by these models. Doing so would internalize the external cost of these emissions, a central (although seldomly applied) environmental policy goal.

11.2.3 Emissions Speed Models

Speed models that also consider emissions in a logistical context are on the average more recent.

Psaraftis and Kontovas (2009a) investigate the simple scenario where a fleet of identical ships, each of which loads from a port A, travels to port B with a known speed, discharges at B and goes back to port A in ballast, with a known speed. A result of the analysis is that total emissions would be always reduced by slowing down, even though more ships would be used. Another result is that if speed is reduced in a Sulphur Emissions Controlled Area (SECA) in order to reduce SO_x emissions and this is compensated by a speed increase outside the SECA so that total transit time is the same, overall emissions increase.

Corbett et al (2010) develop equations relating speed, energy consumption, and total cost to evaluate the impact of speed reduction on emissions. They also explore the relationship between fuel price and optimal speed.

Du et al (2011) use a speed model in the context of a berth allocation problem, in which they assume that the ship operator acts so as to minimize per route leg fuel consumption. A non-linear and not necessarily cubic fuel consumption function is obtained by regression analysis. The regression coefficients are obtained from data provided by a major marine engine manufacturer. Wang et al (2013) improve this model so that general fuel consumption functions can be handled more tractably.

Eefsen and Cerup-Simonsen (2010) examine the tradeoffs between lower fuel costs and higher inventory costs associated with speed reduction, as well as their impact on emissions. The model was used to investigate the transport costs and carbon emissions on a particular container route from China to Europe on a 6,600 TEU containership.

Faber et al (2010) estimate that emissions of bulkers, tankers and container vessels can be reduced maximally by about 30 % in the coming years by using the current oversupply to reduce speed, relative to the situation in 2007.

Fagerholt et al. (2010) consider a single route speed optimization problem with time windows and proposed a solution methodology in which the arrival times are discretized and the solution is based on the shortest path of the directed acyclic graph that is formed. Reduction in ship emissions are also computed. For the same problem, and drawing also from the results of Norstad et al (2011), Hvattum et al (2012) show that if fuel cost is a convex function of vessel speed, optimal speeds can be found in quadratic time.

Qi and Song (2012) investigate the problem of designing an optimal vessel schedule in the liner shipping route to minimize the total expected fuel consumption (hence also emissions) considering uncertain port times and frequency requirements on the liner schedule. The general optimal scheduling problem is formulated and tackled by simulation-based stochastic approximation methods.

Cariou (2011) investigates slow steaming strategies especially in container shipping and measures the reduction of CO_2 achieved in various container trades. In addition, the paper concludes that for the main trades speed reduction is cost beneficial when bunker price is at least $ 350–$400 per tonne.

Kontovas and Psaraftis (2011) examine speed reduction as an operational measure to reduce fuel consumption with a focus on container vessels. Since time at sea increases with slow steaming, there is a parallel and strong interest to investigate possible ways to decrease time in port. To that effect, a related berthing policy was investigated as a measure to reduce waiting time.

Another aspect of the problem is studied in Psaraftis and Kontovas (2010), where the impact of speed reduction on modal split is investigated, in the sense that cargoes that go slower may choose alternative modes of transport, particularly if their inventory costs are high. This may be true not only for short sea trades, but for longer haul ones, for example using the Trans-siberian railway to move cargoes to or from the Far East. Multinomial logit models are introduced.

Lindstad et al (2011) present an analysis at the strategic level. They investigate the impact of lower speeds on the cost and emissions of the world fleet and argue that there is a significant potential for the reduction of GHGs if speed is reduced. They explore Pareto-optimal policies and recommend speed limits as a possible way to achieve speed reduction.

An opposing view is presented by Cariou and Cheaitou (2012), who investigate policy options contemplated by the European Commission and compare speed limits versus a bunker levy as two measures to abate GHGs, with a scenario from the container trades. They conclude that the latter measure is counterproductive for two reasons. First, because it may ultimately generate more emissions and incur a cost per tonne of CO_2 which is more than society is willing to pay. Second, because it is sub-optimal compared to results obtained if an international bunker-levy were to be implemented.

Gkonis and Psaraftis (2012) develop a series of models that optimize speed in both the laden and ballast legs for several tanker categories (VLCC—ULCC, Suezmax, Aframax, product tankers, LNG and LPG carriers) and for a variety of scenarios. The modeling approach consists of two steps. The first step performs a speed optimization for both laden and ballast sailing. This is carried out over certain defined routes and for a certain ship. The second step calculates the annual emissions for the global tanker fleet, broken down into size brackets. The data used is based on actual speed-consumption curves, rather than theoretical or modelling approximations. The impacts of inventory costs, bunker costs, freight rates and other parameters on optimal speeds and emissions are estimated.

Fagerholt and Ronen (2013) develop speed models for a mixed chartering scenario, in which a fleet of ships have the obligation to carry some ‘mandatory’ cargoes under a contract of affreightment scenario, but also have the option to add ‘optional’ cargoes on a spot charter basis. Maximizing profit is the objective.

Last but not least, Psaraftis and Kontovas (2014) clarify some important issues as regards ship speed optimization at the operational level and develop models that optimize ship speed for a spectrum of routing scenarios in a single ship setting. The paper’s main contribution is the incorporation of those fundamental parameters and other considerations that weigh heavily in a ship owner’s or charterer’s speed decision and in his routing decision, wherever relevant. Various examples are given so as to illustrate the properties of the optimal solution and the various trade-offs that are involved.

11.2.4 Taxonomy

A finer-grain taxonomy classifies the literature of the previous two sections according to the following parameters (see also Psaraftis and Kontovas (2013)):

Optimization Criterion The main variants here are cost (to be minimized) and profit (to be maximized). Other variants include fuel consumption, transit time, or others. To be sure, some models in the literature are not cast as optimization problems. In these papers we set 'cost', 'profit', or other, depending on what the model described by the paper tries to measure.

What is the Shipping Market/Context of the Problem? This may be tankers, bulk carriers, containerships, or other ship types. It may even involve the whole commercial fleet.

Who is the Decision Maker? By this we mean who decides what the ship speed should be. This can be the ship owner or the charterer. For weather routing problems, it is typically the ship's master. An attempt to designate who is the decision maker is made even if the model is not an optimization model.

Fuel Price an Explicit Input? Yes if fuel price is explicitly included as one of the explicit inputs of the problem, no otherwise.

Freight rate an Input? Yes if freight rate (spot, or other) is explicitly included as one of the explicit inputs of the problem, no otherwise. There are also models that compute that rate as an equilibrium rate depending on supply and demand.

Fuel Consumption Function It could be cubic, non-linear, linearized, general or unspecified.

Optimal Speeds in Various Legs Whether or not the model computes optimal speeds for each leg of the route (versus a single optimal speed).

Optimal Speed as Function of Payload Whether or not the model can compute the optimal speed as a function of how much full or empty the ship is.

Logistical Context This could be a fixed route scenario, a ship routing and scheduling problem, a fleet deployment problem, or other.

Size of Fleet One ship, or many ships.

Adding more Ships an Option This is so if adding (or subtracting) ships is an option so as to maintain constant throughput.

Inventory Costs Included Yes if cargo carrying (inventory) costs are included in the model, no otherwise.

Emissions Considered Yes or no.

Modal Split Considered Yes if model calculates the split among alternative and competing modes of transport as a function of problem inputs.

Ports Included in Formulation Yes if port times, costs, congestion, port emissions or other port-related variables are included in the model.

The full taxonomy is presented in Psaraftis and Kontovas (2013). A sample table is presented in Table 11.1 below.

11.3 Slow Steaming Fundamentals

We now come to presenting what in our opinion are the fundamentals in slow steaming.

11.3.1 Is ship Speed Fixed?

The first fundamental is something that many papers in the literature seem to ignore: ships do not trade at predetermined speeds. Those who pay for the fuel, that is, a ship owner whose ship trades on the spot market, or a charterer if the ship is on time charter, may want to choose the ship speed as a function of (a) fuel price and (b) market spot rate. In periods of depressed market conditions, as is the typical situation these days, ships tend to slow steam. The same is the case if bunker prices are high. Conversely, in boom periods or in case fuel prices are low, ships tend to sail faster.

An exception to the case that the ship owner or the charterer can freely choose an optimal speed for the ship is in case the ship is *on spot charter* and speed is prescribed in the charter party contract, either explicitly (speed is, say, 15 knots) or implicitly (pickup and delivery dates are prescribed). In spot charters (rental of the ship for a single voyage) the fuel is paid for by the ship owner. Agreeing on a prescribed speed in the charter party involves in most cases only the laden part of the trip, with the owner free to choose his speed on the ballast return leg. The speed that is agreed upon for the laden leg may or may not be the speed that the ship owner would have freely chosen if no explicit agreement were in place. If it is higher, the ship owner may ask for a higher rate than the prevailing spot rate, understanding of course that in this case he may lose the customer to a competitor ship, with which the charterer can obtain more favorable terms. For a discussion of possible distortions and additional emissions that can be caused by charter party speed agreements see Devanney (2011).

11.3.2 Who is the Speed Optimizer?

The second fundamental is perhaps not immediately obvious. This is that even though the owner's and time charterer's speed optimization problems appear at first glance different, the optimal ship speed for both problems turns out to be the same. A proof

Table 11.1 Taxonomy—a sample table. Source: Psaraftis and Kontovas (2013)

Taxonomy parameter paper	Fagerholt (2001)	Kontovas and Psaraftis (2011)	Lindstad et al. (2011)	Notteboom Vernimmen (2013)	Ronen (2010)
Optimization criterion	Cost	Cost	Pareto analysis	Cost	Cost
Shipping market	General	Container	All major ship types	Container	Container
Decision maker	Owner	Charterer	Owner	Owner	Owner
Fuel price an explicit input	No	Yes	Yes	Yes	Yes
Freight rate an input	No	Input	No	No	No
Fuel consumption function	Cubic	Cubic	Cubic	Unspecified	Cubic
Optimal speeds in various legs	Yes	Yes	No	No	No
Optimal speeds as function of payload	No	Yes	Yes	No	No
Logistical context	Pickup and delivery	Fixed route	Fixed route	Fixed route	Fixed route
Size of fleet	One ship	Multiple ships	Multiple ships	Multiple ships	Multiple ships
Add more ships an option	No	Yes	Yes	Yes	Yes
Inventory costs included	No	Yes	Yes	No	No
Emissions considered	No	Yes	Yes	No	No
Modal split considered	No	No	No	No	No
Ports included	No	Yes	Yes	Yes	Yes

is in Devanney (2010) for a rudimentary scenario of a ship hauling cargo from port 1 to port 2 and returning to port 1 on ballast (empty), and goes roughly as follows.

For a given ship, a ship owner in the spot market should operate at a speed that maximizes profit per day. Then his speed optimization problem is the following:

$$\max_v \{sC/(d/v) - pf(v) - E\} \quad (11.1)$$

where

s is the spot rate received by the owner (in \$/tonne)
C is the ship's cargo capacity (in tonnes)
d is the roundtrip distance (in nautical miles)

v is the sailing speed in nautical miles per day[3]
p is the bunker price (in \$/tonne)
f(v) is the daily fuel consumption function at speed v (t/day) and
E are the operating expenses borne by the ship owner other than fuel costs, including crew wages, insurance, etc (in \$/day).

In the above scenario, time in port has been ignored, although including it is a straightforward extension. Also the function f(v) is assumed to be the same in both directions (laden and ballast), although having different functions and different speeds on each leg is also a straightforward extension.

For a time charterer who has chartered the same ship, and who is the effective owner of the vessel during the period of the contract (also known in shipping parlance as the "disponent owner") faces the following problem:

$$\min_v \{s[R - Cv/d] + pf(v) + T\} \tag{11.2}$$

where

R is how much cargo needs to be moved (t/day)
T is the time charter rate paid to the owner (\$/day)

Equation (11.2) above assumes that any difference between the cargo capacity required by the time charterer (R) and what the chartered ship can provide if sailing at speed v (Cv/d) can be chartered in the spot market at a spot rate of s. If the difference [R-Cv/d] is positive (meaning that the chartered ship sailing at speed v cannot fully satisfy the charterer's needs), then additional capacity is chartered in at a rate of s, assuming the spot chartered ship sailing at the same speed v. If this difference is negative (meaning that there is spare capacity in the time chartered ship), then that spare capacity can be chartered out at the same spot rate s.

It can be seen easily that problems (11.1) and (11.2) are mathematically equivalent. In fact, in (11.1) the term E does not depend on speed and can be discarded from the objective function, leading to

$$\max_v \{sC/(d/v) - pf(v)\} \tag{11.3}$$

In (11.2), one can separate the term (sR + T) which does not depend on speed and thus can be discarded as well. What is then left is

$$\min_v \{pf(v) - sCv/d\} \tag{11.4}$$

Problems (11.3) and (11.4) are essentially the same, thus leading to the same optimal speed.

Factoring out the spot rate s, both problems can be rewritten as follows:

$$\min_v \{(p/s)f(v) - Cv/d\} \tag{11.5}$$

---

[3] This is 24 times the speed in knots. We use this unit to avoid carrying the number 24 through the calculations.

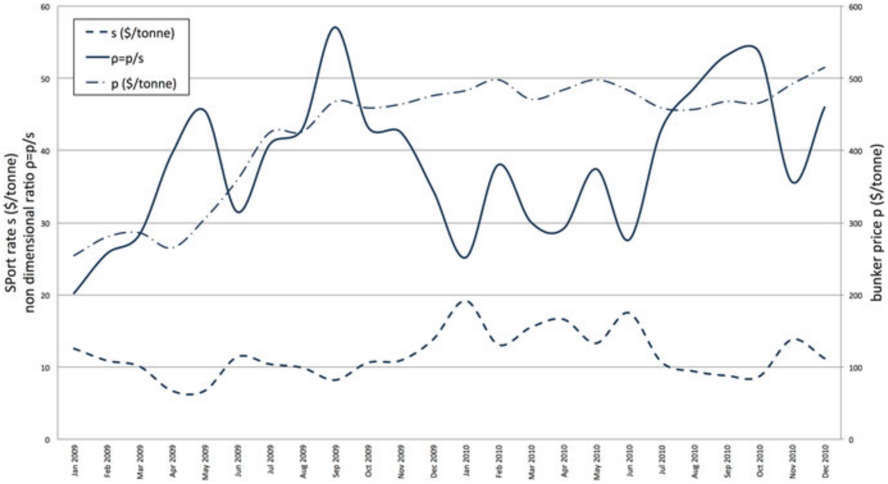


Fig. 11.1 Evolution of bunker price p, spot rate s and their ratio ρ = p/s. Data Source: Drewry Shipping Consultants

Equation (11.5) shows that for both problems, a key determinant parameter of the speed optimization problem is the *nondimensional ratio* $\rho = p/s$ of the bunker price divided by the spot rate, since for a given ship and route the optimal speed will be the same as long as ρ remains constant. Higher ρ ratios will generally induce lower speeds than lower ρ ratios. This corresponds to the typical behavior of shipping lines, which tend to slow steam in periods of depressed market conditions and/or high fuel prices and go faster if the opposite is the case.

Figure 11.1 shows a typical evolution of p, s and ρ for the tanker market. The period is 2009–2010, the route is Persian Gulf to Japan. HFO is the fuel and the fuel supplier is in the Persian Gulf.

11.3.3 Fuel Consumption Functions

The third fundamental of slow steaming is that fuel consumption (and hence fuel costs and emissions) depend non-linearly on ship sailing speed. The simplest model is to assume one type of fuel consumed on the ship, available at a known price of p (in $/tonne). Then the daily *at sea* fuel cost of a ship sailing from port i to port j is equal to $pf(v_{ij}, w_{ij})t_{ij,}$ where $f(v_{ij}, w_{ij})$ is the ship's daily fuel consumption at sea (in t/day), a known function of the ship's speed v_{ij} and payload w_{ij} from i to j, and t_{ij} is the ship's sailing time from i to j, given by the ratio (d_{ij}/v_{ij}), sailing distance divided by speed. Function f depends on many ship parameters, such as type and size of power plant, including main and auxiliary engines, geometry of ship hull, propeller design, and other parameters (weather conditions for instance). It can even be defined for $w_{ij} = 0$ (ship going on ballast). *In port* fuel costs are proportional to overall total port residence time, and these depend on per day fuel consumption of the ship's auxiliary engines while in port. In case the ship uses different fuels for its main engine and auxiliary engines (for instance Heavy Fuel Oil-HFO and Marine Diesel Oil- MDO, respectively), total fuel cost is the summation of all relevant fuel types.

The fact that function f can be a complex function which may not even be defined in closed form does not prevent us from considering some modeling approximations. A usual approximation is that function f is equal to $A + Bv_{ij}^{n}$ with A, B and n input parameters such as $A \geq 0$, $B > 0$ and $n \geq 3$. Another approximation is that for a given speed, f is proportional to $(w_{ij} + L)^{2/3}$, where L is the weight of the ship if empty plus fuel on board and consumables (modified admiralty formula, see also Barass (2005)[4]. A combination of these two approximations can also be considered. Most papers in the literature assume a cubic function, that is, $A = 0$ and $n = 3$ and no dependency on payload. $n = 3$ is usually a good approximation for tankers and bulk carriers and for the range of typical operational speeds of these vessels. A basic drawback of a cubic function is that it is invalid for very low speeds. In fact this function gives zero fuel consumption at zero speed, which is not the case in practice, as a ship, even stationary, consumes some fuel. Another drawback of a cubic function is that it may not be a good approximation for some ship types, containerships being the most notable example. For these ships, exponent n can be 4 or 5 or conceivably even higher.

Figure 11.2 below shows two typical fuel consumption curves for a VLCC, one for the laden condition and one for the ballast condition. Consumption of auxiliary engines is included. The functions in the figure are general and based on real data. Notice also that the curves are not defined below some minimum speed levels (on which more later).

With the above fundamentals in mind, we next examine a basic side effect of slow steaming, the impact of in-transit inventory costs.

11.4 Impact of In-Transit Inventory Costs

Problem (11.2) of the previous section does not include the in-transit inventory costs of cargo, to be borne by the charterer and due to the fact that the cargo is in transit for d/2v days (again, d is the roundtrip distance and cargo travels only one way). These costs depend on transit time and hence on speed, a lower speed entailing higher such costs. If these costs are not already factored in the negotiated market spot rate s, they are equal to $\beta C/2$ (\$/day) if β is the per day and per tonne inventory cost of the cargo. The latter is equal to PR/365 if P is the CIF value of the cargo in \$/tonne and R is the charterer's cost of capital.

---

[4] A first order approximation is that f does not take into account the reduction in the ship's total displacement due to fuel, lubricating oil or other consumables (such as fresh water) being consumed along the ship's route, since displacement would not change much as a result of that consumption.

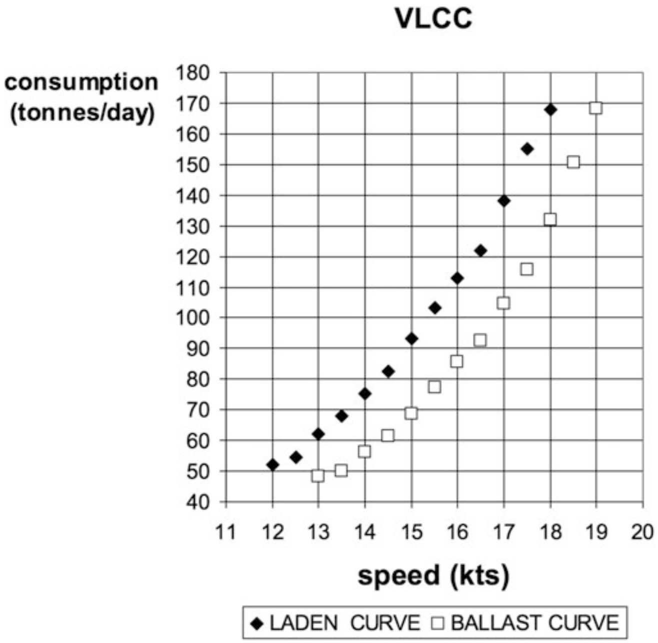


Fig. 11.2 Fuel consumption versus speed (in knots) for a VLCC. (Source: Gkonis and Psaraftis (2012))

Cargo inventory costs can be important, mainly in the liner business which involves trades of higher valued goods than bulk trades. The unit value of the top 20 containerized imports at the Los Angeles and Long Beach Ports in 2004 varied from about $ 14,000/t for furniture and bedding to $ 95,000/t for optic, photographic and medical instruments (CBO 2006). Delaying one tonne of the latter category of cargo by one week because of reduced speed would cost some $ 91 if the cost of capital is 5 %. For a $ 75,000/t payload this would amount to some $ 6.8 million. This may or may not be greater than the reduction of cost due to reduced speed (see Kontovas and Psaraftis (2011) for some examples).

It is straightforward to check that if inventory costs are included, equation (11.4) can be modified by replacing the spot rate s by (s-βd/2v). This problem is tantamount to the owner's problem (11.3) if the spot rate s is replaced by (s-βd/2v), if in fact cargo inventory costs are not factored in when the ship owner negotiates the spot charter with the charterer.

We mention these rudimentary problems because many models that we have reviewed assume (explicitly or implicitly) a fixed revenue for the ship owner and hence ignore the first term in (11.3). This is typically the case for routing and scheduling models in which the set of cargoes is fixed. If the amount of cargo to be transported in a year or within a given time period is fixed, then the ship owner's revenue is also fixed and then obviously the speed optimization problem of the ship owner is a cost minimization problem, subject to the constraint that this fixed quantity of cargo should be hauled. However, a ship owner may like to take advantage of high spot rates by hauling as much cargo as possible within a given period of time. In that case, the set of cargoes is not fixed. Conversely, if the market is low, ships tend to slow steam, as the additional revenue from hauling more cargo is less than the additional cost of the fuel. Even a charterer or an industrial shipping company may conceivably want to take advantage of such opportunities. Not factoring in the state of the market in a speed model means that the model may not capture one of the fundamental facets of shipping industry behavior, according to which the state of the market, along with the price of fuel, are the two main determinants of the speed of vessels.

The above simple model can be extended to the case in which speeds are optimized separately for the laden and ballast legs of a route, assuming different fuel consumption functions for each leg, and port times and costs are included. Figure 11.3 shows how optimal speeds in the laden and ballast leg conditions may vary as a function of fuel price and market rate for a modern VLCC operating from the Persian Gulf to Japan. Spot rates are expressed in terms of World Scale (WS) equivalents[5]. In-transit inventory costs are being included as an option in Fig. 11.4.

One can observe that optimal ballast speeds are typically higher (by 1–1.5 knots) than optimal laden speeds, except if cargo inventory costs are accounted for, in which case laden speeds can be higher than ballast speeds (depending on fuel price). In practice however, many tankers sail faster on the laden leg than on the ballast leg, which is sub-optimal. The reason for this is more likely to be attributed to charter party speeds than inventory costs (Devanney 2011).

In an even more general case, in which the ship is intermediately full at each route leg (a typical situation with containerships), different speeds can be chosen for different legs of the route, so long as they are within a "speed window" $[v_{LB}(w_{ij}), v_{UB}(w_{ij})]$, where $v_{LB}(w_{ij})$and $v_{UB}(w_{ij})$ are lower and upper bounds (respectively) on ship speed if the ship's payload from i to j is w_{ij}. Typically both bounds are dictated by the maximum power and technology of the engine and by the ship's payload when sailing from i to j. Practically both speed bounds are decreasing functions of w_{ij} (a more heavily loaded ship is not able to run as fast as an emptier ship). The upper bound exists because of limits in the ship's power. The lower bound exists because it is simply impossible for a ship engine to run lower than a certain power, below which the engine simply stalls. For a given payload, modern, electronically controlled engines, possibly equipped with 'slow steaming kits', generally have a lower v_{LB} than older, mechanical camshaft engines. Weather also plays a role in both bounds, with a usual approximation involving a 'speed margin' for anything else than calm weather.

Other model formulations do not optimize on a per day basis, but in terms of total costs or profits for a prescribed set of cargoes, for instance on a fixed route scenario or even in a routing and scheduling problem for which the ship route needs to be optimized.

---

[5] For a certain tanker route, WS is defined as 100 times the ratio of the prevailing spot rate on that route divided by the 'base rate' on that route (see Stopford (2004)).

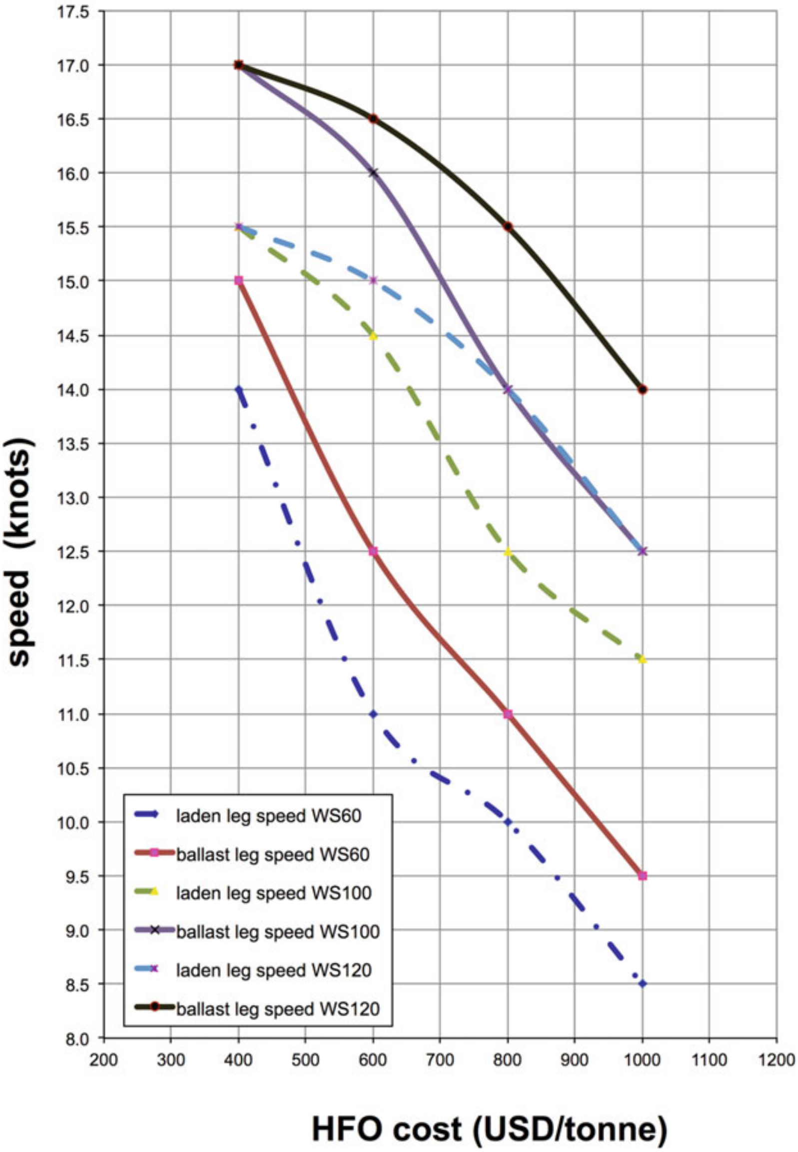


Fig. 11.3 Optimal VLCC laden and ballast speeds as functions of fuel price and spot rate. Spot rates are in WS. (Source: Gkonis and Psaraftis (2012))

Take for instance the case in which a ship on a fixed route wants to minimize costs over a specific route leg of length d. If v is the ship speed (miles per day), w is the ship payload during the leg (tonnes), p is the fuel price ($/tonne), f(v, w) is the fuel consumption function (t/day), T is the time charter rate the charterer is paying ($/day),

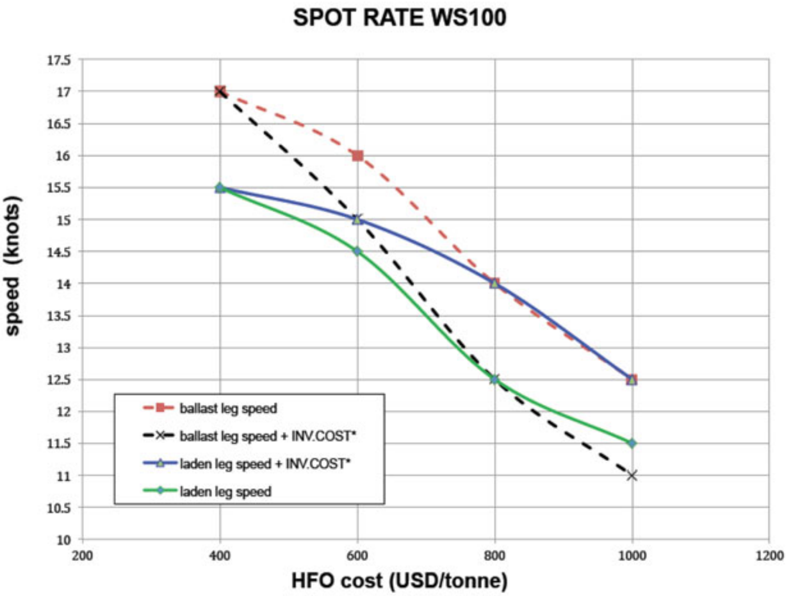


Fig. 11.4 Optimal VLCC laden and ballast speeds with and without inventory costs. (Source: Gkonis and Psaraftis (2012))

β is the inventory cost of the cargo ($/tonne/day), and $V = \{v: v_{LB}(w) \leq v \leq v_{UB}(w)\}$ is the set of allowable speeds, then the speed optimization problem for the charterer (who is the party paying for the fuel) for the specific leg is

$$\min_{v \in V} \{[pf(v, w) + \beta w + T](d/v)\}$$

As d is constant, this problem reduces to

$$\min_{v \in V}\{[pf(v, w) + \beta w \ + T]d/v\}$$

If function f is general (for instance given as a pointwise function), this problem can be solved by complete enumeration over all feasible values of v. If function f is given by a mathematical expression (cubic or other), more can be said.

Assuming for instance that $f(v, w) = (A + Bv^n)(w + L)^{2/3}$, the problem's objective function becomes

$$\{[pf(v, w) + \beta w + T\,]/v\}=\{[\,p(A + Bv^n)(w + L)^{2/3} + \beta w + T]/v\} = K/v + Mv^{n-1} \tag{11.6}$$

with $K= pA(w + L)^{2/3} + \beta w + T$

and $M = pB(w + L)^{2/3}$

Define v* the speed that makes the 1st derivative of expression (11.8) with respect to v equal to zero.

Then $v^* = \{K/[M(n-1)]\}^{1/n}$
If U is the optimal speed, then
If $v^* \leq v_{LB}(w)$, $U = v_{LB}(w)$
If $v_{LB}(w) \leq v^* \leq v_{UB}(w)\}$, $U = v^*$
If $v_{UB}(w) \leq v^*$, $U = v_{UB}(w)$

Even though this is a different model than the previous one, here too it can be seen that the higher the freight rate T, the lower the fuel price p, and the higher the value of the cargo (and hence β), the higher is the optimal speed U. This would seem to fit the pattern observed in many container trades. For instance, in the Far East to Europe trunk route, the busiest in the world, freight rates and average value of cargo are about double in the westbound direction than in the eastbound direction. This is reflected in the operational speeds, as most of slow steaming can be observed eastbound (Journal of Commerce 2010).

11.5 Speed Models for VLCCs

Reference was made in the previous section to the work of Gkonis and Psaraftis (2012), in which speed models were developed for several tanker classes, including Very Large Crude Carriers (VLCCs). These models pertain to both the single ship scenario and the fleet segment scenario. In this section we describe these models in more detail.

11.5.1 The Single Tanker Tool

The objective of this tool is the speed optimization for a known single crude oil tanker over a defined route. This follows the rationale outlined below:

- Typical route(s) per tanker segment are considered, where a typical size ship operates and to which some "average characteristics" are attributed
- The tool is run for the considered route and associated typical ship, under the defined assumptions and scenario (e.g. regarding freight rate levels & bunker prices)
- Output of the tool is the set of laden and ballast leg speeds for the considered route and associated typical ship, and also several emissions statistics.

The model is established in Microsoft Excel^{TM} and structurally consists of a number of main sections or "sheets", see Fig. 11.5.

A variety of runs and sensitivity analyses have been made, see Gkonis and Psaraftis (2012) for more details.

Figure 11.6 below depicts the effect of varying freight rates and fuel prices on annual CO_2 emissions for a specific VLCC running the route Ras Tanura- Yokohama. It can be seen that as the freight rate level decreases, the emissions decrease, as they are proportional to fuel consumption.

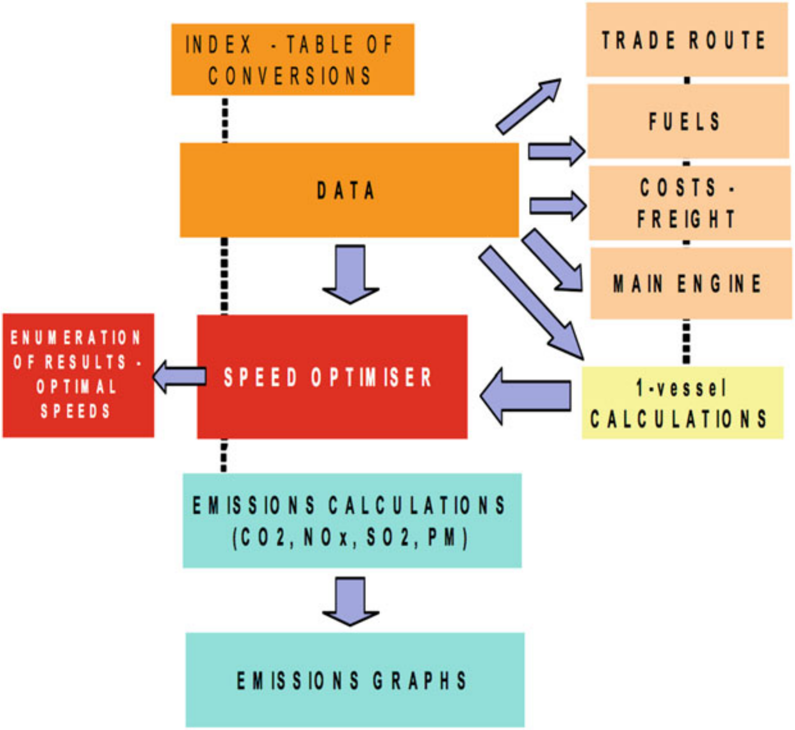


Fig. 11.5 Single tankers speed optimization tool structure. (Source: Gkonis and Psaraftis (2012))

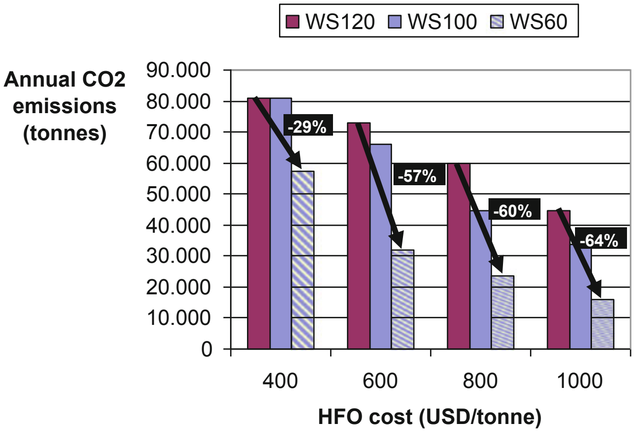


Fig. 11.6 Reduction in annual CO_2 emissions (single VLCC tanker) for a 50 % reduction in freight rates. (Source: Gkonis and Psaraftis (2012))

11.5.2 The Fleet Segment Tool

This tool has as main objective the calculation of emissions for the global fleet of the tanker segment in question (VLCC in our case). By fleet segment we mean the collection of ships that can be characterized as VLCCs, that is, tankers of 200,000 DWT and above. There were 525 such ships in the Lloyd's SeaWeb database in 2010. The tool works as follows.

- For the tanker segment in question, the "tanker segment emissions calculator" tool is run. This is a modified version of the single ship model, so that it does not refer to a specific route, but having as a constraint the annual tonne*miles throughput of a fleet of the typical ship considered.
- The laden & ballast leg speed input is taken from the single ship model.
- The output is annual emissions and operational characteristics (e.g. fuel consumption, operational days) for the tanker segment in question.

This model structurally resembles the previous one, but it is simpler as it does not perform any sort of optimization, see Fig. 11.7. It basically calculates emissions and related indices for the VLCC fleet segment.

As with the single ship tool, a variety of runs has been carried out. One is shown in Fig. 11.8 below.

In the above figure, Case 1 involves speed optimization only in the ballast leg, with the laden leg speed in the neighborhood of the recorded average service speed and assumed fixed in the charter party agreement. In Case 2, both speeds are free to be optimized. In all cases, upper and lower bounds on speeds are applicable. It can be seen that taking on board inventory costs generally reduces emissions for Case 1 (or leaves them constant) and increases emissions for Case 2. In all cases, Case 1 emits more than Case 2, which is something to be expected.

11.6 Multiple Optimal Speeds

If ship payload varies along the ship's route, optimizing ship speed at each leg of the route is better than finding a single optimal speed, the same for all legs.

Assume a cargo ship of lightship weight equal to $A = 6$ and capacity equal to $Q = 12$ (in thousands of tonnes), whose daily fuel consumption (in tonnes) is equal to $FC = kv^3(w + A)^{2/3}$, where v is the ship speed, w is the payload and k is a constant such that at full capacity and at a speed of 14 knots fuel consumption is 35 t/day. For simplicity also assume that the ship's maximum and minimum speeds are 16 and 9 knots respectively, and are independent of payload. Assume that P_{FUEL} = $ 800/t and that $\beta = 0$.

Consider a fixed-route scenario in which the above ship visits, in this order, ports 1–6.

Assume the ship starts empty at port 1 and has to collect cargo shipments of sizes 3, 2, 2 and 5 (in thousand tonnes) at ports 2,3,4 and 5 respectively, and deliver all of

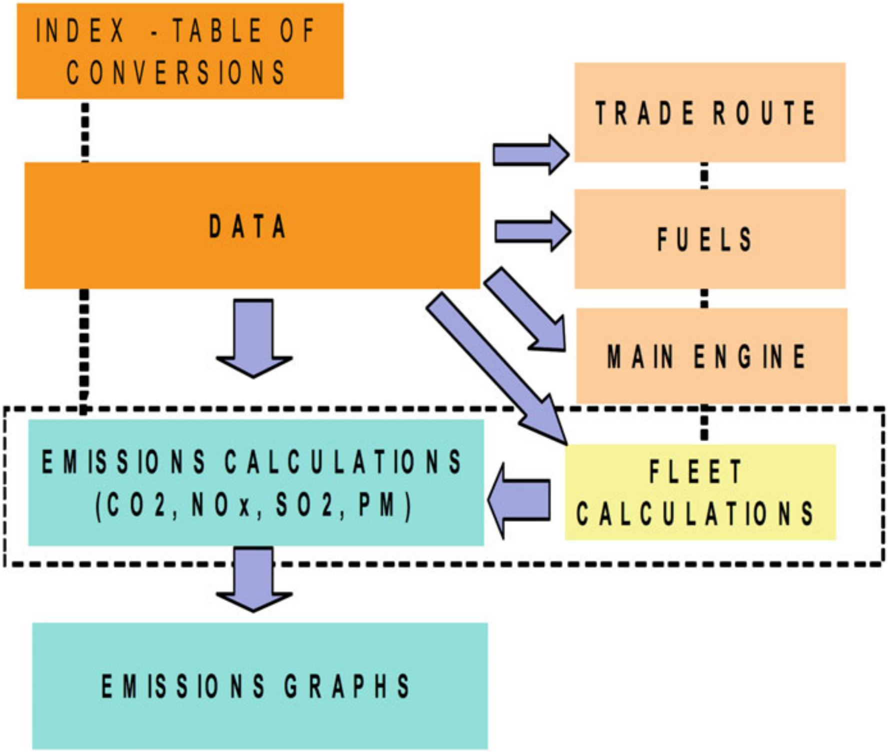


Fig. 11.7 "Fleet segment emissions calculator" tool structure. (Source: Gkonis and Psaraftis (2012))

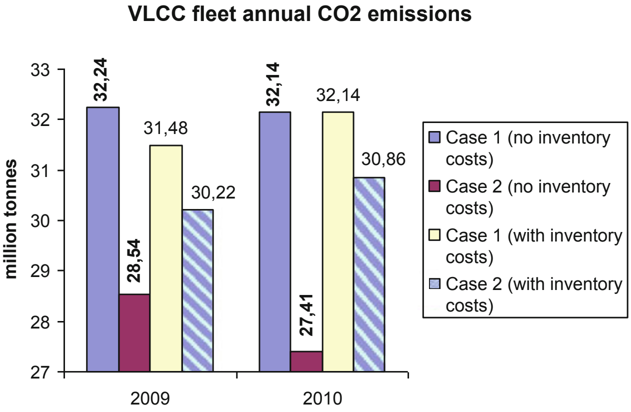


Fig. 11.8 VLCC fleet annual CO_2 emissions, with and without cargo inventory costs. (Source: Gkonis and Psaraftis (2012))

Table 11.2 Results of the variable speed per leg scenario

Leg	Distance (nm)	Speed (knots)	Fuel cost ($)	Charter cost ($)	Total cost ($)	Trip time (days)
1–2	400	13.17	9,494	18,989	28,483	1.27
2–3	150	12.03	3,896	7,792	11,688	0.52
3–4	200	11.51	5,432	10,863	16,295	0.72
4–5	250	11.09	7,046	14,093	21,139	0.94
5–6	300	10.31	9,090	18,179	27,269	1.21
TOTAL	1300	–	*34,958*	*69,916*	*104,874*	*4.66*

Table 11.3 Results of the fixed speed scenario

Leg	Distance (nm)	Speed (knots)	Fuel cost ($)	Charter cost ($)	Total cost ($)	Trip time (days)
1–2	400	11.541	7,296	21,661	28,957	1.44
2–3	150	11.541	3,585	8,123	11,708	0.54
3–4	200	11.541	5,465	10,830	16,295	0.72
4–5	250	11.541	7,636	13,538	21,174	0.90
5–6	300	11.541	11,383	16,245	27,628	1.08
TOTAL	1300	–	*35,365*	*70,397*	*105,762*	*4.69*

them to port 6. As before, assume that P_{FUEL} is $ 800/t, that F is $ 15,000/day and that port dwell times can be ignored.

Note that in this scenario, if the fuel consumption function is assumed independent of ship payload, the ship's optimal speed will be the same on each leg of the route. However, with a payload-dependent fuel consumption function, different speeds on each leg would generally be warranted. For the particular assumed fuel consumption function, the ship consumes about half the fuel in the ballast condition vis-a-vis that consumed in the fully laden condition if sailing at the same speed.

Table 11.2 shows the results of the variable speed scenario.

A pertinent question is, with the same fuel consumption function, if for whatever reason the ship is to keep the same speed along the route, can we at least find the common speed that minimizes total cost? It turns out that this speed is 11.541 knots. Table 11.3 shows detailed results of this scenario.

11.7 Slow Steaming and Ports

This section discusses the possible role of ports in slow steaming.

11.7.1 The Role of Ports in the Supply Chain

As mentioned earlier, speed reduction seems to be an easy fuel cost (and hence emission) reduction measure that can be implemented. A pertinent question is whether this is overall cost-effective or not. When talking about a single roundtrip, a delayed delivery of cargo will distort the current status-quo and may cause inventory costs to the shipper. In the case of containers and passenger vessels this may also lead to a modal shift (cargoes from sea to land) and may put the shipping company in an unfavorable competitive position. This is especially true for short sea shipping, but may also be true for longer distances, at least in principle. Although this may be far-fetched, one would not want to see, for instance, some cargoes from the Far East to Europe being shifted to rail, or (even worse) to road or air, as a result of a drastic speed reduction in the maritime mode. See Psaraftis and Kontovas (2010) for an example. Furthermore, to maintain constant annual throughput, in most of the cases, more ships will have to be used.

Psaraftis and Kontovas(2009b) investigated the simple scenario where a fleet of N identical ships (N: integer), each of capacity (payload) W loads from a port A (time in port $T_{A,}$, days), travels to port B with known speed V_1, discharges at B (time in port T_B, days) and goes back to port A in ballast, with speed V_2. The main result of the analysis was that total emissions would be always reduced by slowing down, even though more ships would be used. Psaraftis et al. (2009) focused on the case where total trip time was kept constant. Given the fact that time at sea increases with slow steaming, possible ways to decrease time in port were investigated. Emissions can be reduced even further if port time can be reduced so that there is no need for additional vessels. But this may be a more difficult proposition. For instance, in the example illustrated above, when speed is reduced by 5 %, time in port has to be reduced by 11 % to maintain a constant total trip time. If this sounds feasible, it is non-trivial nonetheless. For a speed reduction of 15 % the total time in port has to be reduced from 10.8 days down to 6.81, which is almost a 37 % reduction. This is a much more difficult proposition, possibly entailing drastic port re-engineering and/or infrastructure improvements. Obviously speed reductions of more than 5–10 % cannot be implemented without the need of adding more ships to maintain same service level.

Port time is only a small portion of the total turnaround time and reducing port time is not as easy as with time in sea. In the following section, a closer look at container terminal operations will be presented in order to identify areas that can be optimized so that we can easily implement speed reduction scenarios.

There has been tremendous growth in the worldwide container transshipment. Up to 2008, the top 10 container terminals in the world have shown an average relative increase of more than 10 % per year with respect to the total number of TEUs handled. However, in light of the severe negative ongoing economic crisis, the outlook for global container trade has darkened after mid 2008. The overall picture that emerges after the crisis is that while Asia continues to lead the global demand for container port services, growth is slowing.

While most of the top 10 ports experienced an increase in throughput during the last decade, there exists high competition among terminals even within the same geographical region, see for example the European ports of Antwerp, Rotterdam, Bremen and Hamburg and the Asian ports of Shanghai, Busan and Hong Kong. Seaports compete for ocean carrier patronage and feeders as well as land-based carriers (trucks and railways). According to Muller (1995) and Steenken et al. (2004), the most crucial competitive advantage of a terminal is the rapid turnover of container vessels, thus the key factors are low transshipment times combined with high rates for loading and unloading operations.

Usually container traffic networks overlay with each other and terminals can be part of more than one network. Containers can be transshipped between different modes of transport. Forwarding a container from a shipper to a recipient requires the use of one or more traffic networks and a transshipment of the container in a CT in the case where different transport vehicles are involved. The main purpose of seaport container terminals is to serve container vessels, put is simply, that is loading and unloading containers. Beside the large ocean-going container vessels, terminals also serve barges and feeder vessels. Barges are used for the container transport on inland waterways and feeder vessels connect ports with low transport volume or insufficient accessibility for large vessels to so-called hub ports.

According to Steenken et al. (2004) the four major areas of a seaport CT are:

Quay area for berthing container vessles

Transport area where internal transportation of containers takes place

Yard area where containers are transferred to and stored

Truck and train area for service the land-based vehicles

For more information on container terminal operations the reader is referred to Vis and de Koster (2003); Steenken et al.(2004) and Agerschou (1983, 2004), among others.

11.7.2 Bottlenecks in Container Operations

As discussed above, identifying ways to reduce time in port is crucial in order to be able to implement speed reduction measures that are necessary to reduce emissions. The tasks performed vis-a-vis the four major areas described above are definitely time consuming. In a liner service, vessels follow a predefined schedule that gives the order of ports to visit and the calling times. Which ports to connect on a route and how to design the liner network is part of the decision problem of the liner company. Furthermore, by deciding on the frequency of port calls within a schedule (e.g., on a weekly basis, or other) the number of vessels to deploy on a route is determined. More on this subject can be found in Agarwal and Ergun (2008); Notteboom and Vernimmen (2010); and Alvarez (2009) among others. A review of these problems is presented in Meng et al. (2013).

Assuming that the number of vessels deployed and the ports to be visited are known, the total time of the route can be estimated. Laine and Vepsalainen(1994)

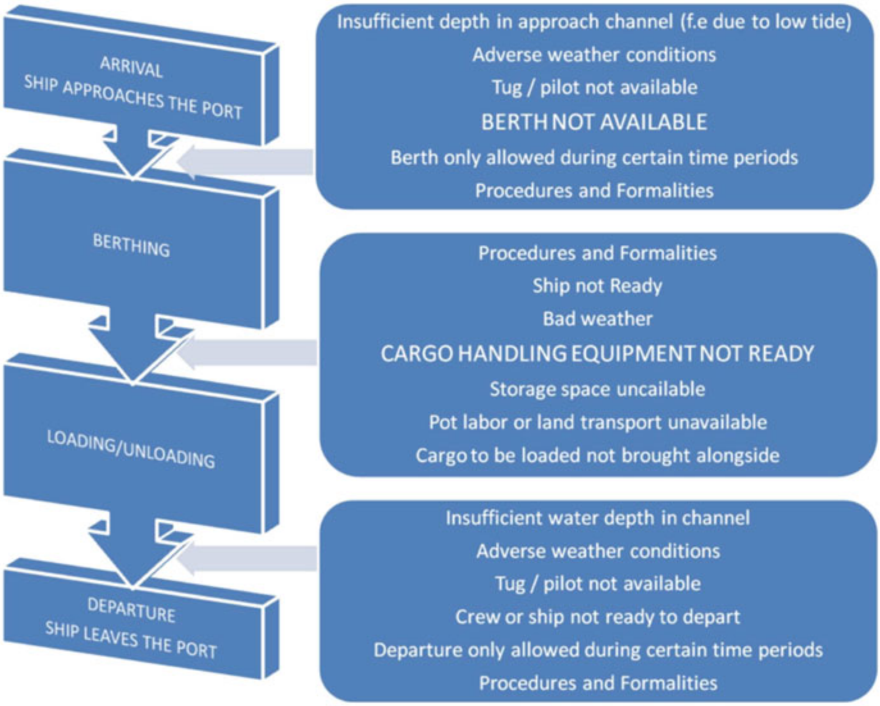


Fig. 11.9 Phases of operations and related problems

state that a high revenue can basically be obtained by an increase in the travel speed. However, they conclude that speed up of cargo handling seems to be more profitable compared to the increase in travel speed. Consequently, the performance of terminal operations is crucial for the profitability of liner services. This is in line with Notteboom (2006), who states that unexpected waiting times of vessels before berthing and unexpected low transshipment productivity at terminals are responsible for about 86 % of liner schedule disturbances.

In order to identify areas and operations that are time consuming, we shall graphically list the sequence of operations that usually take place from the time that a ship approaches a terminal until its departure. The list of procedures is mainly based on Agreschou (1981). There are four possible states for a vessel: arriving, berthing, loading and unloading and, finally, departing. When a vessel that carries cargo is approaching the port, she may berth immediately or wait for one of the reasons illustrated in Fig. 11.9, the most important of which is that the berth may not be available. After berthing, there are also several reasons that can prevent the immediate start of loading or unloading operations. One of them is related to the cargo handling equipment, which may be allocated elsewhere. Cargo handling equipment plays also an important role in the loading/unloading operation itself. Laine and Vepsalainen (1994) and Notteboom (2006) note that the most feasible way to reduce time in port is through operational decisions regarding quayside operations (berth allocation, quay cranes scheduling and vessel stowage).

11.7.3 The Time Factor in Port

Depending on the state of port congestion, the ship may or may not have to wait in an anchorage area outside the port, and the amount of total waiting time is t_W. After berthing, containers are unloaded/loaded from/onto the ship. Finally, when service is completed, the ship leaves the port. The amount of time that the vessels spends from berthing to departure is called service time (t_S).

According to Notteboom (2006), unexpected waiting times of vessels before berthing and unexpected low transshipment productivity at terminals are responsible for about 86 % of liner schedule disturbances. Furthermore, waiting time due to weather, procedure and formalities, canal transits etc will be assumed as constant before and after implementing a speed reduction measure and therefore they do not play any substantial role in the analysis.

We now perform a more detailed analysis of the total port time. In the scope of this work the main time components are the waiting time before berthing (t_w) and the service time (t_s). The average time that ships spend in port is defined as the sum of the average waiting time and average service time, i.e. $t_{ws} = t_w + t_s$.

First of all, the waiting time as seen by the view of port planners and constructors is analyzed. The occupancy rate of a group of berths expresses the percentage of time that berth positions are occupied by ships being serviced. The effect of berth occupancy on waiting time depends on the probability distributions of arrivals and of servicing times as well as on the number of berths available.

According to Tsinker (2004), the assumption that is usually made for container terminals is that the time intervals between successive vessel arrivals do not follow the negative exponential distribution applicable to general cargo terminals, but rather follow an Erlang distribution, with $K = 2$, because of the regularity of container ship arrivals. Furthermore, Agerschou (2004) states that arrival rates for container ports which are used by more than one shipping lines conform to Poisson distribution. He presents waiting time to service time ratios assuming $\mathrm{K} = 4$ and ∞ distributions for multi-user container terminal. The ratios are empirical values resulting from economic feasibility studies and are in general lower than those for general cargo ports due to the value of time for container ships. For example, for more than four berths the ratios for container terminals are 0.12 and 0.10 in the case of $\mathrm{K} = 4$ and $\mathrm{K} = 2\infty$ distributions respectively. Finally, Thoresen (2003) assumes a ratio of the average waiting time or congestion time to the average berth service rime of not higher than between 5–20 %.

On the other hand, Dragović et al. (2005) uses Queuing Theory (QT) models to analyse movements of ships in port. According to the authors and the studies that were analyzed the types ($M/M/n_b$) and ($M/E_k/n_b$) of queuing models were most practical to explain ship movements in port. Note that M denotes the Poisson distribution of arrivals. Furthermore, input data for both the simulation and analytical models were based on the actual ship arrivals at the Pusan East Container Terminal (PECT) for the six-month period from 6 September 2004 to 27 February 2005 -that is approximately 711 ship calls. The ships were categorized into the following three classes according to the number of lifts: under 500 lifts; 501–1000 lifts; and over 1000 lifts per ship. Ship arrival probabilities were as follows: 28.1 % for first class, 42.3 % for second and 29.6 % for third class of ships and an average ship arrival rate of $\lambda = 0.175$ ships/hour. The following table 11.4 presents the results of the work of Dragović et al. (2005) including the calculation of the waiting time to service time ratios for each ship category.

The above results are found to be in line with those of Agerschou (2004) and Thoresen (2003). Also, the berthing/unberthing time of ships was assumed to be one hour (Dragović et al. 2005). The total time at port is the sum of service time and waiting time plus this one hour. As a result, large container vessels (belonging to III class) spend at about 10 % of their total time in port waiting to occupy a berth. For all classes, the time waiting represents about 13 % of the total time in port.

11.7.4 Berthing Priority Policies: An Alternative Approach

Port managers in container terminals attempt to reduce costs by efficiently utilizing resources, including berths, yards, cranes, yard equipment and human personnel. Among all the resources, berths are the most important resources and good berth scheduling improves customers' satisfaction and increases port throughput, thus, leads to higher revenues. The usage of berths is scheduled by an intuitive trial-and-error method and varies from terminal to terminal.

The traditional birth allocation problems (BAP) focuses on the First-Come-First-Served (FCFC) policy. However, lately, many customers have contracts with the terminal operators that ensure them guaranteed berth-on-arrival (BOA) service—that is the actual berthing occurs within two hours of arrival. In this case, the objective of berth scheduling is to minimize the penalty cost resulting from delays in the departures of those vessels and the additional handling costs resulting from non-optimal locations of vessels. Carriers usually inform the terminal operator on the expected arrival time (ETA) and the requested departure time of vessels. Based on the information, the terminal operator tries to meet the requested departure time of all other vessels.

A related strategy is a policy in which a line could book a berthing time slot in advance and guaranteed service in that slot. In a seminal paper Psaraftis (1998) describes his experience from the real world when he was pulled out of the classroom and put in charge of the Piraeus Port Authority (PPA). Back in 1998, he proposed a scheduling berthing priority reform that he as a general manager of the port was thinking to implement. The original motivation was that this system would streamline utilization of cranes during peak periods and would effectively increase the capacity of the terminal. This scheme is referred to as "*Booking by rendez vous*".

Table 11.4 Average service and wait time ships—(Adopted from Dragović et al. 2005)

		(All classes)			(II class)			(III class)		
Results	ρ	t_s	t_w	*tw/ts*	t_s	t_w	*tw/ts*	t_s	t_w	*tw/ts*
Real data	0.643	15.200	2.443	0.161	13.550	2.470	0.182	22.990	2.451	0.107
Simulation	0.641	15.120	2.438	0.161	13.500	2.467	0.183	22.170	2.445	0.110
Analytical	0.720	16.080	3.211	0.200	13.210	3.306	0.250	21.920	5.033	0.230

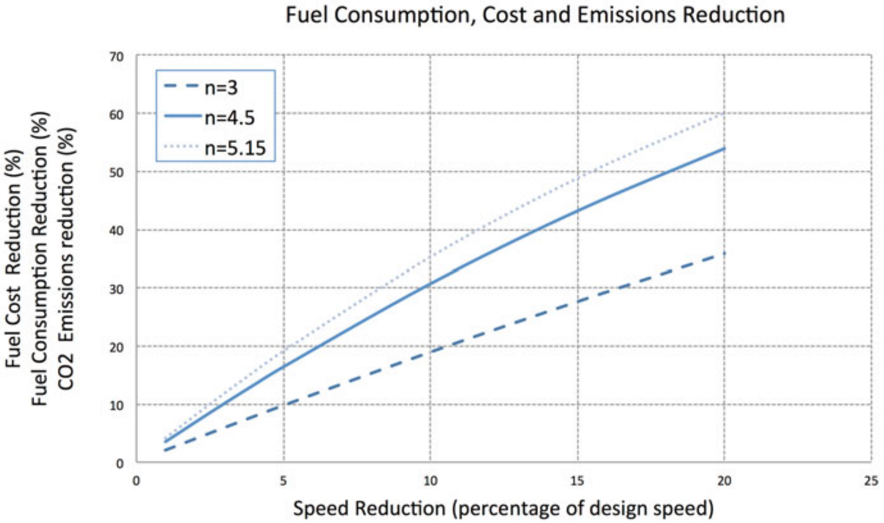


Fig. 11.10 Reduction percentage due to speed reduction

We present the following rudimentary example: A vessel employed in the Far East-Europe AE1 route served by CKYH, will decrease its speed in the Singapore-Rotterdam leg (that is the last asian port—first european port route). We also assume that the manoeuvering time and canal transit time (Suez) will be constant before and after the implementation of speed reduction[6].

The inputs are as follows:

Distance Singapore-Rotterdam: L = 8353 nm
Average Speed: Vo = 23 kn
Fuel consumption: at sea Fo = 150 t/day and in port f = 8.4 t/day and
Time in port of Rotterdam: to = 1.93 days

For reasons of simplicity we omit the detailed calculations and we illustrate the results in the following figures.

Figure 11.10 presents the percentage of reduction in fuel consumption and cost, and CO_2 emissions for the whole trip from Singapore to Rotterdam. As discussed in Sect. 3.2 the power requirement P is proportional to the speed V to the power of n. In the above figure the results are shown for $n = 3$ (cubic relation), $n = 4.5$ (according to MAN Diesel (2006)) and $n = 5.15$ as proposed by the regression analysis that we performed. In most of these cases, the implementation of speed reduction will lead in such an increase of total time that extra ships will be needed.

---

[6] Ships transiting Suez are grouped in convoys that transit the canal every several hours, therefore in practice this assumption may not always be correct.

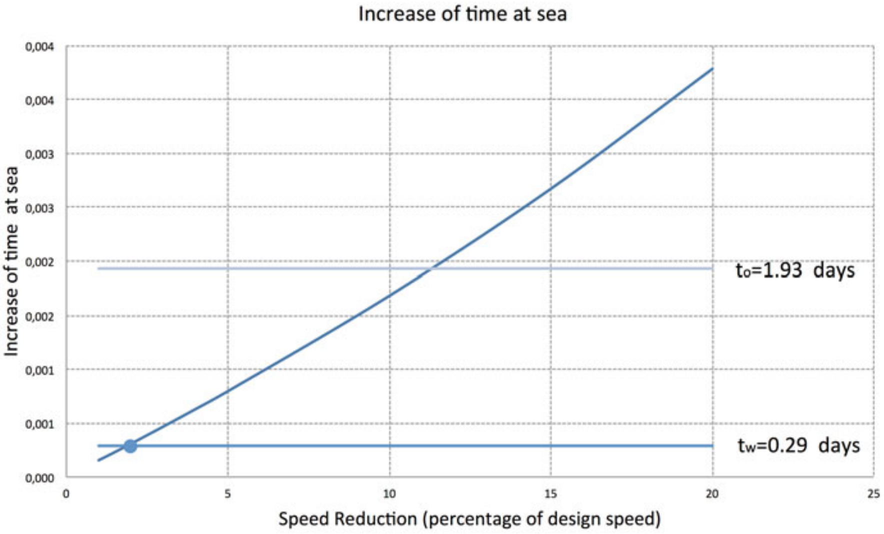


Fig. 11.11 Time implications of speed reduction

As discussed in the previous sections, a speed reduction will lead to an increase in the time at sea but some scenarios can be implemented without the need to add more ships to maintain the same throughput. The scenario of not adding extra vessels is the case when the total turnaround time can be kept constant. Here comes the role of the port in making this scenario feasible.

The waiting and service times for each port vary. For the sake of simplicity, suppose that in the same time we investigate the role of a single port in the route –that is of the first European destination, the Port of Rotterdam. We assume a port time (including service and waiting time) in the Port of Rotterdam equal to 1.93 days based on Notteboom and Vernimmen (2010). Optimizing the port operation in just this one port a speed reduction of 5 % sound realistic and can be implemented without the need to add extra ships.

Now, assume there is no congestion at the port and that service time is the same. By implementing a "booking by rendezvous" system, the vessel will berth as soon as it will arrive. No waiting time is equal to a reduction of 10–15 % of the total port time, that is $t_w = 0.193–0.29$ days (or about 4.6 to 7 h). In the best case scenario of the "booking by rendezvous" scheme, speed reductions of less than 2 % can be implemented without the need to add more ships (see Fig. 11.11). If this 2 % speed reduction sounds small, just imagine if we could introduce this scheme into every port. The potential savings in absolute numbers are great. Just a single figure: when fuel is expensive, let's say 600 $/t, a 5 % reduction saves a total of more than 200 thousand USD per route.

One initiative that is related to the scheme described above is the so-called 'Virtual Arrival' which has been employed firstly by tankers in order to manage the vessels' arrival time based on the experience of congestion at some discharging ports, see Kontovas and Psaraftis (2013) for more. This initiative recognises known inefficiencies in the supply chain, such as waiting to discharge because of port delays and reduces fuel consumption and, consequently, emissions by implementing a mutually-agreed reduction in a vessel's speed in order to achieve an agreed arrival time at a port. This scheme in order to work needs mutually agreement by both the owner and charterer to agree a speed to meet the terminal booking that maximises fuel efficiency and minimises port waiting time. To ensure the accuracy and independence of the calculations and to avoid the risk of disputes it is proposed to use a weather routing analysis company. After the agreement of both parties the ship slows to the economical speed based on the revised arrival time. Once the voyage is completed, demurrage is calculated based on the original plans and bunker savings are split between the parties.

'Virtual Arrival' seems profitable especially given the fact that there are indeed serious delays in discharging in some ports in the world. According to the Global Ports Congestion Index (GPCI) and its weekly newsletter publication that provides details on berthing delays at the major coal and ore ports worldwide, the average delay can be as high as 5 days. Obviously there is no point for vessels to steam at full speed when the have to wait a couple of days in order to discharge. Sailing at a slower speed and arriving on time entails benefits both for the owner and the charterer but also for the environment.

11.8 Combined Speed and Route Decisions

For pickup and delivery scenarios, the decomposition property of the optimal solution identified in sect. 4.1 is valid whatever the scenario and whetever algorithm is used to produce the ship's route, either exact or heuristic. Even if a heuristic algorithm is used, it would not make sense for the ship that sails a specific leg of the route to have an operating speed that is different from the optimal speed for that leg, as computed in sect. 4.1. A fortiori, the same is true if an exact algorithm is used.

An exact algorithm that can easily embed the above property, along with all other input parameters, and is flexible enough to allow for the spectrum of variants as regards the possible objective functions is one that is based on dynamic programming, and is a straightforward extension of the approach developed in Kontovas and Psaraftis (2011). That paper had assumed a general pick up and delivery setting with known and fixed arc traversal costs and times, zero dwell times and vehicle speed was not a decision variable. Further, there was no effect of F to be taken into account.

The approach goes as follows (see also Psaraftis and Kontovas (2014)).

Define the matrix $[k_{ij}]$ and optimal value function V as follows:

$$k_{ij} = \begin{cases} 3 \text{ if cargo from } i \text{ to } j \text{ has not been picked up yet} \\ 2 \text{ if cargo from } i \text{ to } j \text{ is on board the ship} \\ 1 \text{ if cargo from } i \text{ to } j \text{ has been delivered.} \end{cases}$$

$V(L,\ [k_{ij}])$ = Minimum possible total cost to complete the trip from port L to port 0, by executing all pending actions on pickup and delivery of the cargoes, choosing optimal speeds and observing capacity constraints, given that the current status of the cargoes is described by matrix $[k_{ij}]$.

Define $M =$ large number.
For a specific state $(L,\ [k_{ij}])$, define set $R = \{(i, j) :\ i \neq j,\ k_{ij} \neq 1\}$
If $R = \emptyset$, then

$$V(L, [k_{ij}]) = s_{L0}\, \underset{v \in S}{min} \left\{ \frac{P_{FUEL} f(v, 0) + F}{v} \right\}$$

(boundary condition: ship returning to home port)

$$\text{If } R \neq \emptyset, \text{then} V\left(L, \left[k_{ij}\right]\right) = \begin{cases} M \quad if\ w > Q \\ \min\limits_{(x,y)\hat{I}R} \left\{ s_{LL}C^* + \lambda d_{xy}(\alpha u + \beta w) + V\left(L^{'}, \left[k^{'}_{ij}\right]\right) \right\} \\ otherwise \text{ (11.1)} \end{cases} \tag{11.7}$$

where C^* is the optimal value of the optimization problem defined by

$$C^* = \underset{v \in S}{min} \left\{ \frac{P_{FUEL} f(v, w) + \alpha u + \beta w + F}{v} \right\}$$

with $S = \{v : v_{LB}(w) \leq v \leq v_{UB}(w)\}$

In the above recursion,

$$u = \sum_{(i,j):k_{ij}=3} d_{ij}$$

$$w = \sum_{(i,j):k_{ij}=2} d_{ij}$$

λd_{xy} is the port dwell time,

and for all pairs (i, j) with $i \neq j$, it is:

$$k^{'}_{ij} = \begin{cases} k_{ij} - 1\ if\ i = x\ and\ j = y \\ k_{ij} \quad otherwise \end{cases}$$

$$L^{'} = \begin{cases} x\ if\ k_{xy} = 3 \\ y\ \ if\ k_{xy} = 2 \end{cases}$$

To solve the problem, the recursion is executed backwards, by lexicographic ordering of the state variable vector and solving by moving to lexicographically increasing states. An alternative is to solve the recursion stage by stage, by defining an appropriate stage variable m as follows:

Ship is at port 0 (start), $m = 0$
Ship is port 0 (end), $m = 2n(n - 1) + 1$
Ship is at any intermediate port, $m = 3n(n - 1) - \sum_{\mathrm{(i,j):\ i \neq j}} \mathrm{kij}$

The stage-by-stage method is computationally more cumbersome than the lexicographic approach (in which m is not necessary) and as a result we have not used it. The algorithm was coded in Fortran 95 and implemented on a PC.

As in Kontovas and Psaraftis (2011), the computational effort of this method is as follows. Regarding memory, L grows as $O(n)$, and the number of possible combinations of values of the $[k]$ matrix is $O(3^{\mathrm{r}})$, where r is the number of non-zero O/D pairs, hence memory grows as $O(n3^{\mathrm{r}})$. For a complete graph, $r = n(n - 1)$. Each iteration of the recursion takes $O(r)$ time, bringing the total computational effort to $O(r^2 3^{\mathrm{r}})$. This can be as high as $O(n^4 3^{\mathrm{n2}})$ in the most general case. An exception is if both α and β are zero, in which case there are no summations to be taken. In this case the computational effort reduces to $O(n^2 3^r)$.

It can be seen that it is mainly r, the number of cargoes, rather than n, the number of ports, that dictates computational effort. Obviously such effort is on the high side for anything but small values of r, especially if matrix $[d]$ is complete. Lower computational times can be achieved in special cases, for instance in sparse graphs or for low values of Q.

Extensions The following extensions should be straighforward to implement:

1. Include cargo handling costs: Assuming that these are proportional to port dwell time, one can add a term equal to $C_p \lambda d_{xy}$ within the large bracket in recursion (1), where C_p is the per unit time cargo handling cost. However, as total port dwell time is proportional to total cargo volume and therefore fixed, adding this cost component will not change the optimal solution.
2. Include per call port costs: If there is a fixed component to port cost, say a per call cost of CC_K, then one can add this term within the large bracket in recursion (1), but only whenever $L' \neq L$, in the sense that this cost is accounted for only if the state transition involves moving from port to port. If $L' = L$, the state transition involves loading or unloading cargo and this cost should not be accounted for.
3. Include different loading and unloading rates λ, different cargo handling costs C_p and different per port call costs CC_K at each port.
4. Include different inventory coefficients α and β for each cargo, if for instance the cargoes have different values.
5. Last but not least, one can even include different fuel consumption functions for different legs of the route, say, due to different average weather conditions, sea currents, etc.

Table 11.5 Interport distances (in nautical miles)

i\j	0	1	2	3
0	–	255	175	10
1	255	–	200	250
2	175	200	–	170
3	10	250	170	–

Table 11.6 Cargo O/D matrix [d] (in thousand tonnes)

i\j	1	2	3
1	–	5	3
2	2	–	4
3	11	1	–

11.9 Sailing the Minimum Distance Route at Minimum Speed may not Minimize Fuel Costs

In the quest for reducing fuel costs but also obtaining environmentally optimal solutions, one might assume that if the minimum distance route is sailed at the minimum possible speed in all legs, this would minimize emissions. After all, daily emissions are an increasing function of ship speed, and more days at sea would seem to imply more emissions. However, it turns out that this is not necessarily the case, as shown in the rudimentary example below, involving a pickup and delivery scenario.

Assume a 4-port problem (the home port 0 plus 3 other ports) with the distance matrix given by Table 11.5 as follows:

Also assume an asymmetric O/D table for six (6) cargoes to be transported among ports 1–3 as given by Table 11.6:

We again assume the same feeder ship of the previous examples. The ship starts and ends at port 0, and has to visit the three ports as many times as necessary in order to carry all cargoes as shown in the O/D table. Note that one of the cargoes (from port 3 to port 1) is of size equal to the capacity of the ship. In this example we ignore cargo inventory costs, meaning that $\alpha = \beta = 0$.

If the objective is minimum trip time (this is achieved if we set $P_{FUEL} = 0$), all legs are sailed (as expected) at the maximum speed of 14 knots, and the ship makes a total of 6 port calls (once at port 2, twice at port 1 and three times at port 3) as follows (Table 11.7):

In Tables 7 to 12, by "Pxy" we mean "at port x pick up cargo destined to port y," and by "Dxy" we mean "at port y deliver cargo originating from port x."

In this case total distance traveled is also minimized and equal to 1,140 nautical miles, and total CO_2 emitted is 260 t. Total trip time is equal to 3.39 days. This solution is independent of F, so long as F is not zero.

At the other extreme of this example is if we examine the minimum emissions (or minimum fuel consumption) solution. We can do this by setting $F = 0$ and assuming any nonzero fuel price.

Table 11.7 Minimum trip time solution

Port stop	Pickup & delivery operations	Next Leg	Payload w at beginning of leg (000 t)
0	–	0–3	0
3	P31	3–1	11
1	D31, P12, P13	1–3	8
3	D13, P32	3–2	6
2	D12, D32, P21, P23	2–1	6
1	D21	1–3	4
3	D23	3–0	0
0	–	–	–

Table 11.8 Minimum emissions solution

Port stop	Pickup/ delivery operations	Next Leg	Payload w at beginning of leg (000 t)
0	–	0–3	0
3	P31	3–1	11
1	D31, P12	1–2	5
2	D12, P21	2–1	2
1	D21, P13	1–3	3
3	D13, P32	3–2	1
2	D32, P23	2–3	4
3	D23	3–0	0
0	–	–	–

If this is the case, the ship will make 7 port calls instead of 6 (twice at ports 1 and 2 and three times at port 3), and will sail all legs at the minimum speed of 8 knots. The solution will be as follows (Table 11.8):

Total distance traveled in this case will be 1260 nautical miles and total trip time will be 6.56 days, both higher than before. But total CO_2 emitted will only be 80 t, much lower. Obviously the lower emissions are mainly due to the lower speed. However, it is interesting to note that *the amount of CO_2 emitted in this case is lower than the 84.90 t of CO_2 that would be emitted if the ship had sailed the minimum distance route of Table 11.10 at the minimum speed of 8 knots*[7]. The reason that sailing the minimum distance route at minimum speed is suboptimal with respect to

---

[7] For a cubic fuel consumption function, total fuel consumed (and hence CO_2 produced) is proportional to the square of the speed, everything else (including payloads at each leg) being equal. $260(8/14)^2 = 84.90$.

Table 11.9 Solutions for non-zero fuel price and varying freight rates

Port stop	Pickup & delivery operations	Next Leg	Payload w at beginning of leg (000 t)	Optimal speed (knots)	
				F = $ 5,000/day	F = $ 20,000/day
0	–	0–3	0	9.39	14.00
3	P31	3–1	11	8.00	11.51
1	D31, P12, P13	1–3	8	8.00	12.05
3	D13, P32	3–2	6	8.00	12.51
2	D12, D32, P21, P23	2–1	6	8.00	12.51
1	D21	1–3	4	8.24	13.08
3	D23	3–0	0	9.39	14.00
0	–	–	–	–	–

emissions is that it involves more legs in which the ship is more laden as compared to the case it sails the alternate, longer route. A heavier load profile results in higher fuel consumption (and emissions) overall, even though the route is shorter. So in this case what would intuitively seem like an optimal policy is actually suboptimal.

Other solutions may be produced for different values of the input data. Table 11.9 shows two cases where $P_{FUEL} = \$\ 600/t$ (in both cases) and F is either $ 5,000/day or $ 20,000 day. Both cases produce the same optimal route, but speeds along the legs of the route will vary for different values of F.

As expected, the ship goes faster when F is higher, with the lower speed bound active in 4 legs of the $F = \$\ 5{,}000$/day case and the upper speed bound active in 2 legs of the $F = \$\ 20{,}000$/day case.

We mention that the above examples were solved by dynamic programming, in a straightforward extension of the algorithm of Kontovas and Psaraftis (2011), so as to embed speed optimization.

11.10 Policy Implications

In this section we discuss slow steaming from a policy perspective.

11.10.1 The Adoption of EEDI

Perhaps the most sweeping piece of regulation that will have an impact on ship speeds (and in fact at the strategic level) is the recent adoption of Energy Efficiency Design Index (EEDI) by the IMO. Indeed, after years of discussion and intensive and highly

Table 11.10 Parameters for determination of EEDI reference values for different ship types

Ship type	a	c
Bulk carrier	961.79	0.477
Gas carrier	1120.00	0.456
Tanker	1218.80	0.488
Container ship (65 % DWT)	186.52	0.200
General cargo ship	107.48	0.216
Reefer	227.01	0.244
Combination carrier	1219.0	0.488

political debate between developed and developing countries, the finalization of the regulatory text on the EEDI for new ships was agreed upon at the 62nd session of IMO's Marine Environment Protection Committee—MEPC 62 in July 2011.

For a given ship, the EEDI is provided by a complex formula, of which the numerator is a function of all power generated by the ship (main engine and auxiliaries), and the denominator is a product of the ship's deadweight (or payload) and the ship's 'reference speed', defined as the speed corresponding to 75 % of MCR, the maximum power of the ship's main engine. The units of EEDI are grams of CO_2 per tonne mile. The EEDI of a new ship is to be compared with the so-called "EEDI (reference line)," which is defined as EEDI (reference line) = $aDWT^{-c}$, where DWT is the deadweight of the ship and a and c are positive coefficients determined by regression from the world fleet database, per major ship category.

For a given ship, the attained EEDI value should be equal or less than the required EEDI value which is provided by the following formula.

$$\text{Attained EEDI} \leq \text{Required EEDI} = (1 - X/100)\text{aDWT}^{-c} \tag{11.8}$$

where X is a "reduction factor" specified for the required EEDI compared to the EEDI Reference line[8].

The reference line parameters a and c in (11.9), which have been finalized by regression analysis after a long debate within the IMO are presented in Table 11.10 below, although they are subject to revision.

It is interesting to note that Ro/ro vessels are thus far excluded from EEDI, because no adequate regression coefficients have been obtained for this class of vessels. This is an open subject that the IMO hopes to close in the foreseeable future.

A basic problem with EEDI is that compliance effectively imposes a limit on a ship's *design speed*, as the left-hand side of inequality (11.9) is a polynomial function of the design speed whereas the right-hand side is independent of speed. Thus,

---

[8] The values of X specified by the IMO are 0 % for ships built from 2013–2015, 10 % for ships built from 2016–2020, 20 % for ships built from 2020–2025 and 30 % for ships built from 2025–2030. This means that it will be more stringent to be EEDI compliant in the years ahead.

whereas the real goal of EEDI is to design ships with better hulls, engines and propellers so as to be more energy efficient, an easy solution might be to reduce design speed, and, as a consequence, installed power. This may have negative ramifications on ship safety. It may also have negative effects on total CO_2 emitted, as an undepowered ship would burn more fuel and hence emit more CO_2 at the same speed, particularly if it tries to maintain speed in bad weather.

11.10.2 Market Based Measures

A parallel effort at the IMO concerns the so-called Market Based Measures, or MBMs. MBMs are economic instruments that entice the ship owner to adopt measures to make the ship emit less CO_2. MBMs are also used to raise money to invest in carbon-reducing technologies outside the shipping sector.

At this point there are 10 distinct MBM proposals before the IMO. An Expert Group has been formed and some initial discussions have been held, but no final decision has been reached as of yet. These MBMs include a levy on fuel, an emissions trading scheme, various hybrid proposals based on EEDI, and others.

In terms of what has been described in this chapter, it is interesting to note that among the various MBM proposals, the levy proposal is perhaps the only one that can handle slow steaming automatically. In the short run, a levy on fuel would effectively raise the price of fuel and as a result would make the ship go slower. If the levy is equal to the social cost of CO_2, this would fully internalize its external cost. In the long run, the same measure would encourage a ship owner to invest in technologies that would make the ship burn less fuel. For an analysis of the MBMs of the table at the IMO, see IMO (2010) and Psaraftis (2012a).

11.10.3 Instituting Speed Limits?

Realizing that reducing speed also reduces emissions, some researchers and some lobbying groups have recommended instituting speed limits on shipping. Among the researchers, see Lindstad et al. (2011) for an argument and Cariou and Cheaitou (2012) for a counter argument. Among the lobbying groups, the Clean Shipping Coalition (CSC), a Non-Governmental Organization, advocated at IMO/MEPC 61 that *“speed reduction should be pursued as a regulatory option in its own right and not only as possible consequences of market-based instruments or the EEDI.”* However, that proposal was rejected by the IMO. In spite of this decision, lobbying for speed limits has continued by CSC and other groups.

It is clear that slow steaming and speed limits are two different things, as the first is a voluntary response and the second is an imposed measure. If the speed limit is above the optimal speed that is voluntarily chosen, then it is superfluous. If it is below, it may cause distortions in the market and costs that exceed the benefits of speed reduction. Possible side-effects include

- Building more ships to match demand throughput, with more CO_2 associated with shipbuilding and recycling
- Increasing cargo inventory costs due to delayed delivery
- Increasing freight rates due to a reduction in tonne-mile capacity
- Inducing reverse modal shifts to land-based modes (mainly road) that would increase the overall CO_2 level
- Implications on ship safety

It is clear that imposing speed limits, either on a global or on a regional level, is an emissions abatement measure that should be studied very carefully in terms of its possible side effects, as it is quite conceivable that its overall costs might exceed its benefits.

11.11 Conclusions

This chapter has examined the practice of slow steaming from various angles. In that context, a taxonomy of models was presented, some fundamentals were outlined, the main trade-offs were analysed, and some decision models were presented. Some examples were finally presented so as to highlight the main issues that are at play.

The chapter has confirmed that solutions for optimal environmental performance are not necessarily the same as those for optimal economic performance. Also policies that may seem at first glance optimal from an environmental viewpoint may actually be suboptimal. As a private operator would most certainly choose optimal economic performance as a criterion, if policy-makers want to influence the operator in his decision so as to achieve results that are good from a societal point of view, they could play with parameters that would internalize the external costs of CO_2 produced and move the solution closer to what is deemed more appropriate for the environment and for the benefit of society.

In the quest for a balanced economic and environmental performance of maritime transport, we think that this chapter can provide useful insights.

Acknowledgments Work on this chapter has been supported in part by various sources, including Det Norske Veritas in the context of a project at the National Technical University of Athens (NTUA), the authors' former affiliation, and the Lloyd's Register Foundation (LRF) in the context of the Centre of Excellence in Ship Total Energy-Emissions-Economy at NTUA. LRF helps to protect life and property by supporting engineering-related education, public engagement and the application of research. This work has also been supported in part by an internal grant at the Technical University of Denmark (DTU).

References

Agarwal, R., & Ergun Ö. (2008). Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science, 42*(2), 175–196.

Agerschou, H. (Ed.) (1983). *Planning and design of ports and marine terminals*, 1nd Edn. Chichester: Wiley.

Agerschou, H. (Ed.) (2004). *Planning and design of ports and marine terminals*, 2nd Edn. London: Thomas Telford.

Alderton, P. M. (1981). The optimum speed of ships. *Journal of Navigation, 34,* 341–355.

Alphaliner. (October 2013). Extra and super slow steaming help absorb 7.4 % of fleet. *Alphaliner Weekly Newsletter, 2013*(44).

Alvarez, J. (2009). Joint routing and deployment of a fleet of container vessels. *Maritime Economics & Logistics, 11,* 186–208.

Barrass, C. B. (2005). *Ship design and performance for masters and mates*. UK: Butterworth-Heinemann.

Bausch, D., Brown, G., & Ronen, D. (1998). Scheduling short-term marine transport of bulk products. *Maritime Policy and Management, 25,* 335–348.

Benford, H. (1981). A simple approach to fleet deployment. *Maritime Policy & Management, 8,* 223–228.

Bierwirth, C., & Meisel, F. 2010. A survey of berth allocation and quay crane scheduling problems in container terminals. *European Journal of Operational Research, 202,* 615–627.

Brown, G., Graves, G., & Ronen, D. (1987). Scheduling ocean transportation of crude oil. *Management Science, 33,* 335–346.

Cariou, P. (2011). Is slow steaming a sustainable means of reducing CO_2 emissions from container shipping? *Transportation Research Part D, 16*(3), 260–264.

Cariou, P., & Cheaitou, A. (2012). The effectiveness of a European speed limit versus an international bunker-levy to reduce CO_2 emissions from container shipping. *Transportation Research Part D, 17,* 116–123.

CBO. (2006). The Economic Costs of Disruptions in Container Shipments, U.S. Congress, Congressional Budget Office, Washington, DC.

Corbett, J., Wang, H., & Winebrake, J. (2010). The effectiveness and costs of speed reductions on emissions from international shipping. *Transportation Research Part D: Transport and Environment, 14,* 593–598.

Devanney, J. W. (2007). Solving Elastic Transportation Networks. Center for Tankship Excellence. www.c4tx.org. Accessed 30 Sept 2014.

Devanney, J. W. (2010). The Impact of Bunker Price on VLCC Spot Rates. Proceedings of the 3rd International Symposium on Ship Operations, Management and Economics. SNAME Greek Section, Athens, Greece, October.

Devanney, J. W. (2011). The Impact of Charter Party Speeds on CO_2 Emissions. Center for Tankship Excellence. www.c4tx.org. Accessed 30 Sept 2014.

Dragović, B., Park, N. K., Radmilović, Z. R., & Maraš, V. (2005). Simulation modeling of ship-berth link with priority service. *Maritime Economics & Logistics, 7*(4), 316–335.

Du, Y., Chen, Q., Quan, X., Long, L., & Fung, R. Y. K. (2011). Berth allocation considering fuel consumption and vessel emissions. *Transportation Research Part E, 47,* 1021–1037.

Eefsen, T., & Cerup-Simonsen, B. (2010). Speed, Carbon Emissions and Supply Chain in Container Shipping. Proceedings of the Annual Conference of the International Association of Maritime Economists, IAME 2010. Lisbon, Portugal, July.

Faber, J., Freund, M., Köpke, M., & Nelissen, D. (2010). Going Slow to Reduce Emissions: can the current surplus of maritime transport capacity be turned into an opportunity to reduce GHG emissions? Seas at Risk.

Fagerholt, K. (2001). Ship scheduling with soft time windows—an optimization based approach. *European Journal of Operational Research, 131,* 559–571.

Fagerholt, K., & Ronen, D. (2013). Bulk ship routing and scheduling: Solving practical problems may provide better results. *Maritime Policy and Management, 40*(1), 48–64.

Fagerholt, K., Laporte, G., & Norstad, I. (2010). Reducing fuel emissions by optimizing speed on shipping routes. *Journal of the Operational Research Society, 61,* 523–529.

Gelareh, S., & Meng, Q. (2010). A novel modeling approach for the fleet deployment problem within a short-term planning horizon. *Transportation Research E, 46,* 76–89.

Gkonis, K. G., & Psaraftis, H. N. (2012). Modelling tankers' optimal speed and emissions. Proceedings SNAME 2012 Annual Meeting, Providence, RI, October.

Halvorsen-Weare, E. E., & Fagerholt, K.. (2010). Routing and scheduling in a liquefied natural gas shipping problem with inventory and berth constraints. Annals of Operations Research. doi:10.1007/s10479-010-0794-y.

Hvattum, L. M., Norstad, I., Fagerholt, K., & Laporte, G.. (2012). Analysis of an exact algorithm for the vessel speed optimization problem. Networks, in press.

IMO. (2010). Full report of the work undertaken by the Expert Group on Feasibility Study and Impact Assessment of possible Market-based Measures—doc. MEPC 61/INF.2, International Maritime Organization (IMO). London, UK.

Journal of Commerce. (2010). CKYH Begins Slow Steaming Eastbound Europe-Asia, Journal of Commerce. http://www.joc.com/maritime. Accessed 26 Jan 2012.

Kontovas, C. A., & Psaraftis, H. N. (2011). Reduction of emissions along the maritime intermodal container chain. *Maritime Policy & Management, 38*(4), 455–473.

Laine, J. T., & Vepsalainen, A. P. J. (1994). Economies of speed in sea transportation. *International Journal of Physical Distribution and Logistics Management, 24*(8), 33–41.

Lang, N., & Veenstra, A. (2010). A quantitative analysis of container vessel arrival planning strategies. *OR Spectrum, 32,* 477–499.

Lindstad, H.,/ Asbjørnslett, B. E., & Strømman, A. H. (2011). Reductions in greenhouse gas emissions and cost by shipping at lower speeds. *Energy Policy, 39*(6), 3456–3464.

Lloyds, L. (2009). CKYH carriers agree to super-slow steaming. Lloyds List, 16 November.

Lo, H. K., & McCord, M. R. (1998). Adaptive ship routing through stochastic ocean currents: General formulations and empirical results. *Transportation Research Part A, 32,* 547–561.

Maersk (2013). Building the World's Biggest Ship. http://www.maersk.com/innovation/leadingthroughinnovation/pages/buildingtheworldsbiggestship.aspx. Accessed 30 Sept 2014.

MAN DIESEL. (2006). Low Container Ship Speed Facilitated by Versatile ME/ME-C Engines. MAN Diesel A/S Report, Copenhagen, Denmark.

Meng, Q., & Wang, S. (2011). Optimal operating strategy for a long-haul liner service route. *European Journal of Operational Research, 215,* 105–114.

Meng, Q., Wang, S., Andersson, H., & Thun, K., (2013). Containership routing and scheduling in liner shipping: overview and future research directions. *Transportation Science, 48,* 265–280.

Muller, G. (1995). *Intermodal freight transportation* (3rd edn.). Intermodal Association of North America, EnoTransportation Foundation, Inc.

Norstad, I., Fagerholt, K., & Laporte, G. (2011). Tramp ship routing and scheduling with speed optimization. *Transportation Research Part C, 19,* 853–865.

Notteboom, T. E. (2006). The time factor in liner shipping services. *Maritime Economics & Logistics, 8,* 19–39.

Notteboom, T. E., & Vernimmen, B. (2010). The effect of high fuel costs on liner service configuration in container shipping. *Journal of Transport Geography, 17,* 325–337.

Papadakis, N. A., & Perakis, A. N. (1989). A nonlinear approach to multiorigin, multidestination fleet deployment problem. *Naval Research Logistics, 36,* 515–528.

Perakis, A. N. (1985). A second look at fleet deployment. *Maritime Policy & Management, 12,* 209–214.

Perakis, A. N., & Jaramillo, D. I. (1991). Fleet deployment optimization for liner shipping. Part 1: Background, problem formulation and solution approaches. *Maritime Policy and Management, 18*(3), 183–200.

Perakis, A. N., & Papadakis, N. A. (1987a). Fleet deployment models, part 1. *Maritime Policy & Management, 14,* 127–144.

Perakis, A. N., & Papadakis, N. A. (1987b). Fleet deployment models part 2. *Maritime Policy & Management, 14,* 145–155.

Perakis, A. N., & Papadakis, N. A. (1989). Minimal time vessel routing in a time-dependent environment. *Transportation Science, 23*(4), 266–277.

Psaraftis, H. N. 1998. When a port calls an operations researcher answers. OR/MS Today. pp 38–41.

Psaraftis, H. N. (2012a). Market based measures for greenhouse gas emissions from ships. *WMU Journal of Maritime Affairs*. doi:10.1007/s13437-012-0030-5.

Psaraftis, H. N., & Kontovas, C. A. (2009a). Green Maritime Logistics: Cost-effectiveness of Speed Reductions and other Emissions Reduction Measures, MLOG 2009. The International Symposium on Maritime Logistics and Supply Chain Systems 2009, Singapore, April 23–24.

Psaraftis, H. N., & Kontovas, C. A. (2009b) Ship Emissions: Logistics and Other Tradeoffs. 10th International Marine Design Conference (IMDC 2009), Trondheim, Norway, 26–29 May.

Psaraftis, H. N., Kontovas, C. A., & Kakalis, N. (2009). Cost-effectiveness of Speed Reduction as an Emissions Reduction Measure for Fast Ships. Proceedings of the 10th International Conference on Fast Sea Transportation, Athens, Greece, October 5–8.

Psaraftis, H. N., & Kontovas, C. A. (2010). Balancing the economic and environmental performance of maritime transportation. *Transportation Research Part D, 15*(8), 458–462.

Psaraftis, H. N., & Kontovas, C. A. (2013). Speed models for energy-efficient maritime transportation: A taxomomy and survey. *Transportation Research Part C, 26,* 331–351.

Psaraftis, H. N., & Kontovas, C. A. (2014). Ship speed optimization: Concepts, models and combined speed-routing scenarios. *Transportation Research Part C, 44,* 52–69.

Qi, X., & Song, D.-P. (2012). Minimizing fuel emissions by optimizing vessel schedules in liner shipping with uncertain port times. *Transportation research part E. Logistics and Transportation Review, 48*(4), 863–880.

Ronen, D. (1982). The effect of oil price on the optimal speed of ships. *Journal of the Operational Research Society, 33,* 1035–1040.

Ronen, D. (2011). The effect of oil price on containership speed and fleet size. *Journal of the Operational Research Society, 62*(1), 211–216.

Steenken, D., Voß, S., & Stahlbock, R. (2004). Container terminal operation research – a classification and literature review. *OR Spectrum, 26*(1), 3–49.

Stopford, M. (2004). *Maritime economics*, 2nd Edn. London: Routledge. (Taylor and Francis, and New York).

Thoresen, C. (2003). *Port designer's handbook*. London: Thomas Telford.

Tirado, G., Hvattum, L. M., Fagerholt, K., Cordeau, J.-F. (2012). Heuristics for dynamic and stochastic routing in industrial shipping. Computers & Operations Research. (Forth coming, Available online 5 July 2012)

TradeWinds. (2009). Maersk insists on slow speeds. TradeWinds magazine, 30 October.

TradeWinds. (2010). Slow spur for Maersk VLCCs. TradeWinds magazine, 13 December.

Tsinker, G. P. (Ed.). (2004). *Port engineering: Planning, construction, maintenance, and security*. Hoboken: Wiley.

Vis, I. F. A., & de Koster, R. (2003). Transshipment of containers at a container terminal: An overview. *European Journal of Operational Research, 147,* 1–16.

Wang, S., & Meng, Q. (2012a). Sailing speed optimization for container ships in a liner shipping network. *Transportation Research Part E, 48*(3), 701–714.

Wang, S., & Meng, Q. (2012b). Liner ship route schedule design with sea contingency time and port time uncertainty. *Transportation Research Part B, 46*(5), 615–633.

Wang, S., & Meng, Q. (2012c). Robust schedule design for liner shipping services. *Transportation Research Part E, 48,* 1093–1106.

Wang, S., Meng, Q., & Liu, Z. (2013). A note on "Berth allocation considering fuel consumption and vessel emissions". *Transportation Research Part E, 49,* 48–54.

Yao, Z., Ng, S. H., & Lee, L. H. (2012). A study on bunker fuel management for the shipping liner services. *Computers & Operations Research, 39*(5), 1160–1172.

Chapter 12
Efficient Global Container Transport Network Design

Shuaian Wang and Zhiyuan Liu

Abstract This chapter gives a comprehensive overview of existing research works on global container transport network design. In view of competitive pressures and complexity of container shipping networks, global liner shipping companies are seeking optimization-based decision support tools for designing efficient container transport networks. Nevertheless, the status quo research is far lagging, especially in terms of solving practical-size problems while capturing essential operating characteristics, developing efficient solution algorithms, and application by liner shipping companies.

12.1 Introduction

12.1.1 Container Trade

Maritime transportation is the backbone of world trade: around 80 % of global trade by volume and over 70 % by value are carried by sea (UNCTAD 2012). There are three modes of operations in shipping: industrial, tramp, and liner. In industrial shipping, cargo owners control the ships and seek to transport their cargo at minimal cost. In tramp shipping, the ship operator does not own the cargo, but selects available cargoes to transport so as to maximize its revenue. Liner shipping services are similar to bus services: both have fixed ports of call (bus stops) and transport containers (passengers) from many origins and destinations. Liner shipping mainly involves the transportation of containerized cargos. In fact, among all the sea cargos, 52 % in monetary terms are containerized (UNCTAD 2012). Containerized cargos, such as electronics, appliances, furniture, garments, auto parts and toys, generally have much

---

S. Wang (✉)
Strome College of Business, Old Dominion University,
Norfolk, VA 23529, USA
e-mail: wangshuaian@gmail.com

Z. Liu
Institute of Transport Studies, Department of Civil Engineering, Monash University,
Clayton, VIC 3800, Australia
e-mail: zhiyuan.liu@monash.edu

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_12

Table 12.1 Top 20 container ports in the world and their throughputs in 2012. (World Shipping Council 2014)

Rank	Port	Country	Volume (million TEUs)
1	Shanghai	China	32.53
2	Singapore	Singapore	31.65
3	Hong Kong	Hong Kong, China	23.10
4	Shenzhen	China	22.94
5	Busan	South Korea	17.04
6	Ningbo-Zhoushan	China	16.83
7	Guangzhou	China	14.74
8	Qingdao	China	14.50
9	Jebel Ali	United Arab Emirates	13.30
10	Tianjin	China	12.30
11	Rotterdam	Netherlands	11.87
12	Port Kelang	Malaysia	10.00
13	Kaohsiung	Taiwan Province of China	9.78
14	Hamburg	Germany	8.86
15	Antwerp	Belgium	8.64
16	Los Angeles	U.S.A.	8.08
17	Dalian	China	8.06
18	Keihin ports	Japan	7.85
19	Tanjung Pelepas	Malaysia	7.70
20	Xiamen	China	7.20

higher unit values than other sea cargos. The total container trade volume in 2011 amounted in 151 million twenty-foot equivalent units (TEUs) (UNCTAD 2012). The top 20 container ports in terms of throughput in 2012 are shown in Table 12.1. The throughputs of Shanghai and Singapore were more than 30 million TEUs, and the throughputs of Hong Kong and Shenzhen were higher than 20 million TEUs. The top 20 liner shipping companies in terms of total shipboard capacity deployed (TEUs) on 2 April 2014 are shown in Table 12.2. The shipboard capacities of APM-Maersk and MSC were both larger than 2 million TEUs and the shipboard capacity of CMA CGM was larger than 1 million TEUs.

12.1.2 Containers

There are many types of containers. First, dry containers are used for carrying cargo without special needs, and reefer containers are refrigerated for carrying fresh food

Table 12.2 Top 20 liner shipping companies in terms of TEU capacity available on board operated ships on 2 April 2014. (Alphaliner 2014)

Rank	Liner operator	Country	TEUs
1	APM-Maersk	Denmark	2,608,176
2	MSC	Switzerland	2,424,070
3	CMA CGM group	France	1,495,139
4	Evergreen line	Taiwan Province of China	874,892
5	COSCO container line	China	766,094
6	Hapag-Lloyd	Germany	737,767
7	APL	Singapore	663,999
8	Hanjin shipping	Korea	605,113
9	CSCL	China	601,174
10	MOL	Japan	558,113
11	NYK line	Japan	475,167
12	Hamburg Süd group	Germany	474,077
13	OOCL	Hong Kong, China	451,442
14	Yang Ming marine transport corp	Taiwan Province of China	384,320
15	Hyundai merchant marine	Korea	364,584
16	PIL	Singapore	357,053
17	K line	Japan	352,294
18	Zim	Israel	330,343
19	UASC	United Arabic Emirates	275,834
20	CSAV group	Chile	249,732

such as sea food and bananas. Since reefer containers need special equipment to maintain a low temperature and must be airtight, the cost of a reefer container is much higher than a dry container of the same size. Moreover, in transport, reefer containers need to use electricity generated by the ship engine, and hence the freight rate charged for transporting a reefer container is much higher than a dry container of the same size. Second, contains have different sizes, 20-foot long (TEU), 40-foot long (FEU), 45-foot long, etc. An FEU slot on a ship can accommodate 2 TEUs. Therefore, it is common to use TEU as the unit for containers and 1 FEU is equivalent to 2 TEUs. Other than TEU and FEU, it is not quite common to see other container sizes in real world.

How many containers a ship can transport mainly depends on three factors. The first factor is the number of container slots on the ship. We can refer to it as the volume capacity. The second factor is the total weight of all containers, which must be smaller than the weight capacity of a ship. It should be noted that reefer containers are generally much heavier than dry containers, because the cargo in reefer containers (e.g., sea food) usually has a higher unit weight than cargo in dry containers (e.g., garment). The last factor is container storage plan. The container storage plan is about where to store each container on a ship. There are a few rules for determining the storage plan. For instance, heavier containers should be at the bottom, the weight of the ship has to be balanced, and containers to be discharged at the next port of call should be placed on the top to minimize reshuffling. The detailed container storage plan is too detailed to be incorporated in liner shipping service design.

Compared to conventional cargo units, such as boxes, pallets and cartons, containers have a few significant advantages than enable the rapid growth of containerization. The most important advantage is the reduction of cargo handling time at sea ports. A quay crane can move 20–30 TEUs per hour and a large containership can accommodate four quay cranes at the same time. As a consequence, a ship only spends 1 day for loading or discharging 2000 TEUs at a port. Before containerization, a ship may spend 2 months at a port for cargo handling, which significantly reduces the productivity of ships. For instance, assume that a ship can carry 30,000 tons of cargo and needs 40 days to sail between two ports. If the cargo handling time at a port is 30 days, the ship can transport a total of $365/(30+40) \times 30,000 = 156{,}428$ tons of cargo each year. If the cargo handling time at a port is 1 day, the ship can transport a total of $365/(1+40) \times 30{,}000 = 267{,}073$ tons of cargo each year. The second advantage is intermodalization. Besides dedicated cellular gearless container ships, dedicated trucks and trains are also available for transporting containers in inland areas. As a consequence, it is very convenient to move a container of cargo between ships, trains, and trucks. The last advantage is the reduction of cargo damage and pilferage. Consequently, more and more cargos are being containerized.

12.1.3 Liner Shipping Services

Containers are usually transported by liner shipping services with fixed sequences of ports of call at a regular service frequency, which are published by liner shipping companies in their websites in advance to attract cargoes of shippers. Shippers or freight forwarders can pick up and deliver their cargos at any port covered by the liner services. A single shipper usually has far less than a full shipload of cargo whereas containerships have to adhere to their published departure dates even when a full payload is not available.

One can appreciate the liner shipping services by likening them to bus services. First, bus services have fixed routes and visit fixed sequences of bus stops. Liner shipping services also have fixed routes and visit fixed sequences of ports. Buses transport many passengers at the same time (in contrast to taxis), and container ships also transport containers from many customers at the same time. At each bus stop, there are passengers alighting and boarding, and at each port there are containers discharged and loaded.

There are also some differences between liner shipping services and bus services. First, liner shipping services are usually weekly, and therefore the arrival day at each port of call is published to facilitate inland transportation. For instance, suppose that ships visit Pusan on Thursday. If the cargos from Seoul missed the planned ship, they have to wait until next Thursday. By contrast, bus services in metropolitans have a frequency of e.g. 5–15 min. Therefore, it does not matter much to miss one bus. Moreover, buses visit many bus stops and are subject to the uncertain traffic conditions in the city. Hence, the uncertainties of the arrival time may be larger than the average waiting time due to the regular service frequency. Consequently, many bus services do not publish the planned arrival time at each bus stop. Suburban bus services are much more infrequent, e.g., with the frequency of 30 min or 1 h. As a result, many suburban bus services have the schedule of the arrival time at each bus stop. Second, liner shipping services transport containers and bus services transport customers. Passengers choose their own routes from origin to destination, whereas containers can be transported in a system optimal manner. Third, in bus services, the boarding, alighting, and transfer of passengers do not incur cost. However, in liner shipping, loading, discharging and transshipment of containers are very expensive and they are the main revenue for port operators. Fourth, buses usually operate in the day time, whereas containerships operate 24 h a day and 7 days a week. Fifth, bus driver scheduling is a challenging problem for bus companies as each bus must have a driver and drivers have work shifts. By contrast, the crew of a containership works on the ship for a long period of time. There are also other differences between liner shipping services and bus services.

12.1.4 Liner Service Design

Once a liner shipping network is designed, the bunker cost, capital cost, port cost, canal dues, crew cost and consumables are determined. The bunker cost is related to the consumption of bunker. The daily bunker consumption is approximately proportional to the speed cubed (Qi and Song 2012; Wang and Meng 2012b). Hence, to save bunker, containerships generally do not sail at the highest speed. In fact, the bunker cost constitutes a large proportion of the total cost due to the high bunker price in the recent years. Ronen (2011) estimated that when bunker fuel price is around 500 USD per ton the bunker cost constitutes about three quarters of the operating cost of a large containership. The capital cost of a ship can be considered as the ship chartering cost. The port cost usually includes the cost for pilotage and towage to and from the port, the cost of mooring and unmooring, and the cost per visit associated with the tonnage of the ship. Container handling cost and container storage cost are generally excluded from port cost because they occur at the operating level. The canal dues refer to the charges if the ship transits via the Suez Canal or the Panama Canal. Note that the Suez Canal can accommodate larger ships than the Panama Canal. Hence, for example, containers from Shanghai to New York/New Jersey have the following choices. (i) 8000-TEU ships transport the containers from Shanghai, via the Malacca Strait, Suez Canal, the Atlantic Ocean, and to New York/New Jersey. (ii) 4000-TEU ships transport the containers from Shanghai, via the Pacific Ocean, Panama Canal and to New York/New Jersey. (iii) 8000-TEU ships transport the containers from

Shanghai, via the Pacific Ocean, and to Balboa. At Balboa, the containers are discharged and reloaded to 4000-TEU ships that sail via the Panama Canal, Caribbean Sea, to New York/New Jersey. The crew cost and consumables are generally fixed for each ship. Since most of the cost components are determined at the network design stage, it is vital for a liner shipping company to design an efficient global container transport network.

Liner shipping network design can occur at the strategic level or the tactical level. At the strategic level, network design refers to the creation of a shipping network (port rotations, ship deployment, frequency and schedule) from scratch. This may happen if a new liner shipping company is established or if a liner shipping company restructures its overall network. A company may restructure its overall network when a large fleet of newly booked ships is delivered, when it is merged with another shipping company, or when the business environment has dramatically changed. At the tactical level, network design is the alteration of the existing network in a local sense. For instance, the addition of a new service to cope with the increased demand before Christmas, the removal of a port of call from a service because its demand is not sufficient to sustain a direct call, and the restructuring of the feeder network of a particular hub are all possible network alterations. In reality, network design at the tactical level (network alteration) is much more frequent than the strategic network design. However, almost all existing studies examine strategic network design rather than network alteration. There may be attributed to two reasons. First, strategic network design problems are better-structured. By contrast, in network alteration one needs to define to what extent the network is allowed to be changed. For example, Wang and Meng (2013) investigated a network alteration problem by optimizing the port rotation directions of services in a liner shipping network. The network can be changed by reversing the port rotation directions of at most a given number of services. Second, in general, strategic network design is more difficult and the methods for strategic network design may also be applied to some network alteration problems by fixing some decision variables.

12.1.5 Objectives and Organization

The objective of this chapter is to give an overview on studies devoted to global container transport network design and point out future research directions. We stress that we focus on the maritime transportation network although the origins and destinations of containers are usually located in inland (Meng et al. 2012a; Yang et al. 2012). Readers can refer to the three review papers by Christiansen et al. (2004, 2013) and Meng et al. (2014) for discussions about other topics in liner shipping operations.

This chapter is organized as follows. The next section discusses the liner shipping network. The section that follows introduces container shipment demand and container routing. We subsequently discuss the assumptions in network design models. After that, special network design problems are reviewed. We then formulate the mathematical models for general network design problems. The last section points out future research directions.

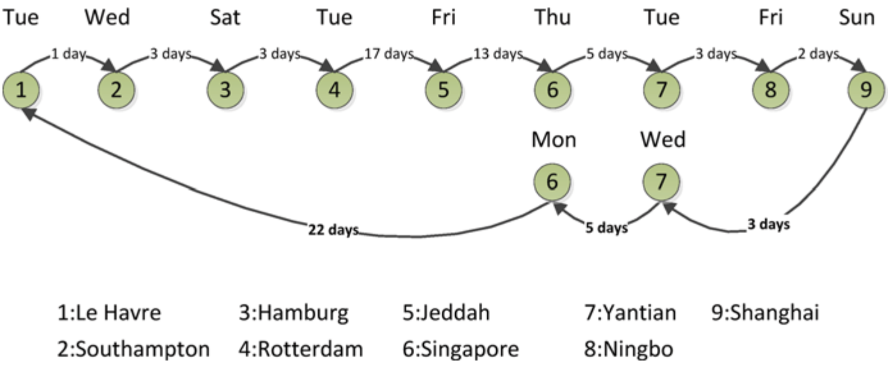


Fig. 12.1 Representation of Asia-Europe Loop 4

12.2 Container Liner Shipping Network

12.2.1 A Single Service

A liner shipping network consists of many services (ship routes). A service has fixed port rotation and schedule. For example, Fig. 12.1 give port rotation (sequence of ports of call) and transit times of APL's current Asia-Europe Loop 4 service route (APL 2013). This ship route is divided into east bound and west bound directions for commercial purposes with some overlaps in Fig. 12.2. Ships depart from each port of call on a given day of a week. In other words, the port rotation of the ship route forms a loop and the ship route has a weekly service frequency. The round-trip time of the Asia-Europe Loop 4 is equal to 77 days. The number of ships serving this ship route, denoted by m, can be calculated as follows:

$$m = \frac{77 \text{ days}}{7 \text{ days}} = 11 \tag{12.1}$$

Therefore, a ship route has the following information: port rotation, service frequency, schedule (arrival and departure time at each port of call, and transit time between two consecutive ports of call). The service frequency and schedule determine the number of ships to deploy similar to Eq. (12.1). It should be noted that a ship route also has the information of deployed ships (in particular, the sizes of the deployed ships). For example, the website of Maersk Line (2013b) has detailed information about the schedule of each ship. Although which ships are deployed on the service is not included in Fig. 12.1 and Fig. 12.2, such information is available (we conjecture that APL does not publish the information because customers are not concerned about it. For instance, many passengers do not care about which aircraft will be operated on their flight).

Most ship routes provided by global shipping companies have a weekly frequency. Some feeder services may have a twice or thrice weekly frequency. Note that the latter could be transformed to a weekly frequency in modeling. For instance, a service

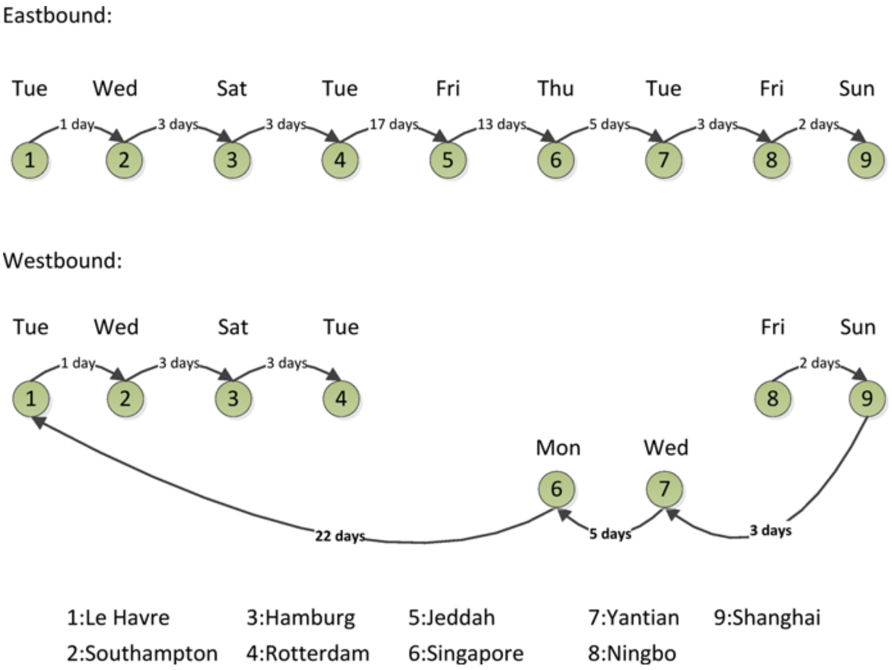


Fig. 12.2 East bound and west bound of Asia-Europe Loop 4

with a twice weekly frequency and port rotation Singapore → Jakarta → (return to) Singapore is equivalent to a weekly service with port rotation Singapore → Jakarta → Singapore → Jakarta → (return to) Singapore. Another typical exception is the daily service between Asia and Europe launched by Maersk Line (2013a) in 2012. Nevertheless, weekly frequency is still the convention and most mathematical models require at least a weekly frequency for all ship routes.

As shown by Eq. (12.1), a weekly service means that fewer ships can be deployed if the round-trip journey time is reduced. In particular, a typical way of reducing the round-trip journey time is to increase the sailing speed of ships as the port time is not easy to be shortened. However, as mentioned above, increasing the sailing speed means burning a large amount of fuel. Therefore, a liner shipping company must balance the trade-off between the number of ships and the bunker cost. This problem is further complicated because if the speed is low, the transit time of both laden and empty containers will be long. As a consequence, customers will incur a high inventory cost, and a larger container fleet has to be available for holding the cargo.

It should be noted that the total port charges associated with all ships on a ship route that has a weekly service frequency is independent of the round-trip journey time or equivalently the number of ships deployed. This is because the total port charges on all the ships in a week are equal to the total charges at all ports for a single ship in a round-trip. For instance, suppose that a ship visits Jakarta on Monday, visits

Singapore on Thursday, and returns to Jakarta next Monday. Suppose further that the port of Jakarta charges 5000 USD for each call, and the port of Singapore charges 10,000 USD for each call. Hence, each week the liner shipping company needs to pay 15,000 USD to ports. Now suppose that the company decides to slow down the speed of the ships and deploy two ships to maintain a weekly frequency: Each ship visits Jakarta on Monday, visits Singapore on the next Monday, and returns to Jakarta on the Monday 2 weeks later. In such a case, each ship either visits Jakarta or visits Singapore every Monday, and hence the total port charges for the liner shipping company in one week is still 15,000 USD. Similarly, to calculate the weekly bunker cost of all ships on a ship route, one only needs to calculate the total bunker cost for one ship to finish a round-trip journey.

12.2.2 Liner Shipping Network

A liner service cannot be designed or analyzed independently. This is because a liner shipping company cannot provide direct shipping services for each OD pair as the container shipment demand of most OD pairs is too small for a containership. In reality, a liner shipping company provides a shipping network which consists of a set of connected ship routes. For example, Fig. 12.3 shows an illustrative liner shipping network. A liner shipping network fulfills the container shipment demand of an OD pair by using one or more than one ship route. For example, although the network in Fig. 12.3 does not provide shipping services between Jakarta and Cochin, it can transport containers between the two ports by transshipping at the port of Singapore and the port of Colombo. As a consequence of transshipment, container shipment demand from any port covered by the network to any other port can be fulfilled. However, transshipment considerably increases the complexity of network design because there are numerous ways to transport containers from their origins to their destinations as a consequence of transshipment.

12.2.3 Slot Purchasing and Vessel Sharing

Due to the economies of scale, liner shipping companies cooperate in various forms (Agarwal and Ergun 2010; Brouer et al. 2014). For example, the company that operates the network in Fig. 12.3 serves the port of Colombo in South Asia. It can buy ship slots from local feeder services such as services connecting Colombo and Bangladeshi ports to transport containers for Bangladeshi ports. In terms of modeling, this is equivalent to adding the demands associated with Bangladeshi ports to Colombo. A liner shipping company may also charter slots on a trunk service (mainline service) from other shipping companies. Mathematically, this case is also easy to model. Companies in an alliance may have vessel sharing agreement (VSA) which means several companies contribute ships to a service and the number of slots each company controls is proportional to its contribution. VSA means that a

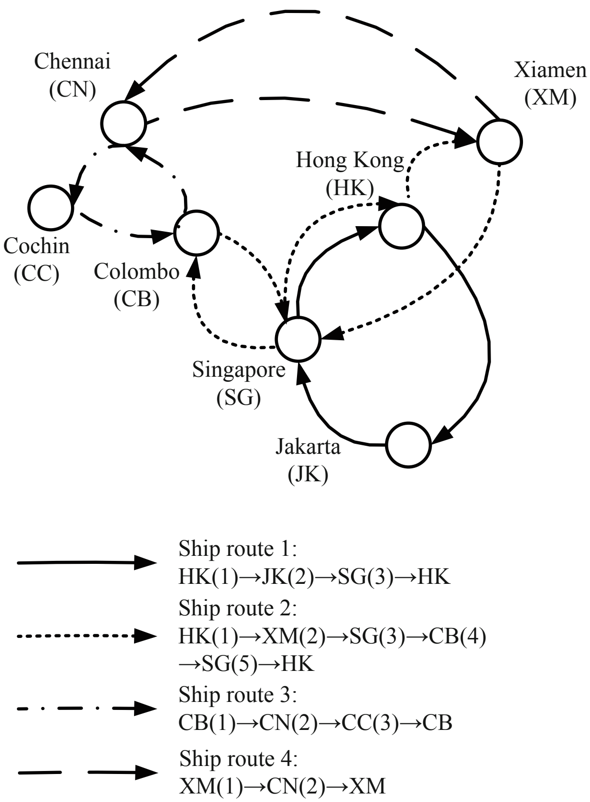


Fig. 12.3 An illustrative liner shipping network

liner shipping company has to discuss with other liner shipping companies before it could change their joint service. Few studies take into consideration VSA in network design.

12.3 Container Shipment Demand and Routing

12.3.1 Container Shipment Demand

To evaluate the efficiency of a liner shipping network, how many containers can be transported and the total operating cost for transporting the containers must be examined. Therefore, container shipment demand is an indispensable input. As liner shipping services are regular and fixed at least for a period of 3 to 6 months, the network is designed before knowing the exact container shipment demand or given known demand distribution. However, container shipment demand is an indispensable input for network design because the objective of network design is to minimize the total cost of fulfilling the demand or to maximize the profit by transporting only profitable containers. Consequently, the container shipment demand has to be estimated and the estimated demand is used as input for network design, although it is difficult to predict the demand.

In the literature on container liner shipping operations, most studies model weekly container shipment demand, that is, the demand is the same in each week. This modeling approach has two advantages: first, a liner shipping company only needs to predict one value for each OD pair; second, the resulting model is compact. Some other works consider a finite planning horizon and model the demand from one port to another in each day, for example, Brouer et al. (2011). This approach is more accurate whereas it requires a large number of inputs that a liner shipping company must provide. There are also studies that incorporate stochastic container shipment demand (Meng et al. 2012b) or elastic container shipment demand (Wang et al. 2013b). Nevertheless, we argue that given the difficulty of the network design problem, focusing on weekly demand is sufficient, at least in the current stage.

Note that in reality a liner shipping company cannot predict the demand for each OD pair. In fact, it predicts the increase of the demand from one region to another and allocates the increase proportionally to all OD pairs. For instance, suppose that in 2012, the company that ran the network in Fig. 12.3 had the demand of 500 TEUs/week from Xiamen to Cochin, 600 TEUs/week from Xiamen to Chennai, 700 TEUs/week from Hong Kong to Cochin, and no demand from Hong Kong to Cochin. It predicted that in 2013 the demand from China to India will increase by 20 %, then the predicted demand would be: 600 TEUs/week from Xiamen to Cochin, 720 TEUs/week from Xiamen to Chennai, and 840 TEUs/week from Hong Kong to Cochin.

12.3.2 Container Routing

A liner shipping company has to consider service factors and respect regulatory rules when transporting containers from origin to destination. The transit time of containers from origin port to destination port affects the inventory cost and depreciation of cargos and hence is a competitive factor for liner shipping companies (Notteboom 2006). For instance, Maersk Line (2013a) has been propagating that the "transportation time" of its Daily Maersk is shorter than other companies. However, most studies do not account for the transit time. In other words, they assume that the only requirement from shippers is that containers are transported from their origin ports to their destination ports.

Shippers may also have requirement on the number of transshipments. Transshipment brings in the risk of missing connection. Therefore, shippers prefer that their containers are delivered to the destination without transshipment. For instance, if a container is to be transported from Shanghai to Rotterdam on one ship route, the only risk associated with the transit time is that the ship arrival at Shanghai may be late. Nevertheless, even if the arrival at Shanghai is late, the liner shipping company may still manage to catch up the schedule by sailing faster or skipping ports. However, if the container is to be transshipped at Singapore, and the ship arrival at Shanghai is delayed, then it is highly possible that the ship from Shanghai to Singapore cannot catch up the schedule of the connecting ship that sails from Singapore to Rotterdam, and hence the container has to wait at Singapore for 1 week until the next ship that sails from Singapore to Rotterdam arrives. Consequently, transshipment adds the risk of missing connection and hence leads to higher uncertainty of the transit time. In research, although there are studies that require at most two transshipments for a container (e.g., Meng and Wang 2011; Song and Dong 2012), the main motivation is to simplify the model rather than to incorporate the level of service.

There may be various cabotage rules within countries restricting internal transport within the country by foreign ships/companies. For example, APL can transship the containers from Dalian, Tianjian, and Qingdao of China at the port of Pusan, South Korea, but it cannot transship the containers at Shanghai, as that involves internal transport within China. In other words, if APL could transport containers from Dalian to Shanghai, it would be able to compete with Chinese shipping companies in the domestic shipping market of China. Hence, the main purpose of cabotage rules is to protect the domestic shipping companies. Wang et al. (2013c) compared the routing of containers with and without cabotage and their case study demonstrated that cabotage considerably reduces the flexibility of container routing. Zheng et al. (2014) compared the designed network with and without cabotage. Other than these two studies, there are few quantitative works on maritime cabotage.

Another law that may affect container routing is embargoes: when a country A is embargoing on a country B, cargo from/to B cannot be transshipped in A (Brouer et al. 2014). For example, US embargos on several countries. We are aware of not study that incorporates embargo in network design, which is mainly because the affected container volume is insignificant.

12.4 Framework of Container Transport Network Design

Due to the complexity of container transport operations, mathematical models on container transport network design could only capture the most important factors. The designed network has to be revised by network planners before put to use. In this section we elaborate on the assumptions explicitly or implicitly required in most existing studies. Some studies have relaxed a few of these assumptions that will be pointed out as follows.

12.4.1 General Modeling Assumptions

There are a few frequently used assumptions in liner shipping network design modeling. First, ships in the fleet are categorized into types and ships in the same type are considered having the same capacity, cost structure and other properties. A string of ships in the same type is deployed on a ship route to maintain a weekly frequency. Weekly frequency of each ship route is the only requirement of level of service. The speed of ships is generally fixed and known, rather than a decision variable. An exception is Álvarez (2009), who treated the speed as a decision variable in network design by considering ships of different speeds as different types of ships. Using this modeling approach, the network design problem with ship speed optimization is equivalent to network design with fixed speed and with more types of ships. Here it should be noted that in reality two ships can hardly be identical and hence ships deployed on a ship route may also have some differences. Nevertheless, to simplify the modeling process and design tractable algorithms, it is advantageous to group ships into different types rather than treat each ship individually.

Second, port operations are usually highly simplified in network design. It is generally assumed that the time a ship spends at a port is a known parameter, although in reality this parameter is random and depends on many factors such as the container storage plan, the number of quay cranes deployed, the efficiency of quay crane operators, and the volume of containers handled. The port cost is also assumed to be fixed (depending on the port and the type of ship). The unit container handling cost (loading cost, discharge cost, transshipment cost) is known and there is no discount if a large volume of containers are handled at a particular port. It is usually assumed that any ship can visit any port at any time. Similarly, the canal dues are either not modeled or considered as a fixed parameter for each canal and each ship type. The additional time for transiting a canal is generally assumed to be 0.

Third, competitions from other shipping companies, purchasing slots and vessel sharing agreements are usually not considered in existing models.

12.4.2 Other Practical Considerations

Based on the above assumptions in modeling, the global container transport network design problem can be defined as follows: Given a set of ports and weekly container shipment demand between the ports, a fleet of ships, design a liner shipping network in which each ship route has a string of homogeneous ships to maintain a weekly frequency, to maximize the profit for fulfilling all or part of the container shipment demand.

Since the time associated with container storage at origin ports due to the weekly frequency, the transit time of containers and the connection time at transshipment ports are not incorporated, the designed ship routes has information on port rotations, type and number of ships to deploy, and a weekly frequency. The schedules of the ship routes are irrelevant.

We stress that the designed network can only be used as a benchmark for the existing network or a starting point for network planners to design an implementable network based on their experience. Besides the aforementioned assumptions, there are a number of reasons that the designed network cannot be put to use immediately. First, liner shipping companies cannot reshuffle their networks overnight. In this aspect, network alteration is more practical. Second, there are a number of business rules in each liner shipping company for designing ship routes. For example, Brouer et al. (2014) stated a few of Maersk Line’s rules: (i) Services with ships of capacity of at least 2400 TEUs must have a weekly frequency. (ii) The round-trip journey time of services with ships of capacity of at least 8400 TEUs is at least 4 weeks. Based on the first rule, the services have a weekly frequency. Hence, the number of ships on each of the services is at least four. (iii) Services with ships of capacity of at least 1600 TEUs must have at least two feeder ports of call between any two hub ports of call. Another issue that dramatically affects network design is the choice of hubs. For instance, some liner shipping companies operate container terminals at some ports and hence these ports will naturally serve as the hubs (the money paid for container handling is simply transferred from the left hand to the right hand of the company). In mathematical models, the pre-determined choice of hubs could be reflected by setting a lower container handling price at the hubs. Nevertheless, there should be more appropriate modeling techniques.

12.5 Special Network Design Models

In this section we look at models that are focused on network design under special settings. These models are generally aimed at designing shipping services in a local sense by fixing the other parts of the services in a global container transport network. These models shed insights into the global container transport network design problem and hence are reviewed here.

12.5.1 Ship Route Design Without Transshipment

Early models usually do not allow container transshipment. For example, Rana and Vickson (1988) built a mixed-integer nonlinear programming model for routing a single containership. They then linearized the model by enumerating all possible round trips and solved the resulting mixed-integer linear programming model using Benders’ decomposition. Rana and Vickson (1991) extended the work to multiple ship routes. They employed Lagrangian relaxation and a decomposition technique. Wang and Meng (2014) designed multiple ship routes with weekly frequency. The demand for each OD pair has a shipping deadline (maximum allowable transit time). Moreover, the time a ship spends at a port is a linear function of the number of containers handled. A column generation based heuristic method was developed to address the problem. In the above three studies, the port calling sequence is predetermined. Take Fig. 12.1 as an example, the predetermined port calling sequence requires that on the East Bound, if Yantian is to be visited, it must be visited after Singapore and before Shanghai. Such an assumption is reasonable when there is a clear geographical order of the ports. However, it may be unreasonable to pre-specify any sequence for European ports.

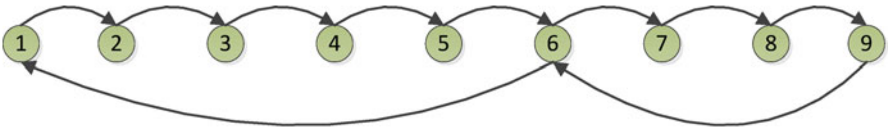


Fig. 12.4 A butterfly ship route

Shintani et al. (2007) relaxed the port calling precedence relation and considered empty container repositioning to design a single ship route. They assumed that all container shipment demand emanating from a port is satisfied if that port is visited, that is, the ship capacity is sufficient. The problem was formulated as a bi-level model, where the upper level is a knapsack problem choosing the best set of ports of call, and the lower level determines the optimal sequence to visit the ports chosen in the upper level. A genetic algorithm is employed to solve the upper-level and lower-level problems simultaneously.

Gelareh et al. (2013b) developed a heuristic approach for designing a single ship route. The ship route was designed in the form of a controlled re-sequencing, insertion and elimination of ports with regard to the current ship route. The outcome determines the required ship capacity, frequency, and port rotation.

Plum et al. (2014a) proposed a branch-and-cut-and-price to design a single liner shipping service. The numerical experiments demonstrate that the algorithm could solve instances with up to 25 ports to optimality.

The above ship route design models are applicable when a liner shipping company fixes all other ship routes in its network and allocates the residual demand that cannot be fulfilled by the fixed ship routes to the ship route(s) to be designed.

12.5.2 Ship Route Design with Transshipment Within the Ship Route

Reinhardt and Pisinger (2012) presented a model and a branch-and-cut method for designing butterfly ship routes to optimality. A butterfly ship route is a ship route with a port visited twice in a round trip and containers can be transshipped at the port, as shown in Fig. 12.4. We stress that the containers are transshipped from a ship on a ship route to another ship (could also be the same ship) on the same ship route. In Fig. 12.4, containers from port 5 to port 1 will be transshipped at port 6. Otherwise, the containers have to be transported for a much longer distance, which leads to the occupation of ship slots and a much longer transit time. The model of Reinhardt and Pisinger (2012) selects a route for each individual ship in a fleet. They commented that the configuration is suitable for smaller liner shipping companies as global liner shipping companies tend to group a set of ships with similar characteristics to a single service to reduce the complexity of the network design and to provide a regular schedule to customers.

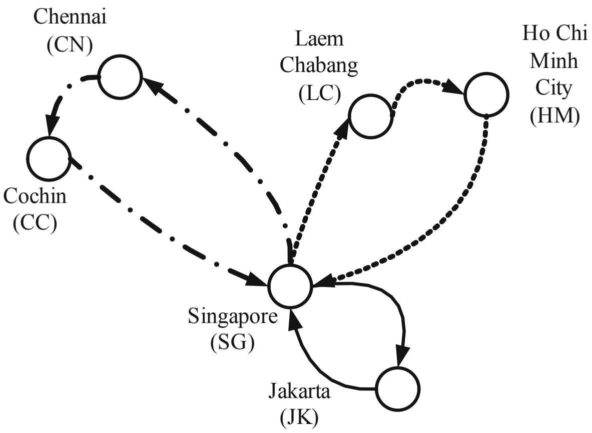


Fig. 12.5 A feeder network

Song and Dong (2013) investigated a liner long-haul ship route design problem. They assumed that the set of ports to be serviced is determined at the strategic level and proposed a three-stage optimization method. The first stage is the topological structure. For example, the butterfly ship route in Fig. 12.4 has two cycles, the ship route in Fig. 12.1 has three cycles. In view of the observation that most long-haul routes in practice have no greater than three directed cycles, the network design decisions are significantly simplified compared to enumerating all possible loops. The second stage is laden and empty flow optimization and the third stage is ship deployment. Although Song and Dong (2013) did not mention the term “butterfly ship route”, the designed ship routes are more general than those in Reinhardt and Pisinger (2012) in that the former may have more than two directed cycles. The ship routes of Reinhardt and Pisinger (2012) and Song and Dong (2013) are also more general than those in Shintani et al. (2007) as the former two allow container transshipment within a ship route.

12.5.3 Feeder Network Design

Some works have examined the feeder container shipping network design problem, which consists of a hub port and many feeder ports, see Fig. 12.5. Containers either originate from or are destined for the hub port, and transshipment is excluded within the feeder network. A feeder network covers a much smaller geographical region compared with a global shipping network. Consequently, a feeder network tends to have fewer ports and can be changed more frequently. In this sense, a feeder network can be designed from scratch.

Fagerholt (1999) contributed a pioneering study by proposing a set-partitioning model. The model requires enumerating all possible shipping service routes and combining these single shipping service routes into multiple shipping service routes if possible. Fagerholt (2004) later extended the set-partitioning model to address a heterogeneous ship fleet with a given cost structure, capacity and, in particular, sailing speed for each type of ship. He reported results for 40 ports and 20 ships. Sambracos et al. (2004) carried out a case study on the feeder ship route design for dispatching small containers in the Aegean Sea, from one depot port (Piraeus) to 12 other ports (islands). A list-based threshold acceptance meta-heuristic method was employed and results show that at least a 5.1 % cost saving could be realized over existing shipping practices. Karlaftis et al. (2009) generalized the above problem to account for container pick-up and delivery operations as well as time deadlines. A hybrid genetic algorithm was applied to solve the problem.

12.5.4 Hub-and-Spoke Network Design

In view of the importance of container transshipment in liner shipping operations, some efforts are devoted to design a hub-and-spoke liner shipping network, as shown in Fig. 12.6. A hub-and-spoke network generally requires that containers can only be transshipped at the pre-specified hubs and container shipment demand between two feeder ports belonging to two different hubs must be transshipped at the two hubs. Nevertheless, it should be stressed that the hub-and-spoke structure in liner shipping is not as evident as in airline. In fact, as Notteboom (2004) pointed out, a global network will not have a pure hub-and-spoke structure or a pure multi-port-call structure. This implies that in liner shipping network design, one should seek to design a network that is complex enough to have both the properties of hub-and-spoke structure and multi-port-call structure so that the network can have practical implications.

Still, there are some studies that have examined pure hub-and-spoke and multi-port-call structures, which shed lights to liner shipping network design. Imai et al. (2009) compared the efficiency of hub-and-spoke networks and multi-port-call networks using a six-port example, with three ports in one region (e.g. North America), and the other three in the other region (e.g., Europe), as shown in Fig. 12.5. In the hub-and-spoke network, one port in each region is chosen as a hub. In the multi-port-call network, all six ports are visited sequentially. Results show that the multi-port-call network is superior to the hub-and-spoke network in most scenarios of European or North American trade lanes. Gelareh et al. (2010) examined a hub-and-spoke network design problem in a competitive environment with a newcomer liner shipping company and an existing dominating operator. The newcomer chooses hub ports so as to maximize its market share, which depends on the cost and transit time. A mixed-integer programming formulation and a Lagrangian method combined with a primal heuristic were developed. Based on this work, Gelareh and Pisinger (2011) presented a mixed-integer linear programming formulation for the simultaneous design of a hub-and-spoke network and deployment of containerships. A Benders' decomposition-based algorithm was developed that outperformed general-purpose mixed-integer linear programming solvers. Gelareh et al. (2013a) proposed a mixed-integer linear programming model to minimize the weighted sum of transit time and

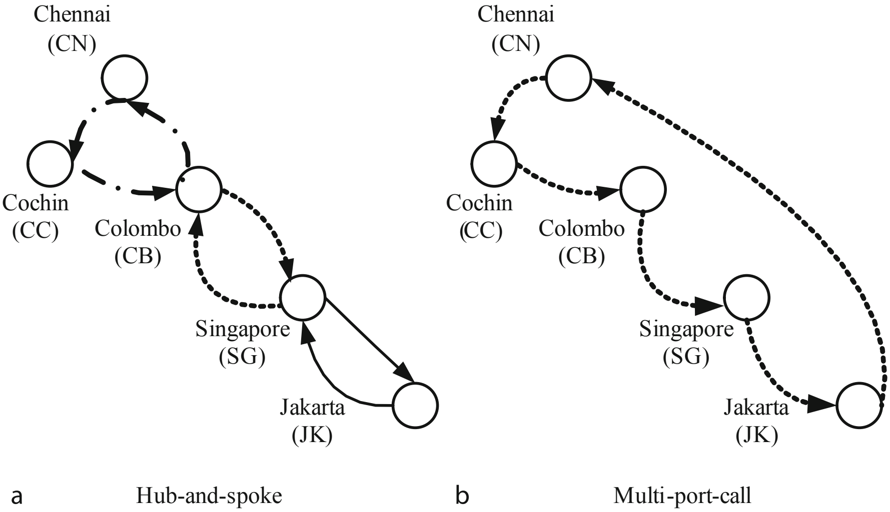


Fig. 12.6 Hub-and-spoke versus multi-port-call structures

fixed deployment costs for a hub-and-spoke liner shipping network. A Lagrangian decomposition approach which uses a heuristic procedure was applied to obtain practical solutions in reasonable time.

12.6 General Network Design Models

We use the term "general network design models" to refer to models for designing multiple ship routes and in which containers can be transshipped at any ports that are visited more than once a week. This is practical: every port can transship containers since transshipment is simply a discharge operation plus a loading operation. If transshipping at a port that has no transshipment before is more advantageous, the liner shipping company should consider transshipping containers at the port.

Agarwal and Ergun (2008) proved that the general network design problem is weakly NP-hard as it can be reduced to a knapsack problem. Brouer et al. (2014) strengthened the result by proving that the general network design problem is strongly NP-hard as it can be reduced to a traveling salesman problem. Brouer et al. (2014) further proved that the general network design problem with a set of candidate port rotations is strongly NP-hard as it can be reduced to a set covering problem. In fact, solving a large-scale general network design problem is difficult even heuristically.

In general network design, there are two fundamental problems: (i) How to generate ship routes; (ii) Given a set of ship routes to operate, how to evaluate its efficiency. The efficiency of a liner shipping network is the maximum profit that can be gained from transporting containers. Therefore, the second problem actually decides how to choose the most profitable containers to transport and how to transport the containers.

The second problem is a special case of the following problem: Given a set of candidate ship routes, determine the optimal subset of ship routes to operate and determine the optimal container routing with the chosen ship routes. We first formulate three types of models for the ship route selection and container routing problem, and then describe how existing studies design a general container transportation network.

12.6.1 Ship Route Selection and Container Routing

Consider a set $\mathcal{R}$ of candidate ship routes, regularly serving a group of ports denoted by the set $\mathcal{P}$. Ship route $r \in \mathcal{R}$ can be expressed as:

$$p_{r1} \to p_{r2} \to \cdots \to p_{rN_r} \to p_{r1} \tag{12.2}$$

where N_r is the number of ports of call and p_{ri} is the ith port of call, $i = 1,2,\cdots,N_r$. Define $I_r = \{1,2,\cdots,N_r\}$. The voyage from port i to port $i + 1$ is called leg i and leg N_r is the voyage from port N_r to port 1. Figure 12.3 shows four ship routes: ship route 1 has three legs, ship route 2 has five legs, ship route 3 has three legs, and ship route 4 has two legs. The container capacity of a ship deployed on ship route $r \in \mathcal{R}$ is denoted by V_r (TEUs). The operating cost of ship route r is c_r (USD/week), including capital cost, bunker cost, port and canal dues, manning cost and consumables.

Represent by W the set of OD pairs, $W \subset P \times P$. The demand for OD pair $(o, d) \in W$ is denoted by q^{od} (TEUs/week). The revenue for shipping a container is g^{od} (USD/TEU). Containers can be transshipped at any port from origins to destinations. The load, transshipment, discharge cost (USD/TEU) at port $p \in P$ is denoted by $\hat{c}_p$, $\bar{c}_p$ and $\tilde{c}_p$, respectively. We have $\bar{c}_p < \hat{c}_p + \tilde{c}_p$ because ports encourage liner shipping companies to transship and transshipment involves much less paper work than exporting and importing.

Path-Based Formulation

A straightforward way of formulating container flow is to use paths. We refer to a path as a container route. For example, the followings are three container routes with respect to the ship routes shown in Fig. 12.3:

$$h_1 = p_{1,3}(\text{SG}) \xrightarrow{\text{Ship Route 1}} p_{1,1}(\text{HK}) \tag{12.3}$$

$$h_2 = p_{2,5}(\text{SG}) \xrightarrow{\text{Ship Route 2}} p_{2,1}(\text{HK}) \tag{12.4}$$

$$h_3 = p_{2,2}(\text{XM}) \xrightarrow{\text{Ship Route 2}} p_{2,4}(\text{CB}) \mapsto p_{3,1}(\text{CB}) \xrightarrow{\text{Ship Route 3}} p_{3,2}(\text{CN}) \tag{12.5}$$

Container route h_1 is used to directly deliver containers from Singapore to Hong Kong which are loaded at the 3rd port of call of the ship route 1 (Singapore) and discharged at the 1st port of call of the ship route 1 (Hong Kong). Containers along the container route h_2 are delivered by the ship route 2. Container route h_3 involves container transshipment operations: Containers are first loaded at the 2nd port of call of the ship route 2 (Xiamen) and delivered to the 4th port of call of the ship route 2 (Colombo). At Colombo, these containers are discharged and reloaded (transshipped) to a ship deployed on the ship route 3, and transported to the destination Chennai.

The set of container routes for $(o,d) \in W$ is denoted by H^{od}. Define $H := \bigcup_{(o,d)\in W} H^{od}$ to be the set of all container routes for all the OD pairs. The container handling cost of $h \in H^{od}$ is c_h (USD/TEU). For instance, in Eq. (12.5), $c_{h_3} = \hat{c}_{\text{XM}} + \bar{c}_{\text{CB}} + \tilde{c}_{\text{CN}}$. We further let binary coefficient ρ_h^{ri} be 1 if containers on container route h are transported on leg i of ship route r, and 0 otherwise. For example, container route h_3 consists of the 2nd and the 3rd legs of the ship route 2 and the 1st leg of the ship route 3. We hence have $\rho_{h_3}^{2,2} = 1$, $\rho_{h_3}^{2,3} = 1$, $\rho_{h_3}^{3,1} = 1$.

The decision variables are as follows. x_r is a binary variable which equals 1 if and only if candidate ship route r is operated and 0 otherwise; y_h is the volume (TEUs/week) of containers transported on container route h. The path-based formulation is a mixed-integer linear programming model:

$$\max_{x_r, y_h} \sum_{(o,d)\in\mathcal{W}} g^{od} \sum_{h\in H^{od}} y_h - \sum_{r\in\mathcal{R}} c_r x_r - \sum_{h\in H} c_h y_h \tag{12.6}$$

subject to:

$$\sum_{h\in H} \rho_h^{ri} y_h \le V_r x_r, \forall r \in R, \forall i \in I_r \tag{12.7}$$

$$\sum_{h\in H^{od}} y_h \le q^{od}, \forall (o,d) \in \mathcal{W} \tag{12.8}$$

$$x_r \in \{0,1\}, \forall r \in R \tag{12.9}$$

$$y_h \ge 0, \forall h \in H \tag{12.10}$$

The objective function (12.6) maximizes the total profit, which is the revenue for shipping containers minus the operating cost of ship routes and container handling cost. Eq. (12.7) imposes the ship capacity constraint. Constraint (12.8) enforces that the volume of transported containers cannot exceed the demand. Constraint (12.9) defines x_r as a binary variable and constraint (12.10) defines y_h as a nonnegative continuous variable.

The path-based model [P1] only captures the most essential liner shipping features for better readability. For example, it assumes an unlimited number of ships available in the fleet and does not incorporate empty container repositioning. y_h is a continuous number rather than an integer in constraint (12.10) because the magnitude of y_h is usually several tens and the error caused by rounding down to the nearest integer is much smaller than the estimation error of the demand (Brouer et al. 2014; Wang 2013).

The path-based model is adopted in Meng et al. (2012a) and Wang et al. (2013a). Moreover, Meng and Wang (2011) employed a segment-based model which is very similar to path-based model: a path consists of one to three segments. The path-based model has a very small number of constraints. However, the cardinality of container routes $|H|$ is exponentially large, which implies the large number of binary decision variables x_r. Another disadvantage of the path-based model is that the set of paths H depends on the set of candidate ship routes $\mathcal{R}$. However, in network design the set of candidate ship routes $\mathcal{R}$ is updated in each iteration. Consequently, the set of paths H has to be updated in each iteration too, which leads to intensive computation time. However, the path-based model has the advantage that side constraints such as maximum allowable transit time and maritime cabotage can be easily incorporated (Wang et al., 2013c). Therefore, the path-based model is commonly used in container routing problems (Brouer et al. 2011; Song and Dong 2012).

OD-Link-Based Formulation

A compact model that does not need path enumeration or generation is OD-link-based. Since the OD-link-based formulation is not as intuitive as the path-based formulation, we first present the OD-link-based formulation for a single OD denoted by (o, d), that is, $W = \{(o, d)\}$.

The decision variables are as follows. x_r is a binary variable which equals 1 if and only if ship route r is operated and 0 otherwise; $\hat{z}_{ri}$ and $\tilde{z}_{ri}$ are the volume of containers (TEUs/week) loaded and discharged at port of call i on ship route r, respectively (note that when calculating $\hat{z}_{ri}$ and $\tilde{z}_{ri}$, a transshipped container is considered as being discharged once and being loaded once); f_{ri} is the volume of containers (TEUs/week) flowing on leg i on ship route r (we define $f_{r0} := f_{rN_r}$). y is the fulfilled demand (TEUs/week); $\hat{z}_p$, $\tilde{z}_p$, and $\bar{z}_p$ are the total volume of loaded, discharged, and transshipped containers (TEUs/week) at port $p \in P$, respectively. The OD-link-based formulation for a single OD is a mixed-integer linear programming problem:

$$\max_{x_r, \hat{z}_{ri}, \tilde{z}_{ri}, f_{ri}, y, \hat{z}_p, \bar{z}_p, \tilde{z}_p} g^{od} y - \sum_{r \in \mathcal{R}} c_r x_r - \sum_{p \in \mathcal{P}} (\hat{z}_p \hat{c}_p + \bar{z}_p \bar{c}_p + \tilde{z}_p \tilde{c}_p) \tag{12.11}$$

subject to:

$$f_{r,i-1} + \hat{z}_{ri} = f_{ri} + \tilde{z}_{ri}, \forall r \in R, \forall i \in I_r \tag{12.12}$$

$$\hat{z}_p = \begin{cases} y, p = o \\ 0, \forall p \in \mathcal{P} \backslash \{\mathrm{o}\} \end{cases} \tag{12.13}$$

$$\tilde{z}_p = \begin{cases} y, p = d \\ 0, \forall p \in \mathcal{P} \backslash \{\mathrm{d}\} \end{cases} \tag{12.14}$$

$$\bar{z}_p = \sum_{r \in \mathcal{R}} \sum_{i \in I_r, p_{ri} = p} \hat{z}_{ri} - \hat{z}_p, \forall p \in \mathcal{P} \tag{12.15}$$

$$\sum_{r\in\mathcal{R}}\sum_{i\in I_r,p_{ri}=p}(\hat{z}_{ri}-\tilde{z}_{ri})=\begin{cases}y, p=o\\ -y, p=d\\ 0, \text{otherwise}\end{cases},\forall p\in\mathcal{P} \tag{12.16}$$

$$f_{ri}\leq V_r x_r,\forall r\in R,\forall i\in I_r \tag{12.17}$$

$$y\leq q^{od} \tag{12.18}$$

$$x_r\in\{0,1\},\forall r\in R \tag{12.19}$$

$$\hat{z}_{ri}\geq 0,\tilde{z}_{ri}\geq 0,f_{ri}\geq 0,\forall r\in R,\forall i\in I_r \tag{12.20}$$

$$y\geq 0 \tag{12.21}$$

The objective function (12.11) maximizes the total profit. Eq. (12.12) is the flow conservation constraint. Constraint (12.13) defines the total loaded containers at a port. Constraint (12.14) defines the total discharged containers at a port. Constraint (12.15) defines the total transshipped containers at a port. Constraint (12.16) computes the fulfilled demand. Constraint (12.17) enforces the ship capacity constraint. Constraint (12.18) defines the upper limit of the fulfilled demand. Constraint (12.19) defines x_r as a binary variable. Constraint (12.20) defines $\hat{z}_{ri}$, $\tilde{z}_{ri}$ and f_{ri} as nonnegative continuous variables. Constraint (12.21) defines y as a nonnegative continuous variable. Note that the nonnegativity of $\hat{z}_p$, $\tilde{z}_p$ and $\bar{z}_p$ is implicitly incorporated.

Having formulated the OD-link-based model for a single OD, we are now ready to formulate the OD-link-based model for a set of OD pairs W. The interactions between different OD pairs lie in that they need to share the same ship slots. The decision variables are as follows. x_r is a binary variable which equals 1 if and only if ship route r is operated and 0 otherwise; $\hat{z}^{od}_{ri}$ and $\tilde{z}^{od}_{ri}$ are the volume of containers (TEUs/week) from $(o,d)\in W$ loaded and discharged at port of call i on ship route r, respectively (note that when calculating $\hat{z}^{od}_{ri}$ and $\tilde{z}^{od}_{ri}$, a transshipped container is considered as being discharged once and being loaded once); f^{od}_{ri} is the volume of containers (TEUs/week) from (o,d) flowing on leg i on ship route r (we define $f^{od}_{r0}:=f^{od}_{rN_r}$). y^{od} is the fulfilled demand (TEUs/week) for (o,d); $\hat{z}_p$, $\tilde{z}_p$, and $\bar{z}_p$ are the total volume of loaded, discharged, and transshipped containers (TEUs/week) at port $p\in P$, respectively. The OD-link-based formulation for multiple OD pairs is a mixed-integer linear programming problem:

$$\max_{x_r,\hat{z}^{od}_{ri},\tilde{z}^{od}_{ri},f^{od}_{ri},y^{od},\hat{z}_p,\bar{z}_p,\tilde{z}_p}\sum_{(o,d)\in\mathcal{W}}g^{od}y^{od}-\sum_{r\in\mathcal{R}}c_r x_r-\sum_{p\in\mathcal{P}}(\hat{z}_p\hat{c}_p+\bar{z}_p\bar{c}_p+\tilde{z}_p\tilde{c}_p) \tag{12.22}$$

subject to:

$$f^{od}_{r,i-1}+\hat{z}^{od}_{ri}=f^{od}_{ri}+\tilde{z}^{od}_{ri},\forall r\in R,\forall i\in I_r,\forall(o,d)\in\mathcal{W} \tag{12.23}$$

$$\hat{z}_p = \sum_{(p,d)\in\mathcal{W}} y^{pd}, \forall p \in \mathcal{P} \tag{12.24}$$

$$\tilde{z}_p = \sum_{(o,p)\in\mathcal{W}} y^{op}, \forall p \in \mathcal{P} \tag{12.25}$$

$$\bar{z}_p = \sum_{r\in\mathcal{R}} \sum_{i\in I_r, p_{ri}=p} \sum_{(o,d)\in\mathcal{W}} \hat{z}_{ri}^{od} - \hat{z}_p, \forall p \in \mathcal{P} \tag{12.26}$$

$$\sum_{r\in\mathcal{R}} \sum_{i\in I_r, p_{ri}=p} (\hat{z}_{ri}^{od} - \tilde{z}_{ri}^{od}) = \begin{cases} y^{od}, p = o \\ -y^{od}, p = d \\ 0, \text{otherwise} \end{cases} , \forall (o,d) \in \mathcal{W}, \forall p \in \mathcal{P} \tag{12.27}$$

$$\sum_{(o,d)\in\mathcal{W}} f_{ri}^{od} \le V_r x_r, \forall r \in R, \forall i \in I_r \tag{12.28}$$

$$y^{od} \le q^{od}, \forall (o,d) \in \mathcal{W} \tag{12.29}$$

$$x_r \in \{0,1\}, \forall r \in R \tag{12.30}$$

$$\hat{z}_{ri}^{od} \ge 0, \tilde{z}_{ri}^{od} \ge 0, f_{ri}^{od} \ge 0, \forall r \in R, \forall i \in I_r, \forall (o,d) \in \mathcal{W} \tag{12.31}$$

$$y^{od} \ge 0, \forall (o,d) \in \mathcal{W} \tag{12.32}$$

The objective function (12.22) maximizes the total profit. Eq. (12.23) is the flow conservation constraint. Constraint (12.24) defines the total loaded containers at a port. Constraint (12.25) defines the total discharged containers at a port. Constraint (12.26) defines the total transshipped containers at a port. Constraint (12.27) computes the fulfilled demand. Constraint (12.28) enforces the ship capacity constraint. Constraint (12.29) defines the upper limit of the fulfilled demand. Constraint (12.30) defines x_r as a binary variable. Constraint (12.31) defines $\hat{z}_{ri}^{od}$, $\tilde{z}_{ri}^{od}$ and f_{ri}^{od} as nonnegative continuous variables. Constraint (12.32) defines y^{od} as a nonnegative continuous variable. Note that the nonnegativity of $\hat{z}_p$, $\tilde{z}_p$ and $\bar{z}_p$ is implicitly incorporated.

The OD-link-based model is adopted in Agarwal and Ergun (2008): they used a single index to represent the "commodity" rather than two indices o and d. Compared with the path-based formulation, the number of decision variables and constraints in the OD-link-based model are both polynomially bounded by the size of the liner shipping network.

Origin-Link-Based Formulation

The number of flow variables f_{ri}^{od} in the OD-link-based model has the magnitude of $|\mathcal{W}|\sum_{\mathrm{r}\in\mathcal{R}}\mathcal{N}_{\mathrm{r}}$. If we use an origin-link-based model, we will have a total of $|\mathcal{P}|\sum_{\mathrm{r}\in\mathcal{R}}\mathcal{N}_{\mathrm{r}}$ flow variables, which is one order of magnitude smaller than the OD-link-based model.

Similar to the OD-link-based model, we first give the origin-link-based model for containers with the same origin port. In other words, the set of OD pairs W has many elements, but all elements have the same origin port denoted by $o \in \mathcal{P}$. The decision variables are as follows. x_r is a binary variable which equals 1 if and only if candidate ship route r is operated and 0 otherwise; $\hat{z}_{ri}$ and $\tilde{z}_{ri}$ are the volume of containers (TEUs/week) loaded and discharged at port of call i on ship route r, respectively (note that when calculating $\hat{z}_{ri}$ and $\tilde{z}_{ri}$, a transshipped container is considered as being discharged once and being loaded once); f_{ri} is the volume of containers (TEUs/week) flowing on leg i of ship route r (we define $f_{r0} := f_{rN_r}$). We define $\mathcal{W} := \mathcal{P} \times \mathcal{P}$ and $q^{od} = 0$ if there is no demand from port o to port d. y^{od} is the fulfilled demand (TEUs/week) for $(o,d) \in W$; $\hat{z}_p$, $\tilde{z}_p$, and $\bar{z}_p$ are the total volume of loaded, discharged, and transshipped containers (TEUs/week) at port $p \in P$, respectively. The origin-link-based formulation for OD pairs with the same origin o is a mixed-integer linear programming problem:

$$\max_{x_r,\hat{z}_{ri},\tilde{z}_{ri},f_{ri},y^{od},\hat{z}_p,\bar{z}_p,\tilde{z}_p} \sum_{(o,d)\in\mathcal{W}} g^{od}y^{od} - \sum_{r\in\mathcal{R}} c_r x_r - \sum_{p\in\mathcal{P}} (\hat{z}_p\hat{c}_p + \bar{z}_p\bar{c}_p + \tilde{z}_p\tilde{c}_p) \tag{12.33}$$

subject to:

$$f_{r,i-1} + \hat{z}_{ri} = f_{ri} + \tilde{z}_{ri}, \forall r \in R, \forall i \in I_r \tag{12.34}$$

$$\hat{z}_p = \begin{cases} \sum\limits_{(o,d)\in\mathcal{W}} y^{od}, p = o \\ 0, \forall p \in \mathcal{P}\backslash\{\mathrm{o}\} \end{cases} \tag{12.35}$$

$$\tilde{z}_p = y^{op}, \forall p \in \mathcal{P} \tag{12.36}$$

$$\bar{z}_p = \sum_{r\in\mathcal{R}} \sum_{i\in I_r, p_{ri}=p} \hat{z}_{ri} - \hat{z}_p, \forall p \in \mathcal{P} \tag{12.37}$$

$$\sum_{r\in\mathcal{R}} \sum_{i\in I_r, p_{ri}=p} (\hat{z}_{ri} - \tilde{z}_{ri}) = \begin{cases} \sum\limits_{(o,d)\in\mathcal{W}} y^{od}, p = o \\ -y^{op}, p \neq o \end{cases}, \forall p \in \mathcal{P} \tag{12.38}$$

$$f_{ri} \leq V_r x_r, \forall r \in R, \forall i \in I_r \tag{12.39}$$

$$y^{od} \leq q^{od}, \forall (o,d) \in \mathcal{W} \tag{12.40}$$

$$x_r \in \{0,1\}, \forall r \in R \tag{12.41}$$

$$\hat{z}_{ri} \geq 0, \tilde{z}_{ri} \geq 0, f_{ri} \geq 0, \forall r \in R, \forall i \in I_r \tag{12.42}$$

$$y^{od} \geq 0, \forall (o,d) \in \mathcal{W} \tag{12.43}$$

The objective function (12.33) maximizes the total profit. Eq. (12.34) is the flow conservation constraint. Constraint (12.35) defines the total loaded containers at a port. Constraint (12.36) defines the total discharged containers at a port. Constraint (12.37) defines the total transshipped containers at a port. Constraint (12.38) computes the fulfilled demand. Constraint (12.39) enforces the ship capacity constraint. Constraint (12.40) defines the upper limit of the fulfilled demand. Constraint (12.41) defines x_r as a binary variable. Constraint (12.42) defines $\hat{z}_{ri}$, $\tilde{z}_{ri}$ and f_{ri} as nonnegative continuous variables. Constraint (12.43) defines y^{od} as a nonnegative continuous variable. Note that the nonnegativity of $\hat{z}_p$, $\tilde{z}_p$, and $\bar{z}_p$ is implicitly incorporated.

Now we can formulate the origin-link-based model for OD pairs W with many origins. The decision variables are as follows. x_r is a binary variable which equals 1 if and only if candidate ship route r is operated and 0 otherwise; $\hat{z}^o_{ri}$ and $\tilde{z}^o_{ri}$ are the volume of containers (TEUs/week) with origin port $o \in \mathcal{P}$ and any destination port loaded and discharged at port of call i on ship route r, respectively (note that when calculating $\hat{z}^o_{ri}$ and $\tilde{z}^o_{ri}$, a transshipped container is considered as being discharged once and being loaded once); f^o_{ri} is the volume of containers (TEUs/week) with origin port $o \in \mathcal{P}$ and any destination port flowing on leg i on ship route r (we define $f^o_{r0} := f^o_{rN_r}$). We define $\mathcal{W} := \mathcal{P} \times \mathcal{P}$ and $q^{od} = 0$ if there is no demand from port o to port d. y^{od} is the fulfilled demand (TEUs/week) for $(o,d) \in W$; $\hat{z}_p$, $\tilde{z}_p$, and $\bar{z}_p$ are the total volume of loaded, discharged, and transshipped containers (TEUs/week) at port $p \in P$, respectively. The origin-link-based formulation for OD pairs with many origins is a mixed-integer linear programming problem:

$$\max_{x_r, \hat{z}^o_{ri}, \tilde{z}^o_{ri}, f^o_{ri}, y^{od}, \hat{z}_p, \bar{z}_p, \tilde{z}_p} \sum_{(o,d)\in\mathcal{W}} g^{od} y^{od} - \sum_{r\in\mathcal{R}} c_r x_r - \sum_{p\in\mathcal{P}} (\hat{z}_p \hat{c}_p + \bar{z}_p \bar{c}_p + \tilde{z}_p \tilde{c}_p) \quad (12.44)$$

subject to:

$$f^o_{r,i-1} + \hat{z}^o_{ri} = f^o_{ri} + \tilde{z}^o_{ri}, \forall r \in R, \forall i \in I_r, \forall o \in \mathcal{P} \quad (12.45)$$

$$\hat{z}_p = \sum_{(p,d)\in\mathcal{W}} y^{pd}, \forall p \in \mathcal{P} \quad (12.46)$$

$$\tilde{z}_p = \sum_{(o,p)\in\mathcal{W}} y^{op}, \forall p \in \mathcal{P} \quad (12.47)$$

$$\bar{z}_p = \sum_{r\in\mathcal{R}} \sum_{i\in I_r, p_{ri}=p} \sum_{o\in\mathcal{P}} \hat{z}^o_{ri} - \hat{z}_p, \forall p \in \mathcal{P} \quad (12.48)$$

$$\sum_{r\in\mathcal{R}} \sum_{i\in I_r, p_{ri}=p} (\hat{z}^o_{ri} - \tilde{z}^o_{ri}) = \begin{cases} \sum_{(o,d)\in\mathcal{W}} y^{od}, p = o \\ -y^{op}, p \neq o \end{cases}, \forall o \in \mathcal{P}, \forall p \in \mathcal{P} \quad (12.49)$$

$$\sum_{o\in\mathcal{P}} f^o_{ri} \leq V_r x_r, \forall r \in R, \forall i \in I_r \quad (12.50)$$

$$y^{od} \le q^{od}, \forall (o,d) \in \mathcal{W} \tag{12.51}$$

$$x_r \in \{0,1\}, \forall r \in R \tag{12.52}$$

$$\hat{z}^o_{ri} \ge 0, \tilde{z}^o_{ri} \ge 0, f^o_{ri} \ge 0, \forall r \in R, \forall i \in I_r, \forall o \in \mathcal{P} \tag{12.53}$$

$$y^{od} \ge 0, \forall (o,d) \in \mathcal{W} \tag{12.54}$$

The objective function (12.44) maximizes the total profit. Eq. (12.45) is the flow conservation constraint. Constraint (12.46) defines the total loaded containers at a port. Constraint (12.47) defines the total discharged containers at a port. Constraint (12.48) defines the total transshipped containers at a port. Constraint (12.49) computes the fulfilled demand. Constraint (12.50) enforces the ship capacity constraint. Constraint (12.51) defines the upper limit of the fulfilled demand. Constraint (12.52) defines x_r as a binary variable. Constraint (12.53) defines $\hat{z}^o_{ri}$, $\tilde{z}^o_{ri}$ and f^o_{ri} as nonnegative continuous variables. Constraint (12.54) defines y^{od} as a nonnegative continuous variable. Note that the nonnegativity of $\hat{z}_p$, $\tilde{z}_p$, and $\bar{z}_p$ is implicitly incorporated.

The origin-link-based model is used in Wang and Meng (2012a) for a fleet deployment problem.

Destination-Link-Based Formulation

Similar to the origin-link-based model, Álvarez (2009) and Brouer et al. (2014) have applied a destination-link-based formulation in network design. We directly formulate the destination-link-based model for OD pairs W with many destinations. The decision variables are as follows. x_r is a binary variable which equals 1 if and only if candidate ship route r is operated and 0 otherwise; $\hat{z}^d_{ri}$ and $\tilde{z}^d_{ri}$ are the volume of containers (TEUs/week) with destination port $d \in \mathcal{P}$ and any origin port loaded and discharged at port of call i on ship route r, respectively (note that when calculating $\hat{z}^d_{ri}$ and $\tilde{z}^d_{ri}$, a transshipped container is considered as being discharged once and being loaded once); f^d_{ri} is the volume of containers (TEUs/week) with destination port $d \in \mathcal{P}$ and any origin port flowing on leg i on ship route r (we define $f^d_{r0} := f^d_{rN_r}$). We define $\mathcal{W} := \mathcal{P} \times \mathcal{P}$ and $q^{od} = 0$ if there is no demand from port o to port d. y^{od} is the fulfilled demand (TEUs/week) for $(o,d) \in W$; $\hat{z}_p$, $\tilde{z}_p$, and $\bar{z}_p$ are the total volume of loaded, discharged, and transshipped containers (TEUs/week) at port $p \in P$, respectively. The destination-link-based formulation for OD pairs with many destinations is a mixed-integer linear programming problem:

$$\max_{x_r, \hat{z}^d_{ri}, \tilde{z}^d_{ri}, f^d_{ri}, y^{od}, \hat{z}_p, \bar{z}_p, \tilde{z}_p} \sum_{(o,d) \in \mathcal{W}} g^{od} y^{od} - \sum_{r \in \mathcal{R}} c_r x_r - \sum_{p \in \mathcal{P}} (\hat{z}_p \hat{c}_p + \bar{z}_p \bar{c}_p + \tilde{z}_p \tilde{c}_p) \tag{12.55}$$

subject to:

$$f^d_{r,i-1} + \hat{z}^d_{ri} = f^d_{ri} + \tilde{z}^d_{ri}, \forall r \in R, \forall i \in I_r, \forall d \in \mathcal{P} \tag{12.56}$$

$$\hat{z}_p = \sum_{(p,d) \in \mathcal{W}} y^{pd}, \forall p \in \mathcal{P} \tag{12.57}$$

$$\tilde{z}_p = \sum_{(o,p)\in\mathcal{W}} y^{op}, \forall p \in \mathcal{P} \tag{12.58}$$

$$\bar{z}_p = \sum_{r\in\mathcal{R}} \sum_{i\in I_r, p_{ri}=p} \sum_{d\in\mathcal{P}} \hat{z}^d_{ri} - \hat{z}_p, \forall p \in \mathcal{P} \tag{12.59}$$

$$\sum_{r\in\mathcal{R}} \sum_{i\in I_r, p_{ri}=p} (\hat{z}^d_{ri} - \tilde{z}^d_{ri}) = \begin{cases} -\sum_{(o,d)\in\mathcal{W}} y^{od}, p = d \\ y^{op}, p \neq d \end{cases}, \forall d \in \mathcal{P}, \forall p \in \mathcal{P} \tag{12.60}$$

$$\sum_{d\in\mathcal{P}} f^d_{ri} \le V_r x_r, \forall r \in R, \forall i \in I_r \tag{12.61}$$

$$y^{od} \le q^{od}, \forall (o,d) \in \mathcal{W} \tag{12.62}$$

$$x_r \in \{0,1\}, \forall r \in R \tag{12.63}$$

$$\hat{z}^d_{ri} \ge 0, \tilde{z}^d_{ri} \ge 0, f^d_{ri} \ge 0, \forall r \in R, \forall i \in I_r, \forall d \in \mathcal{P} \tag{12.64}$$

$$y^{od} \ge 0, \forall (o,d) \in \mathcal{W} \tag{12.65}$$

The objective function (12.55) maximizes the total profit. Eq. (12.56) is the flow conservation constraint. Constraint (12.57) defines the total loaded containers at a port. Constraint (12.58) defines the total discharged containers at a port. Constraint (12.59) defines the total transshipped containers at a port. Constraint (12.60) computes the fulfilled demand. Constraint (12.61) enforces the ship capacity constraint. Constraint (12.62) defines the upper limit of the fulfilled demand. Constraint (12.63) defines x_r as a binary variable. Constraint (12.64) defines $\hat{z}^o_{ri}$, $\tilde{z}^o_{ri}$ and f^o_{ri} as nonnegative continuous variables. Constraint (12.65) defines y^{od} as a nonnegative continuous variable. Note that the nonnegativity of $\hat{z}_p$, $\tilde{z}_p$, and $\bar{z}_p$ is implicitly incorporated.

Hybrid-Link-Based Formulation

Wang (2014) proposed a compact hybrid-link-based model based on the following observation. As shown in Fig. 12.7, the number of origins $|O| = 5$ and the number of destinations $|D| = 6$, where O and D represent the sets of origins and destinations, respectively. However, if we use origin-link-based formulation for origin port p_4 (and the OD pairs (p_4, p_2), (p_4, p_3), (p_4, p_5) and (p_4, p_6)), and destination-link-based formulation for destination port p_1 (and the OD pairs (p_2, p_1) and (p_3, p_1)) and p_7 (and the OD pairs (p_5, p_7) and (p_6, p_7)), then we only have three origins and destinations. Mathematically, we let $\bar{O}$ be the set of origin ports and $\bar{D}$ be the set of destination ports for the hybrid-link-based model, $\bar{O} \subseteq O$, $\bar{D} \subseteq D$. In Fig. 12.7, $\bar{O} = \{p_4\}$, $\bar{D} = \{p_1, p_7\}$. We further define $\bar{O}^d$ as the set of ports $p \in P$ where the OD pair $(p, d) \in W$ is assigned to destination port $d \in \bar{D}$, and $\bar{D}^o$ as the set of ports $p \in P$ where the OD pair $(o, p) \in W$ is assigned to origin port $o \in \bar{O}$. Hence, in Fig. 12.7, $\bar{D}^{p_4} = \{p_2, p_3, p_5, p_6\}$, $\bar{O}^{p_1} = \{p_2, p_3\}$, and $\bar{O}^{p_7} = \{p_5, p_6\}$.

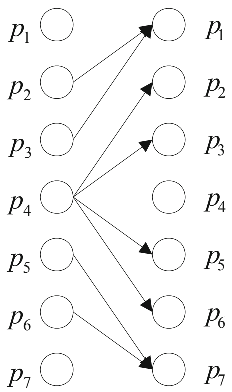


Fig. 12.7 Motivation of the hybrid-link-based model

In the hybrid-link-based model, we use $\hat{z}^o_{ri}$ and $\tilde{z}^o_{ri}$ to represent the total volume of containers with origin port $o \in \bar{O}$ and any destination loaded and discharged at port of call i on ship route r, respectively (transshipped containers are also considered) and use f^o_{ri} to denote the total volume of containers with origin port $o \in \bar{O}$ and any destination flowing on leg i of ship route r. We use $\hat{z}^d_{ri}$ and $\tilde{z}^d_{ri}$ to represent the total volume of containers with destination port $d \in \bar{D}$ and any origin loaded and discharged at port of call i on ship route r, respectively (transshipped containers are also considered) and use f^d_{ri} to denote the total volume of containers with destination port $d \in \bar{D}$ and any origin flowing on leg i of ship route r. y^{od} is the fulfilled demand (TEUs/week) for $(o, d) \in W$; $\hat{z}_p$, $\tilde{z}_p$, and $\bar{z}_p$ are the total volume of loaded, discharged, and transshipped containers (TEUs/week) at port $p \in P$, respectively. The hybrid-link-based formulation is a mixed-integer linear programming problem:

$$\min_{\substack{x_r, \hat{z}^o_{ri}, \tilde{z}^o_{ri}, f^o_{ri}, \hat{z}^d_{ri}, \tilde{z}^d_{ri}, f^d_{ri}, \\ y^{od}, \hat{z}_p, \bar{z}_p, \tilde{z}_p}} \sum_{(o,d) \in W} g^{od} y^{od} - \sum_{r \in R} c_r x_r - \sum_{p \in P} \left(\hat{z}_p \hat{c}_p + \bar{z}_p \bar{c}_p + \tilde{z}_p \tilde{c}_p \right) \quad (12.66)$$

subject to:

$$f^o_{r,i-1} + \hat{z}^o_{ri} = f^o_{ri} + \tilde{z}^o_{ri}, \forall r \in R, \forall i \in I_r, \forall o \in \bar{O} \quad (12.67)$$

$$f^d_{r,i-1} + \hat{z}^d_{ri} = f^d_{ri} + \tilde{z}^d_{ri}, \forall r \in R, \forall i \in I_r, \forall d \in \bar{D} \quad (12.68)$$

$$\hat{z}_p = \begin{cases} \sum\limits_{(p,d) \in W} y^{pd}, \forall p \in O \\ 0, \forall p \in P \backslash O \end{cases} \quad (12.69)$$

$$\tilde{z}_p = \begin{cases} \sum\limits_{(o,p) \in W} y^{op}, \forall p \in D \\ 0, \forall p \in P \backslash D \end{cases} \quad (12.70)$$

$$\bar{z}_p = \sum_{r\in R}\sum_{i\in I_r, p_{ri}=p}\left(\sum_{o\in\bar{O}}\hat{z}^o_{ri} + \sum_{d\in\bar{D}}\hat{z}^d_{ri}\right) - \hat{z}_p, \forall p \in P \tag{12.71}$$

$$\sum_{r\in R}\sum_{i\in I_r, p_{ri}=p}(\hat{z}^o_{ri} - \tilde{z}^o_{ri}) = \begin{cases} \sum\limits_{d\in\bar{D}^o} y^{od}, p = o \\ -y^{op}, p \in \bar{D}^o \\ 0, \text{otherwise} \end{cases} , \forall o \in \bar{O}, \forall p \in P \tag{12.72}$$

$$\sum_{r\in R}\sum_{i\in I_r, p_{ri}=p}(\hat{z}^d_{ri} - \tilde{z}^d_{ri}) = \begin{cases} -\sum\limits_{o\in\bar{O}^d} y^{od}, p = d \\ y^{pd}, p \in \bar{O}^d \\ 0, \text{otherwise} \end{cases} , \forall d \in \bar{D}, \forall p \in P \tag{12.73}$$

$$\sum_{o\in\bar{O}} f^o_{ri} + \sum_{d\in\bar{D}} f^d_{ri} \le V_r x_r, \forall r \in R, \forall i \in I_r \tag{12.74}$$

$$y^{od} \le q^{od}, \forall (o,d) \in W \tag{12.75}$$

$$x_r \in \{0,1\}, \forall r \in R \tag{12.76}$$

$$\hat{z}^o_{ri} \ge 0, \tilde{z}^o_{ri} \ge 0, f^o_{ri} \ge 0, \forall r \in R, \forall i \in I_r, \forall o \in \bar{O} \tag{12.77}$$

$$\hat{z}^d_{ri} \ge 0, \tilde{z}^d_{ri} \ge 0, f^d_{ri} \ge 0, \forall r \in R, \forall i \in I_r, \forall d \in \bar{D} \tag{12.78}$$

$$y^{od} \ge 0, \forall (o,d) \in \mathcal{W} \tag{12.79}$$

The objective function (12.66) maximizes the total profit. Eqs. (12.67) and (12.68) are the flow conservation constraint. Constraint (12.69) defines the total loaded containers at a port. Constraint (12.70) defines the total discharged containers at a port. Constraint (12.71) defines the total transshipped containers at a port. Constraints (12.72) and (12.73) compute the fulfilled demand. Constraint (12.74) enforces the ship capacity constraint. Constraint (12.75) defines the upper limit of the fulfilled demand. Constraint (12.76) defines x_r as a binary variable. Constraint (12.77) defines $\hat{z}^o_{ri}$, $\tilde{z}^o_{ri}$ and f^o_{ri} as nonnegative continuous variables. Constraint (12.78) defines $\hat{z}^d_{ri}$, $\tilde{z}^d_{ri}$ and f^d_{ri} as nonnegative continuous variables. Constraint (12.79) defines y^{od} as a nonnegative continuous variable.

The hybrid-link-based model is at least as compact as the origin-link-based model and the destination-link-based model. Wang (2014) further proposed that finding the least number of origins and destinations $|O| + |D|$ can be solved in polynomial time with regard to the number of ports in the network.

Container Routing Formulation

The above three models aim to select the most efficient subset of ship routes from a set of candidate ship routes to operate. If we are given a set of ship routes to operate, the optimal container routing can be solved by fixing the variables x_r at 1. For example, given a set $\bar{R}$ of ship routes to operate, the container routing problem aims to maximize the total revenue from shipping containers minus the container handling cost. The OD-link-based formulation for container routing is as follows:

$$\max_{\hat{z}_{ri}^{od}, \tilde{z}_{ri}^{od}, f_{ri}^{od}, y^{od}, \hat{z}_p, \bar{z}_p, \tilde{z}_p} \sum_{(o,d)\in\mathcal{W}} g^{od} y^{od} - \sum_{p\in\mathcal{P}} (\hat{z}_p \hat{c}_p + \bar{z}_p \bar{c}_p + \tilde{z}_p \tilde{c}_p) \tag{12.80}$$

subject to:

$$f_{r,i-1}^{od} + \hat{z}_{ri}^{od} = f_{ri}^{od} + \tilde{z}_{ri}^{od}, \forall r \in \bar{R}, \forall i \in I_r, \forall (o,d) \in \mathcal{W} \tag{12.81}$$

$$\hat{z}_p = \sum_{(p,d)\in\mathcal{W}} y^{pd}, \forall p \in \mathcal{P} \tag{12.82}$$

$$\tilde{z}_p = \sum_{(o,p)\in\mathcal{W}} y^{op}, \forall p \in \mathcal{P} \tag{12.83}$$

$$\bar{z}_p = \sum_{r\in\overline{\mathcal{R}}} \sum_{i\in I_r, p_{ri}=p} \sum_{(o,d)\in\mathcal{W}} \hat{z}_{ri}^{od} - \hat{z}_p, \forall p \in \mathcal{P} \tag{12.84}$$

$$\sum_{r\in\overline{\mathcal{R}}} \sum_{i\in I_r, p_{ri}=p} (\hat{z}_{ri}^{od} - \tilde{z}_{ri}^{od}) = \begin{cases} y^{od}, p = o \\ -y^{od}, p = d \\ 0, \text{otherwise} \end{cases}, \forall (o,d) \in \mathcal{W}, \forall p \in \mathcal{P} \tag{12.85}$$

$$\sum_{(o,d)\in\mathcal{W}} f_{ri}^{od} \le V_r, \forall r \in \overline{R}, \forall i \in I_r \tag{12.86}$$

$$y^{od} \le q^{od}, \forall (o,d) \in \mathcal{W} \tag{12.87}$$

$$\hat{z}_{ri}^{od} \ge 0, \tilde{z}_{ri}^{od} \ge 0, f_{ri}^{od} \ge 0, \forall r \in \overline{R}, \forall i \in I_r, \forall (o,d) \in \mathcal{W} \tag{12.88}$$

$$y^{od} \ge 0, \forall (o,d) \in \mathcal{W} \tag{12.89}$$

Note that in the objective function (12.80), the component $\sum_{r\in\overline{\mathcal{R}}} c_r x_r$ is fixed and hence is not incorporated. Path-based, origin-link-based, destination-link-based and hybrid-link-based models for container routing can be formulated in a similar manner.

In contrast to the ship route selection problem that is formulated as a mixed-integer linear programming model, the optimal container routing problem is a linear programming model and dual variables provide useful information for ship route generation. Finally, we note that Brouer et al. (2011) compared the computational efficiency of solving the OD-link-based container routing model using CPLEX's barrier method, and solving the path-based container routing model using a delayed column generation algorithm. Their numerical experiments demonstrate that the latter approach is more efficient. It should be mentioned that in Brouer et al. (2011) a multi-period planning horizon was considered and the demands in different periods are different. As a result, the network is very sparse. This is a different setting from the weekly demand in the models [P1] to [P4] mentioned above.

12.6.2 Generating Ship Routes

Álvarez (2009) pointed out a few difficulties for directly using the ship route selection models to design container transport network by enumerating all ship routes in set $\mathcal{R}$. First, the set $\mathcal{R}$ is too large to enumerate, not to mention of including it in a mixed-integer linear programming model. Second, since larger ships have economies of scale, the linear programming relaxation will favor the use of fractions of the largest ships available. As a consequence of the difficulties, all the existing studies have worked on only a small subset of all possible ship routes, and iteratively improved the quality of the subset. There are three slightly different approaches in terms of ship route generation, as described below.

Agarwal and Ergun (2008) considered the linear programming relaxation of the ship route selection problem [P2], and the resulting model is as follows:

$$\max_{x_r, \hat{z}_{ri}^{od}, \tilde{z}_{ri}^{od}, f_{ri}^{od}, y^{od}, \hat{z}_p, \bar{z}_p, \tilde{z}_p} \sum_{(o,d)\in\mathcal{W}} g^{od} y^{od} - \sum_{r\in\mathcal{R}} c_r x_r - \sum_{p\in\mathcal{P}} (\hat{z}_p \hat{c}_p + \bar{z}_p \bar{c}_p + \tilde{z}_p \tilde{c}_p) \tag{12.90}$$

subject to:

$$f_{r,i-1}^{od} + \hat{z}_{ri}^{od} = f_{ri}^{od} + \tilde{z}_{ri}^{od}, \forall r \in R, \forall i \in I_r, \forall (o,d) \in \mathcal{W} \tag{12.91}$$

$$\hat{z}_p = \sum_{(p,d)\in\mathcal{W}} y^{pd}, \forall p \in \mathcal{P} \tag{12.92}$$

$$\tilde{z}_p = \sum_{(o,p)\in\mathcal{W}} y^{op}, \forall p \in \mathcal{P} \tag{12.93}$$

$$\bar{z}_p = \sum_{r\in\mathcal{R}} \sum_{i\in I_r, p_{ri}=p} \sum_{(o,d)\in\mathcal{W}} \hat{z}_{ri}^{od} - \hat{z}_p, \forall p \in \mathcal{P} \tag{12.94}$$

$$\sum_{r\in\mathcal{R}} \sum_{i\in I_r, p_{ri}=p} (\hat{z}_{ri}^{od} - \tilde{z}_{ri}^{od}) = \begin{cases} y^{od}, p = o \\ -y^{od}, p = d \\ 0, \text{otherwise} \end{cases}, \forall (o,d) \in \mathcal{W}, \forall p \in \mathcal{P} \tag{12.95}$$

$$\sum_{(o,d)\in\mathcal{W}} f_{ri}^{od} \le V_r x_r, \forall r \in R, \forall i \in I_r \tag{12.96}$$

$$y^{od} \le q^{od}, \forall (o,d) \in \mathcal{W} \tag{12.97}$$

$$0 \le x_r \le 1, \forall r \in R \tag{12.98}$$

$$\hat{z}_{ri}^{od} \ge 0, \check{z}_{ri}^{od} \ge 0, f_{ri}^{od} \ge 0, \forall r \in R, \forall i \in I_r, \forall (o,d) \in \mathcal{W} \tag{12.99}$$

$$y^{od} \ge 0, \forall (o,d) \in \mathcal{W} \tag{12.100}$$

The only difference between [P2] and [P5] lies in that x_r is relaxed as a continuous variable in Eq. (12.98) of [P5]. Note that the original model of Agarwal and Ergun (2008) is different from [P5] as (i) the former considers the demand in each day rather than each week; (ii) the transshipment cost is not incorporated; and (iii) a Benders-decomposition based formulation is applied to improve the efficiency.

[P5] is a linear programming model and hence dual variables associated with the constraints could be obtained. Based on the dual variables, Agarwal and Ergun (2008) obtained profitable ship routes that are not included in set $\mathcal{R}^{(\mathrm{k})}$ in iteration k via a column generation scheme. The newly generated ship routes are added to the set $\mathcal{R}^{(\mathrm{k})}$. Hence, the set $\mathcal{R}^{(\mathrm{k})}$ is becoming larger and larger as the algorithm progresses (Note that in the implementation of Agarwal and Ergun (2008), ship routes may also be excluded from $\mathcal{R}^{(\mathrm{k})}$). Finally, integrality constraints on x_r is imposed and model [P2] is solved to find the optimal subset of ship routes to operate.

Álvarez (2009) and Brouer et al. (2014) have considered a set $\bar{R}^{(k)}$ of ship routes that are operated in iteration k. To evaluate the quality of the set $\bar{R}^{(k)}$, they solve a multicommodity flow problem [P4] (note that they have adopted the destination-link-based rather than the OD-link-based formulation). Two types of information obtained from [P4] are taken advantage of: ship route utilization (ratio of used capacity and deployed capacity) and dual variables associated with constraint (12.86). Intuitively, a ship route with low utilization should be discarded or modified, and a larger dual variable associated with constraint (12.86) implies that increasing the capacity of ships on the leg could bring in greater profit. The dual variables are incorporated in a separate column generation problem to obtain new ship routes. A Tabu search is applied to select the most promising new set $\bar{R}^{(k+1)}$ for evaluation in the next iteration.

Wang et al. (2013a) solved [P1] in each iteration k of the algorithm and obtained the optimal subset of ship routes from $\mathcal{R}^{(\mathrm{k})}$ to operate. The set of ship routes that are not chosen are then excluded from $\mathcal{R}^{(\mathrm{k})}$. Since [P1] is a mixed-integer linear programming model, dual information is no longer available. As a consequence, Wang et al. (2013a) designed new ship routes based on the container flow information on the chosen ship routes. For example, if there is considerable unshipped demand between ports, new ship routes are constructed to fulfill the demand; if the number of laden and empty containers handled at some ports of call is very small, for example, fewer than ten containers, then these ports of call can be removed; if the capacity utilization of ship route is very large or small, ships with a string of larger or smaller

Table 12.3 Comparison of algorithmic approaches for general network design

Aspect	Agarwal and Ergun (2008)	Álvarez (2009) and Brouer et al. (2014)	Wang et al. (2013a)
Model providing information for generating new ship routes	Linear programming model [P5]	Linear programming model [P4]	Mixed-integer linear programming model [P1]
Information for generating new ship routes	Dual variables	Capacity utilization and dual variables	All container flow information
How new ship routes are designed	Column generation based heuristic	Column generation based heuristic and Tabu search	Rule-based heuristic
Size of the set of ship routes for evaluation in each iteration	The size keeps increasing after each iteration	The size of is almost constant throughout the iterations	The size is almost constant throughout the iterations
Key challenges	The size of the mixed-integer linear programming model [P2] to be solved at the end is very large	How to improve the efficiency of a network by finding good neighborhood solutions	How to improve the efficiency of a network by finding good neighborhood solutions
Nature and size of the problem tested	Randomly generated instances with at most 20 ports	Randomly generated case study with 120 ports in Álvarez (2009) and Real cases based on data from Maersk Line with at most 197 ports in Brouer et al. (2014)	Real case study based on data from a global liner shipping company with more than 150 ports

ships may be deployed, respectively; if the capacity utilization of a feeder ship route is low, the feeder ports included in the feeder ship route may be removed from existing line-haul ship routes and a new feeder port may be added to the feeder ship route based on the geographical location and unshipped demand; if the capacity utilization of a line-haul ship route is low, a new port may be added to the ship route based on the unshipped demand. The set $\mathcal{R}^{(k+1)}$ is the union of $\mathcal{R}^{(k)}$ and the newly generated ship routes. Hence, $\mathcal{R}^{(k+1)}$ is at least as good as $\mathcal{R}^{(k)}$. The above process is repeated until the stop criterion is satisfied.

A comparison of the three approaches is summarized in Table 12.3.

12.6.3 An Exact Approach

In contrast to the heuristic approaches in Agarwal and Ergun (2008), Álvarez (2009), Brouer et al. (2014), and Wang et al. (2013a); Plum et al. (2014b) took the initiative to develop a compact formulation of the liner shipping network design problem based on “service flows”. The formulation could handle multiple calls to the same port, which are popular for butterfly ship routes. They introduced service nodes, together with port nodes in a graph representation of the problem, and numbered arcs between a port and a dummy service node. An arc from a port node to a service node indicates whether a service is calling the port or not. This representation allows recurrent calls of a service to a port. By imposing upper bounds on the number of services and the number of ports of call on a service, the problem is formulated as a mixed-integer linear programming model, which is solved by existing solvers. The model is solved for the two smallest instances of the benchmark suite in Brouer et al. (2014). Although this approach may not be applicable for slightly larger problem instances, it is an interesting attempt to solve such a difficult problem to optimality.

12.7 Future Research Directions

Despite more and more research works devoted to container transport network design in recent years (Christiansen et al. 2004; 2013; Meng et al. 2014), there are still a number of research avenues to explore in future. In particular, solving practical-size problems while capturing essential operating characteristics, developing efficient solution algorithms, and application by liner shipping companies are three worthwhile research directions.

First, there are a number of essential operating characteristics that are seldom touched by existing studies on network design, for instance, (i) intermodal container transport network design that consists of both inland and maritime networks; (ii) collaboration and competition between liner shipping companies; (iii) how to handle uncertainties in the shipping environment, such as port time uncertainty, demand uncertainty, and bunker price uncertainty; (iv) green shipping in view of the more and more stringent emission regulations imposed by International Maritime Organization (IMO) and governments.

Second, there is large room for designing efficient solution algorithms for global container transport network design. The global container transport network design looks similar to the well-known vehicle routing problem (VRP), however, the former is more challenging due to regular service frequency and container transshipment. We expect that a number of operations research methods will be developed in global container transport network design, which will enrich the operations research theory. To date there are few exact solution methods for global container transport network design. Reinhardt and Pisinger (2012) contributed an exact branch-and-cut approach. Nevertheless, this approach could only address special network design problems of relatively small size. Plum et al. (2014a) proposed an exact branch-and-cut-and-price algorithm for designing a single ship route without container transshipment. Plum et al. (2014b) developed a compact mixed-integer linear programming model that could be solved to optimality for designing small-sized networks. The quality of the heuristic approaches in Agarwal and Ergun (2008), Álvarez (2009), Brouer et al.

(2014) and Wang et al. (2013a) may be hard to evaluate. On one side, one could not find the optimal solution to a large liner shipping network design problem. On the other hand, the operating data from liner shipping companies are disorganized, scattered in different departments, and confidential to the academia, and hence it is difficult to compare in great details the designed network by algorithms and the network operated by shipping lines.

Finally, we highlight that the ultimate objective for developing models and algorithms for global container transport network design is to improve the profitability of liner shipping companies. Some research works that are applied by liner shipping companies may not be published due to confidentiality, and some works have more theoretical contributions than practical significance because of the lack of input from industrial partners. Brouer et al. (2014) contributed a seminal benchmark suite for global container transport network design. This benchmark suite is designed based on the operating data from the largest liner shipping company Maersk Line. As a result, researchers could compare different solution algorithms using the benchmark suite. More importantly, when all or many researchers conduct their works based on the operating data from the largest liner shipping company in the world, academic works will arouse much more attention from the liner shipping industry.

References

Agarwal, R., & Ergun, O. (2008). Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science, 42*(2), 175–196.

Agarwal, R., & Ergun, O. (2010). Network design and allocation mechanisms for carrier alliances in liner shipping. *Operations Research, 58*(6), 1726–1742.

Alphaliner (2014). TOP 100 operated fleets as per 2 April 2014. http://www.alphaliner.com/top100/. Accessed 2 April 2014.

Álvarez, J. F. (2009). Joint routing and deployment of a fleet of container vessels. *Maritime Economics & Logistics, 11*(2), 186–208.

APL. (2013) Asia-Europe Routes. http://www.apl.com/routes/html/asia_europe.html. Accessed 30 April, 2013.

Brouer, B. D., Pisinger, D., & Spoorendonk, S. (2011). Liner shipping cargo allocation with repositioning of empty containers. *INFOR, 49*(2), 109–124.

Brouer, B. D., Alvarez, J. F., Plum, C., Pisinger, D., & Sigurd, M. (2014). A base integer programming model and benchmark suite for linear shipping network design. *Transportation Science. 48*(2), 281–312.

Christiansen, M., Fagerholt, K., & Ronen, D. (2004). Ship routing and scheduling: Status and perspectives. *Transportation Science, 38*(1), 1–18.

Christiansen, M., Fagerholt, K., Nygreen, B., & Ronen, D. (2013). Ship routing and scheduling in the new millennium. *European Journal of Operational Research, 228*(3), 467–478.

Fagerholt, K. (1999). Optimal fleet design in a ship routing problem. *International Transactions in Operational Research, 6*(5), 453–464.

Fagerholt, K. (2004). Designing optimal routes in a liner shipping problem. *Maritime Policy and Management, 31*(4), 259–268.

Gelareh, S., & Pisinger, D. (2011). Fleet deployment, network design and hub location of liner shipping companies. *Transportation Research Part E, 47*(6), 947–964.

Gelareh, S., Nickel, S., & Pisinger, D. (2010). Liner shipping hub network design in a competitive environment. *Transportation Research Part E, 46*(6), 991–1004.

Gelareh, S., Maculan, N., Mahey, P., & Monemi, R. N. (2013a). Hub-and-spoke network design and fleet deployment for string planning of liner shipping. *Applied Mathematical Modeling, 37*(5), 3307–3321.

Gelareh, S., Monemi, R. N., Maculan, N., Mahey, P., & Pisinger, D. (2013b). Single string planning problem arising in liner shipping industries: A heuristic approach. *Computer and Operations Research, 40*(10), 2357–2373.

Imai, A., Shintani, K., & Papadimitriou, S. (2009). Multi-port vs. Hub-and-Spoke port calls by containerships. *Transportation Research Part E, 45*(5), 740–757.

Karlaftis, M. G., Kepaptsoglou, K., & Sambracos, E. (2009). Containership routing with time deadlines and simultaneous deliveries and pick-ups. *Transportation Research Part E, 45*(1), 210–221.

Maersk L. (2013a). Daily Maersk. http://dailymaersk.com/. Accessed 18 Feb 2013a.

Maersk L. (2013b). Service Schedules. https://www.maerskline.com/frameset.jsp. Accessed 30 May 2013b.

Meng, Q., & Wang, S. (2011). Liner shipping service network design with empty container repositioning. *Transportation Research Part E, 47*(5), 695–708.

Meng, Q., Wang, S., & Liu, Z. (2012a). Network design for shipping service of large-scale intermodal liners. *Transportation Research Record, 2269,* 42–50.

Meng, Q., Wang, T., & Wang, S. (2012b). Short-term liner ship fleet planning with container transshipment and uncertain demand. *European Journal of Operational Research, 223*(1), 96–105.

Meng, Q., Wang, S., Andersson, H., & Thun, K. (2014). Containership routing and scheduling in liner shipping: Overview and future research directions. *Transportation Science. 48*(2), 265–280.

Notteboom, T. E. (2004). A carrier's perspective on container network configuration at sea and onland. *Journal of International Trade and Logistics, 2*(1), 65–87.

Notteboom, T. E. (2006). The time factor in liner shipping services. *Maritime Economics and Logistics, 8*(1), 19–39.

Plum, C. E. M., Pisinger, D., Salazar-González, J. J., & Sigurd, M. M. (2014a). Single liner shipping service design. *Computers & Operations Research, 45,* 1–6.

Plum, C. E. M., Pisinger, D., & Sigurd, M. M. (2014b). A service flow model for the liner shipping network design problem. *European Journal of Operational Research, 235*(2), 378–386.

Qi, X., & Song, D. P. (2012). Minimizing fuel emissions by optimizing vessel schedules in liner shipping with uncertain port times. *Transportation Research Part E, 48*(4), 863–880.

Rana, K., & Vickson, R. G. (1988). A model and solution algorithm for optimal routing of a time-chartered containership. *Transportation Science, 22*(2), 83–95.

Rana, K., & Vickson, R. G. (1991). Routing container ships using Lagrangean relaxation and decomposition. *Transportation Science, 25*(3), 201–214.

Reinhardt, L. B., & Pisinger, D. (2012). A branch and cut algorithm for the container shipping network design problem. *Flexible Services and Manufacturing Journal, 24*(3), 349–374.

Ronen, D. (2011). The effect of oil price on containership speed and fleet size. *Journal of the Operational Research Society, 62*(1), 211–216.

Sambracos, E., Paravantis, J. A., Tarantilis, C. D., & Kiranoudis, C. T. (2004). Dispatching of small containers via coastal freight liners: The case of the Aegean Sea. *European Journal of Operational Research, 152*(2), 365–381.

Shintani, K., Imai, A., Nishimura, E., & Papadimitriou, S. (2007). The container shipping network design problem with empty container repositioning. *Transportation Research Part E, 43*(1), 39–59.

Song, D.-P., & Dong, J.-X. (2012). Cargo routing and empty container repositioning in multiple shipping service routes. *Transportation Research Part B, 46*(10), 1556–1575.

Song D.-P., & Dong, J.-X. (2013). Long-haul liner service route design with ship deployment and empty container repositioning. *Transportation Research Part B, 55,* 188–211.

UNCTAD. (2012). Review of Maritime Transportation. Paper presented at the United Nations Conference on Trade and Development. New York and Geneva. http://unctad.org/en/PublicationsLibrary/rmt2012_en.pdf. Accessed 23 April 2013.

Wang, S. (2013). Essential elements in tactical planning models for container liner shipping. *Transportation Research Part B, 54,* 84–99.

Wang, S. (2014). A novel hybrid-link-based container routing model. *Transportation Research Part E, 61,* 165–175.

Wang, S., & Meng, Q. (2012a). Liner ship fleet deployment with container transshipment operations. *Transportation Research Part E, 48*(2), 470–484.

Wang, S., & Meng, Q. (2012b). Sailing speed optimization for container ships in a liner shipping network. *Transportation Research Part E, 48*(3), 701–714.

Wang, S., & Meng, Q. (2013). Reversing port rotation directions in a container liner shipping network. *Transportation Research Part B, 50,* 61–73.

Wang, S., & Meng, Q. (2014). Liner shipping network design with deadlines. *Computers & Operations Research, 41,* 140–149.

Wang, S., Liu, Z., & Meng, Q. (2013a). Systematic network design for liner shipping services. *Transportation Research Record, 2330,* 16–23.

Wang, S., Meng, Q., & Liu, Z. (2013b). Containership scheduling with transit-time-sensitive container shipment demand. *Transportation Research Part B, 54,* 68–83.

Wang, S., Meng, Q., & Sun, Z. (2013c). Container routing in liner shipping. *Transportation Research Part E, 49*(1), 1–7.

World Shipping Council. (2014). Top 50 world container ports. http://www.worldshipping.org/about-the-industry/global-trade/top-50-world-container-ports. Accessed 2 April 2014.

Yang, Z., Chen, K., & Notteboom, T. (2012). Optimal design of container liner services: Interactions with the transport demand in ports. *Maritime Economics and Logistics, 14*(4), 409–434.

Zheng, J., Sun, Z., & Meng, Q. (2014). Impact analysis of maritime cabotage legislations on liner hub-and-spoke shipping network design. *European Journal of Operational Research, 223*(3), 874–884.

Part III
Shippers and Global Supply Chain Management

Chapter 13
Purchasing Transportation Services from Ocean Carriers

Zhou Xu and Xiaofan Lai

Abstract Reducing transportation costs is priority number one for global shippers who need to move their cargo containers all over the world. To achieve such cost reduction, a shipper can use what is called a reverse auction mechanism to purchase transportation services, by inviting carriers, i.e. liner shipping companies, to bid competitively to sell their services. As part of the process, carriers often seek commitments from the shipper, and internal business units of the shipper often express their own preferences when it comes choosing the carriers, which naturally complicates the shipper's decisions. In this chapter, we first review existing studies on the transportation service procurement problem. Based on a new general optimization model, we then discuss extensions to the existing known results, as well as present several results new to the literature.

13.1 Introduction

In the container shipping market, the key players are shippers (as buyers) and carriers (as sellers), where shippers, such as manufacturers and retailers, are companies who need to move their cargo containers, and carriers, such as shipping liners, are companies who provide transportation services to ship the containers. With the huge expansion in global supply chains, shippers today have a huge demand for transportation services from carriers to transport their cargo, which may include raw materials or finished products. As a result, transportation services are often listed among the top categories of spending by global shippers, providing large opportunities for cost savings (Xu 2007). In fact, it is common for a global shipper to spend more than a \$ 100million annually on transportation services (Lim et al. 2012).

Z. Xu (✉) · X. Lai
Department of Logistics and Maritime Studies, The Hong Kong
Polytechnic University, Kowloon, Hong Kong
e-mail: lgtzx@polyu.edu.hk

X. Lai
e-mail: xiaofan.lai.lms@connect.polyu.hk

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_13

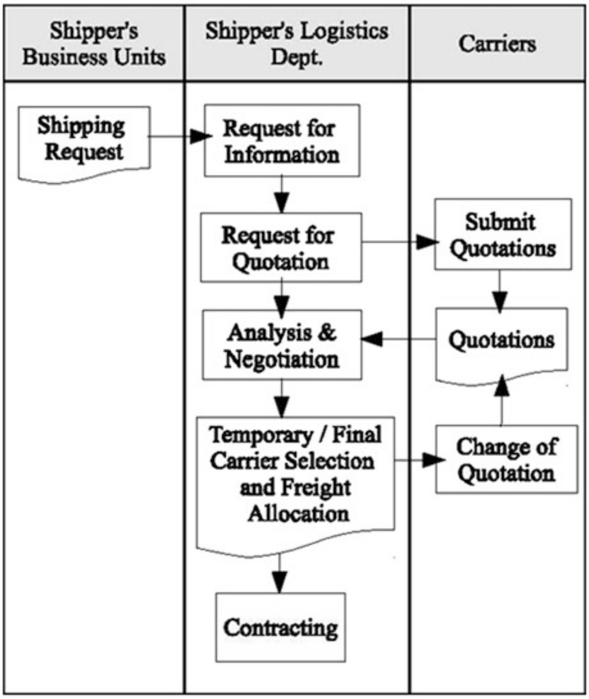


Fig. 13.1 Reverse auction mechanism for transportation service procurement

For shippers, the transportation services are often purchased by their logistics departments, and typically follow a reverse auction mechanism that consists of the following four stages (Xu 2007), as shown in Fig. 13.1:

- Stage 1—Request for Information (RFI): The logistics department collects shipping request information from different business units (or departments) of the company, and based on the information collected, it then forecasts cargo volumes for the coming period.
- Stage 2—Request for Quotation (RFQ): The logistics department invites a number of carriers to make quotations of shipping prices for lanes between different origin-destination pairs, for different service levels in terms of shipping times, and for different weights of cargo, etc.
- Stage 3—Analysis and Negotiation: The logistics department analyzes the quotations from the carriers, estimates the total transportation cost under different scenarios, and negotiates with the carriers by bargaining over the shipping prices and conditions.
- Stage 4—Signing Contract: The logistics department makes its decisions on the selection of carriers and the allocation of shipping volumes to the selected carriers, so as to finalize the prices and conditions with the carriers and then sign contracts.

Before finalizing and signing contracts, the shipper may go through multiple rounds of analysis and negotiations, with a view to minimizing the total transportation cost.

During the above process, particularly at the various stages of analysis and negotiation, as well as at the stage of signing contracts, a shipper often needs to solve optimization problems with regard to selecting carriers and allocating cargo to the selected carriers in order to minimize its total transportation cost. Such problems are challenging, since they are often triggered by various constraints that reflect different business considerations, some imposed by the external carriers who provide the shipping services (Lim et al. 2008a, 2006), and others imposed by the internal business units who require the shipping services (Lim et al. 2012). Moreover, unlike the trading of physical goods, shipping services are of a combinatorial nature, as the price of shipping services is often imposed not only on a single lane but also on a group of lanes of different origin-and-destination pairs (Lim et al. 2008a). It is also well-known that the shipping market is very volatile, as both the demands and the sport-market shipping prices vary significantly all the time. Due to this, the key players, including both shippers and carriers, sometimes may not strictly follow the contracts in actual operation, at times maybe breaking them to suit their own interests.

The procurement of transportation services is challenging, and has raised several interesting research questions that fall into the following three categories:

1. On Models: How should the problems be defined? What are the useful properties of the optimal solutions to these problems? To answer these questions, it is necessary to formulate the corresponding optimization or decision problems as mathematical programming models, as well as to analyze the properties of the models.
2. On Tractability: Do efficient algorithms exist that can solve the problems to optimality? Answering this question, it requires an understanding of the computational complexities of the corresponding optimization or decision problems, as well as being able to identify special cases that have practical applications and that can also be efficiently solved to optimality.
3. On Algorithms: How can exact or near optimal solutions to the problems be found in affordable running times? To answer this question, it is necessary to develop exact algorithms or heuristic algorithms, and to show by either theoretical analysis or numerical experiments that these algorithms can guarantee good performance.

In this chapter, we review recent studies that have addressed some of these research questions related to transportation service procurement problems. Since most of the existing studies are focused only on problems with specific constraints, it is also of interest to know how their results, models, and algorithms can be extended to solving more general problems having broader applications. For this purpose, we also introduce in this chapter a generalized optimization model for the transportation service procurement problem, which is defined and formulated in Sect. 2, and we then discuss how existing results from the literature can be applied to this generalized model. These results include computational complexities, relaxations of mathematical models, exact algorithms, and heuristic algorithms, which are all discussed in Sects. 3–6, respectively. The chapter is concluded in Sect. 7 with discussions on future research directions.

13.2 Problem Formulations: Generic Model, Side Constraints, and Generalization

13.2.1 Generic Model

Consider a shipper who has to make decisions on purchasing transportation services to move containers of its cargo for a time horizon of T periods. The shipper has a set of business units, denoted by B, who use the services potentially from m carriers, which are denoted by $I = \{i{:}1 \le i \le m\}$. The cargo for each period in the time horizon is for n lanes in total, these being defined as pairs of the cargo's origins and destinations, and are denoted by $J = \{j{:}1 \le j \le n\}$. At the beginning of the time horizon, the shipper collects information from its business units so as to forecast demands for transportation services for each period $t \in \{1, 2, \ldots, T\}$, which are given by $d_{bjt} \in R_+$ for each business unit $b \in B$ and each lane j, where R_+ indicates the set of non-negative real numbers. The forecast demand d_{bjt} is then released to the carriers, and each carrier i responds by quoting a quote of a shipping rate $p_{ijt} \in R_+$ according to its bidding strategy. In addition to such a reverse auction mechanism, the shipper can also purchase shipping services from the spot market, so as to minimize its total shipping cost. The forecast spot market shipping rates are represented by $s_{jt} \in R_+$. To reflect the fact that carriers often quote prices on groups of lanes, for each carrier i we assume that the lanes that it operates are partitioned into a collection $J_i = \{J_{i1}, J_{i2}, \ldots, J_{i,|J|}\}$ of lane groups, where $J_{ik} \subseteq J$ for $1 \le k \le |J_i|$. We assume that lane groups are disjoint. Let $c_{ik} \in Z_+$ indicate the capacity that carrier i can ship for all the lanes in J_{ik} for all the T periods of the planning time horizon.

The shipper needs to make the following decisions, so as to minimize its total transportation cost:

- Decisions on selecting carriers: This can be represented by binary variables $y_{ibk} \in \{0, 1\}$ and binary variables y_{ik}, where $y_{ibk} = 1$ if and only if carrier i is selected to serve lanes in J_{ik} for the cargo of business unit b, and $y_{ik} = 1$ if and only if carrier i is selected to serve lanes in J_{ik};
- Decisions on allocating cargo to carriers: This can be represented by variables $x_{ibjt} \in \mathrm{Z}_+$, where Z_+ is the set of non-negative integers, and each x_{ibjt} indicates the number of containers allocated to carrier i for shipping the cargo of business unit b on lane j in period t.

Thus, the total transportation cost for the shipper is $\sum_{b\in B}\sum_{j\in J}\sum_{t=1}^{T} [\sum_{i\in I} p_{ijt}x_{ibjt} + s_{jt}(d_{bjt} - \sum_{i\in I} x_{ibjt})]$. The generic integer programming model (Generic IP) for the transportation service procurement problem, or in short the *TSPP*, can then be formulated as follows:

$$(Generic\ IP) = \min \sum_{b\in B}\sum_{j\in J}\sum_{t=1}^{T}\left[\sum_{i\in I} p_{ijt}x_{ibjt} + s_{jt}\left(d_{bjt} - \sum_{i\in I} x_{ibjt}\right)\right] \quad (13.1)$$

$$\text{s.t.} \sum_{i\in I} x_{ibjt} \le d_{bjt}, \forall b \in B, j \in J, 1 \le t \le T, \quad (13.2)$$

$$\sum_{b \in B} \sum_{j \in J} \sum_{t=1}^{T} x_{ibjt} \leq c_{ik} y_{ik}, \forall i \in I, 1 \leq k \leq |J_i|, \tag{13.3}$$

$$\sum_{b \in B} y_{ibk} \leq |B| y_{ik}, \forall i \in I, 1 \leq k \leq |J_i|, \tag{13.4}$$

$$x_{ibjt} \in Z_+, \forall i \in I, b \in B, j \in J, 1 \leq t \leq T, \tag{13.5}$$

$$y_{ik} \in \{0{,}1\}, \forall i \in I, 1 \leq k \leq |J_i|, \tag{13.6}$$

$$y_{ibk} \in \{0{,}1\}, \forall i \in I, b \in B, 1 \leq k \leq |J_i|. \tag{13.7}$$

In the (Generic IP), the objective (1) is to minimize the total transportation cost for the shipper. Constraint (2) ensures that for each lane j, period t and business unit b, the total volume allocated to all carriers does not exceed the total demand d_{bjt} of b. Constraint (3) ensures that the total volume of cargo allocated to carriers can never exceed their capacities for each of their lane groups. Constraint (4) ensures that y_{ibk} equals zero as long as y_{ik} equals zero. Constraint (5) restricts shipment allocations to be integers, because shippers are often required to buy spaces of full size containers in TEUs (twenty-foot equivalent units), which can be expensive. It is not difficult to see that decision variables y_{ibk} and y_{ik} are redundant in the (Generic IP). We leave y_{ibk} and y_{ik} to model various side constraints that reflect different business considerations in our formulations in the latter part of this chapter. It can be seen that the model above has taken into account carriers' lane groups, and has formulates basic demand and capacity constraints

13.2.2 Side Constraints

There are two main sources of side constraints, those from external carriers and those from internal business units (Xu 2007).

As for external carriers, in addition to the basic capacity constraint, they often also request for a volume guarantee from the shipper. When carriers submit quotations to the shipper, they make assumptions about demand, and their quotations can be either too low or too high (Caplice and Sheffi 2005). As a result, the carrier that wins the quotation can be the one that underestimates its service cost, and will thus suffer from the "winner's curse" (Caplice 2003; Sawhney 2003). A volume guarantee from the shipper can thus be helpful in resolving this problem, since it can reduce risk and uncertainty for the carriers, which enables the carriers to bid in a more realistic manner.

There are two types of volume guarantee that are commonly requested by carriers and have been studied in the literature. One is called a minimum quantity commitment (MQC) constraint, which is motivated by the stipulation of the United States Federal Maritime Commission that restricts a fixed minimum quantity for the total volume of cargo shipped by each carrier to cities in the US (Lim et al. 2006). The commitment to a minimum quantity has commonly been applied in transportation service procurement for various shippers, where carriers negotiate and quote for a minimum volume for each lane group. Let $q_{ik} \in R_+$ indicate the minimum quantity for shipments for lanes in J_{ik}, which limits the volume for carrier i to carry in all the periods to be either none or above q_{ik}. Accordingly, the MQC constraint can be formulated as follows:

$$q_{ik} y_{ik} \leq \sum_{b \in B} \sum_{j \in J_{ik}} \sum_{t=1}^{T} x_{ibjt}, \forall i \in I, 1 \leq k \leq |J_i|, \tag{13.8}$$

where the right hand side of (13.8) is the total volume of cargo allocated to carrier i for lanes in J_{ik}.

The other type of volume guarantee constraint is called the maximum-to-average ratio commitment (MARC), which is motivated by common practice in the shipping industry (Lim et al. 2008a). This commitment requires that the volume of each shipment that the shipper can ship through a carrier cannot exceed a fixed proportion of the average volume shipped through the carrier during the term of the contract. This proportion is usually referred to as the *maximum-to-average ratio*, which is quoted by the carrier and can be negotiated with the shipper. Under this condition, the shipper has to buy sufficient volume from the carrier that can be spread over the duration. This is because the shipper's shipments at any time can be made only if commensurate volume is shipped throughout the duration. For example, the shipper must ship sufficient volume in off-peak seasons if shipments are planned for peak seasons. This stipulation translates into a volume guarantee to the carrier which, in effect, smoothes shipments throughout the duration of the contract.

To formulate a constraint for the (MARC), let us take $r_{ikt} \in R_+$ to be the maximum-to-average ratio for shipments in J_{ik}, which limits the volume that carrier i is willing to carry in period t in excess of the average volume that it will carry for all T periods of the contract (Lim et al. 2008a). We take $\xi \in Z_+$ to represent a small excess that is allowed in the maximum-to-average ratio commitment in practice. Letting $\alpha_{ikt} = r_{ikt}/T$, the constraint related to the maximum-to-average ratio commitment can be formulated as follows:

$$\sum_{b \in B} \sum_{j \in J_{ik}} x_{ibjt} \leq \alpha_{ikt} \left(\sum_{b \in B} \sum_{j \in J_{ik}} \sum_{t=1}^{T} x_{ibjt} \right) + \xi, \forall i \in I, 1 \leq k \leq |J_i|, 1 \leq t \leq T, \tag{13.9}$$

where the left hand side represents the total volume of cargo shipped by carrier i during period t for lanes in J_{ik}, and the right hand side is the average volume shipped by carrier j and for lanes in J_{ik} of all T periods.

Besides constraints from carriers, the shipper's decision maker often also needs to satisfy constraints from its own internal business units, the ones who are the actual users of transportation service for shipping and receiving cargo. In the literature, three types of such constraints have been introduced and studied. One is the carrier number constraint, which imposes a lower bound and/or an upper bound on the number of carriers selected for a group of business units and a group of lanes (Lim et al. 2012). This is motivated by practical considerations, whereby a smaller number of carriers may reduce the management cost, but it also restricts the flexibility of business units in choosing shipping dates. To formulate this constraint, let BG denote a collection of groups of business units, and LG denote a collection of lane groups. For each group of business unit $B_h \in BG$, and lane group $L_g \in LG$, let $\underline{n}_{hg}$ and $\bar{n}_{hg}$ indicate the minimum and maximum numbers of carriers that can be selected to ship cargo for business units in B_h and for lanes in L_g. The carrier number constraint can be formulated as follows:

$$\underline{n}_{hg} \leq \sum_{i \in I} z_{ihg} \leq \bar{n}_{hg}, \forall B_h \in BG, L_g \in LG, \tag{13.10}$$

$$\sum_{b \in B_h} \sum_{k: J_{ik} \cap L_g \neq \varnothing} y_{ibk} \leq M z_{ihg}, \forall i \in I, B_h \in BG, L_g \in LG, \tag{13.11}$$

$$z_{ihg} \in \{0,1\}, \forall i \in I, B_h \in BG, L_g \in LG, \tag{13.12}$$

where z_{ihg} indicates whether or not carrier i is selected to ship cargo for business units in B_h and for lanes in L_g, and M is a sufficiently large constant.

The second type of shipper's constraint is called the preference constraint (Xu 2007), under which cargo of a certain business unit cannot be assigned to particular carriers, or must be assigned to particular other carriers. This is motivated by current practice, where business units have their own preferences as to the choice of carriers based on their previous experience. To impose such constraints, the logistics department needs to collect preference information from business units. For each group of business units $B_h \in BG$, and lane group $L_g \in LG$, let I_{hg}^{+} indicate a set of carriers that cannot be assigned to any business unit in B_h and any lane in L_g, and let I_{hg}^{-} indicate a set of carriers that cannot be assigned to any business unit in B_h and any lane in L_g. Thus, the preference constraint can be formulated as:

$$\sum_{b \in B_h} \sum_{k: J_{ik} \cap L_g \neq \varnothing} y_{ibk} \geq 1, \forall i \in I_{hg}^{+}, B_h \in BG, L_g \in \mathrm{LG}, \tag{13.13}$$

$$\sum_{b \in B_h} \sum_{k: J_{ik} \cap L_g \neq \varnothing} y_{ibk} = 0, \forall i \in I_{hg}^{-}, B_h \in BG, L_g \in LG. \tag{13.14}$$

Another type of shipper's constraint is called the fairness constraint, which restricts the assignment of carriers so as to be fair to different business units. This is motivated by current practice, where different business units may all request to be assigned to carriers that quote the lowest shipping price, but such carriers may only provide limited capacities. As a result, the available capacities of these low-price carriers have to be allocated fairly to the various business units (Lim et al. 2012). One way to reflect such a fairness concern is to impose a constraint such that, for each lane, the gap between the actual total shipping cost to a business unit and its minimum possible shipping cost shall not exceed a given percentage, denoted by $\eta\%$. As a result, we can formulate the fairness constraint as follows:

$$\sum_{i\in I}\sum_{t=1}^{T}[(p_{ijt}-s_{jt})x_{ibjt}+s_{jt}d_{bjt}] \le (1+\eta\%)\sum_{t=1}^{T}[(p_{i'jt}-s_{jt})x_{i'bjt}+s_{jt}d_{bjt}]$$
$$+M(1-y_{i'k}), \forall 1\le k\le |J_{i'}|, j\in J_{i'k}, b\in B, i'\in I, \tag{13.15}$$

where the left hand side indicates the actual shipping cost for business unit *b* and its cargo on lane *j*, the right hand side indicates its shipping cost for cargo on lane *j* if carrier i' is assigned, and *M* is a sufficiently large constant.

13.2.3 Generalization

By generalizing the various side constraints presented earlier, we can obtain a general mathematical programming model (General MP) for the TSPP as follows:

$$\text{(General MP)} = \min \sum_{b\in B}\sum_{j\in J}\sum_{t=1}^{T}\left[\sum_{i\in I} p_{ijt}x_{ibjt} + s_{jt}\left(d_{bjt}-\sum_{i\in I}x_{ibjt}\right)\right]$$

$$\text{s.t.}(2)-(7).\ (\mathbf{x},\mathbf{y}) \in EDom \bigcap IDom,$$

where *EDom* indicates the domain restricted by certain side constraints raised by the carriers, and *IDom* indicates the domain restricted by certain side constraints raised by the shipper's internal business units.

Moreover, the (General MP) can be reformulated into a bi-level optimization model, which separates determinations of decision variables y_{ibk} and x_{ibjt}. In other words, letting $\mathbf{y} = \{y_{ibk}: i\in I, b\in B, 1\le k\le |J_i|\}$, $EDom(\mathbf{y}) = \{\mathbf{x} : (\mathbf{x},\mathbf{y}) \in EDom\}$, and $IDom(\mathbf{y}) = \{\mathbf{x} : (\mathbf{x},\mathbf{y}) \in IDom\}$, we have

$$(\textbf{General MP}) = \min\ (\text{General MP})(y),\ \text{s.t.}7$$

$$(\text{GeneralMP})(y) = \min(1), \text{s.t.}(2)-(15)\mathbf{x} \in EDom(\mathbf{y}) \bigcap IDom(\mathbf{y}).$$

We make the point that the general model given here can include considerations other than those mentioned in Sect. 2.2, such as carriers' capacities in each period, and other commitment constraints on lanes.

13.3 Problem Complexities and Tractability

In this section, we discuss the computational tractability of different variations of the TSPP. Since the (Generic IP) model is equivalent to the classical transportation problem, with carriers as supply points and with triples (*b*, *j*, *t*) for $b\in B$, $j\in J$,

$1 \le t \le T$ as demand points, it can be solved by a min-cost network flow algorithm in polynomial time.

For the problem with the MQC constraint, it is easy to see that when q_{ik} equals one, the MQC constraint can be relaxed, implying that the problem is equivalent to the (Generic IP) model and can thus be solved in polynomial time. Moreover, it is known that the problem is strongly NP-hard whenever the minimum quantity q_{ik} is greater than or equal to three (Lim et al. 2006). The proof is based on a reduction from the set cover problem, which is well known to be strongly NP-complete. However, it still remains an open question as to whether or not the problem is NP-hard when q_{ik} equals two. We now consider a new special case when only one lane and one business unit is taken into account, i.e., $|J| = |B| = 1$, and establish a new tractability result for the TSPP as follows:

Theorem 1 *Solving the (Generic IP) model with the MQC constraint (13.8) and with $|J| = |B| = 1$ is NP-hard in the strong sense. It is NP-hard but has a pseudo-polynomial time algorithm for any fixed T.*

Proof. Given $|J| = |B| = 1$, we can reformulate the problem as follows:

$$(\textbf{MQC1 IP}) = \min \sum_{t=1}^{T} \left[\sum_{i \in I} p_{it} x_{it} + s_t \left(d_t - \sum_{i \in I} x_{it} \right) \right] \tag{13.16}$$

$$\text{s.t.} \sum_{i \in I} x_{it} \le d_t, \forall 1 \le t \le T, \tag{13.17}$$

$$q_i y_i \le \sum_{i \in I} x_{it} \le c_i y_i, \forall i \in I, \tag{13.18}$$

$$x_{it} \in Z_+, \forall i \in I, 1 \le t \le T, \tag{13.19}$$

$$y_i \in \{0,1\}, \forall i \in I, \tag{13.20}$$

where $d_t := d_{11t}$ indicates the demand for period t, x_{it} indicates the volume of cargo assigned to carrier i for period t, and y_i indicates whether or not carrier i is selected.

The strong NP-hardness of model (MQC1 IP) can be shown by a reduction from the following unary NP-complete problem:

Cover By 3-Sets (X3C) (Garey and Johnson 1983): Given a set $X = \{1, \ldots, 3k\}$ and a collection $C = \{C_1, \ldots, C_m\}$ with each member $C_i \subseteq X$ and $|C_i| = 3$ for $i = 1, \ldots, m$, does C contain an exact cover for X, i.e. a sub-collection $C' \subset C$ such that every element of X occurs in exactly one member of C'?

For any arbitrary instance of *X3C*, consider the instance of model (MQC1 IP) where $I = \{i: 1 \le i \le m\}$, $T = |X|$, $d_t = 1$ for $1 \le t \le T$, $q_i = 3$ for $i \in I$, $c_i \ge 3$ for $i \in I$, $p_{it} = 0$ for $i \in I$ and for $t \in C_i$, $p_{it} = \infty$ for $i \in I$ and for $t \notin C_i$, and $s_t = \infty$ for $1 \le t \le T$. We can show as follows that the (MQC1 IP) has a minimum total cost of zero if and only if the *X3C* has an exact cover.

On one hand, if the *X3C* has an exact cover, then we can set $y_i = 1$ if $C_i \in C'$ and $y_i = 0$ otherwise, and set $x_{it} = 1$ if $t \in C_i$ and $C_i \in C'$, and $x_{it} = 0$ otherwise. It can be seen that this leads to a feasible solution to (MQC1 IP) of a total cost equal to zero.

On the other hand, if (MQC1 IP) has a feasible solution of zero cost, let C' include C_i if and only if $y_i = 1$. Since $d_t = 1$, $q_i = 3$, and the total cost equals zero, it can be seen that each element $t \in X$ is covered by C' exactly once, which implies that C' is an exact cover. This completes the proof of the strong NP-hardness of (MQC1 IP).

Next, consider the case when T is fixed, and its NP-hardness can be shown by a reduction from the following NP-complete problem:

Partition (Garey and Johnson 1983): Given a set X, size $s(a)$ for $a \in X$, positive integer S, does X have a subset X' such that $\sum_{a \in X'} s(a) = S$?

For any arbitrary instance of *Partition*, consider the instance of model (MQC1 IP) where $I = X, q_a = c_a = s(a)$ for $a \in X, d_1 = S$, $p_{a1=}0, s_1 = \infty$, and $d_t = p_{at} = s_t = 0$ for $2 \leq t \leq T$. If this instance has a feasible solution of zero cost, then letting X' include a with $y_a = 1$ leads to $\sum_{a \in X'} s(a) = d_1 = S$. On the other hand, if there exists $X' \subseteq X$ with $\sum_{a \in X'} s(a) = S = d_1$, then setting $y_a = 1$ only for $a \in X'$ and setting $x_{a1} = s(a)$ and $x_{at} = 0$ for $2 \leq t \leq T$ lead to a feasible solution of zero cost. Thus, the case with $T = 1$ is NP-hard.

Now, for any fixed T, we can use the following dynamic programming algorithm to solve (MQC1 IP). Define $f\,(i, Q_1, \ldots, Q_T)$ as the minimum total cost to assign the cargo of Q_t for time $t = 1, 2, \ldots, T$ to carriers $1, 2, \ldots, i$, such that the capacity and MQC constraints for carriers $1, 2, \ldots, i$ are satisfied. For this, we can establish recurrence equations as follows:

$$f(i, Q_1, \ldots, Q_T) = \min\,\{f(i-1, Q_1, \ldots, Q_T), g(i, Q_1, \ldots, Q_T)\} \tag{13.21}$$

$$g(i, Q_1, \ldots, Q_T) = \min_{\mathbf{x}} f(i-1, Q_1 - x_{i1}, \ldots, Q_T - x_{iT}) + \sum_{t=1}^{T} p_{it} x_{it}$$

$$\text{s.t. } q_i \leq \sum_{t=1}^{T} x_{it} \leq c_i \tag{13.22}$$

where $f\,(0, 0, \ldots, 0) = 0$. Hence, the optimal objective value equals

$$\min_{Q_1,\ldots,Q_T} f(n, Q_1, \ldots, Q_T) + \sum_{t=1}^{T} s_t(d_t - Q_T) \tag{13.23}$$

It can be seen that the total time complexity of the above dynamic programming algorithm is $O(n(\max\{c_i : i \in I\})^T + (\max\{d_t : 1 \leq t \leq T\})^T)$, which is pseudo-polynomial when T is fixed. □

For the problem with the MARC constraint, it is known that the problem is strongly NP-hard for any fixed $\xi \geq 0$ even when only a single lane is considered (Lim et al. 2008a). The proof is also based on a reduction from the set cover problem. When $T = 1$, it can be seen that the problem with the MARC constraint is equivalent to the (Generic IP) model, and can thus be solved in polynomial time. Moreover, when only a single carrier is considered, i.e., $|I| = 1$, the problem can be reformulated as a min-cost network flow problem, which can also be solved in polynomial time. This special case also has a simple greedy algorithm (Lim et al. 2008a), which can guarantee a polynomial running time if the total demand is polynomially bounded.

For the problem with the constraints on the number of selected carriers, it can be transformed to the classical k-median problem, and thus it is strongly NP-hard (Arya et al. 2001; Bozkaya et al. 2002; de Farias 2001; Hochbaum 1982; Lorena and Senne 2004; Rolland et al. 1997; Senne et al. 2005), where carriers correspond to candidate locations of facilities, and triples (b, j, t) indicate the locations of demand points, even when $|J| = 1$ or $T = 1$. When the maximum number of selected carriers is fixed, the problem can be solved in polynomial time, as one can enumerate all the possible combinations of carriers, and solve model (Generic IP) on only selected carriers to obtain optimal cargo allocations. In this case, even if the MQC constraint is included, the problem can still be solved in polynomial time. Now, consider a new special case where both $|J|$ and T equal to one, for which we can derive a new tractability result as follows.

Theorem 2 *It is strongly NP-hard to solve the (Generic IP) model with constraints (13.10)—(13.12) on the number of carriers and with $|J| = T = 1$, but it has a polynomial time algorithm when $|B|$ is fixed.*

Proof. For the (Generic IP) model with constraints (13.10)—(13.12) on the number of carriers and with $|J| = T = 1$, we can reformulate it as follows.

$$(\textbf{NUM1 IP}) = \min \sum_{\mathrm{b}\in\mathrm{B}} \left[\sum_{\mathrm{i}\in\mathrm{I}} \mathrm{p_i x_{ib}} + \mathrm{s} \left(\mathrm{d_b} - \sum_{\mathrm{i}\in\mathrm{I}} \mathrm{x_{ib}} \right) \right] \tag{13.24}$$

$$\text{s.t.} \sum_{i\in I} x_{ib} \le d_b, \forall b \in B, \tag{13.25}$$

$$\sum_{b\in B} x_{ib} \le c_i y_i, \forall i \in I, \tag{13.26}$$

$$\sum_{b\in B} y_{ib} \le |B| y_i, \forall i \in I, \tag{13.27}$$

$$\underline{n}_b \le \sum_{i\in I} y_{ib} \le \bar{n}_b, \forall b \in B, \tag{13.28}$$

$$x_{ib} \in Z_+, \forall i \in I, b \in B, \tag{13.29}$$

$$y_i \in \{0,1\}, \forall i \in I, \tag{13.30}$$

$$y_{ib} \in \{0,1\}, \forall i \in I, b \in B, \tag{13.31}$$

where x_{ib} indicates the total volume of cargo of business unit b assigned to carrier i, and y_{ib} indicates whether or not carrier i is selected to serve business unit b. We now show its NP-hardness as follows by a reduction from the following NP-complete problem:

3-Partition (Garey and Johnson 1983): Given a set X of $3k$ elements, a bound S, and a size $s(a)$ for $a \in X$ such that $S/4 < s(a) < S/2$ and $\sum_{a \in X} s(a) = kS$, can X be partitioned into $A_1, \ldots, A_k$ such that for $1 \leq r \leq k$, $\sum_{a \in A_r} s(a) = S$?

For any arbitrary instance of *3-Partition*, consider the instance of model (NUM1 IP) where $|J| = T = 1$, $I = X$, $|B| = k$, $c_a = s(a)$ for $a \in X$, $d_b = S$ for $b \in B$, $\underline{n}_b = \bar{n}_b = 3$ for $b \in B$, $p_i = 0$ for $i \in I$, and $s = \infty$. If this instance has a feasible solution of zero cost, then letting $A_b := \{i \in I : y_{ib} = 1\}$ include those carriers i with $y_{ib} = 1$ for $b \in B$. Since $\underline{n}_b = \bar{n}_b = 3$ for $b \in B$, $|B| = k$, and $\sum_{a \in X} s(a) = kS$, it can be seen that no carrier can serve more than one business unit, which implies that $A_1, \ldots, A_k$ is feasible to the *3-Partition* problem. On the other hand, if the *3-Partition* problem has a feasible partition $A_1, \ldots, A_k$, then by setting $y_{ib} = 1$ and $x_{ib} = s(i)$ for $i \in A_b$ and $b \in B$, we obtain a feasible solution to (NUM1 IP) of zero cost. Thus, the case is strongly NP-hard.

Next, consider the case when $|B|$ is fixed. Consider any y_{ib} for $i \in I$ and $b \in B$ that satisfy the constraints on the number of selected carriers. The total number of such possible combinations is polynomially bounded, since $|B|$ is fixed. We can now extend the (Generic IP) model to obtain a min-cost network flow model as follows, so as to determine optimal allocations for the cargo:

$$\min \sum_{b \in B} \left[\sum_{i \in I} p_i x_{ib} + s \left(d_b - \sum_{i \in I} x_{ib} \right) \right] \tag{13.32}$$

$$\text{s.t.} \sum_{i \in I} x_{ib} \leq d_b, \forall b \in B, \tag{13.33}$$

$$\sum_{b \in B} x_{ib} \leq c_i y_i, \forall i \in I, \tag{13.34}$$

$$x_{ib} \in Z_+, \forall i \in I, b \in B. \tag{13.35}$$

Hence, this special case can be solved in polynomial time.

For the problem with the preference constraints, this can still be transformed to a min-cost network flow problem, and thus can be solved in polynomial time. However, in the literature, it is often imposed simultaneously with the constraints on the number of selected carriers (Lim et al. 2012), which in general is strongly NP-hard.

For the problem with the fairness constraints, its tractability is unknown in the literature. Lim et al. (2012) studied a generalized problem that has taken into account fairness constraints, and carrier number constraints, as well as preference constraints, and they showed that it is strongly NP even to find a feasible solution to this problem, and also that it has a polynomial time algorithm that can obtain a feasible solution when all shipping prices are identical.

From the tractability results above, we know that the generalized problem (General MP) is strongly NP-hard. However, for those special cases of (General MP), where only two types of side constraints are taken into account, few results showing their tractability are known in the literature.

13.4 Problem Relaxations

In this section, we discuss various relaxations of the models proposed in Sect. 2 for the TSPP. Although optimal solutions to these relaxations may not be feasible for the original models, they can be used to estimate the optimal objective values of the original models, as well as to construct feasible solutions for the original models.

Consider the problem with the MQC constraints, where $|B|$ is assumed to be one for ease of presentation. We can reformulate this problem as the following integer programming model:

$$(\textbf{MQC IP}) = \min \sum_{j \in J} \sum_{t=1}^{T} \left[\sum_{i \in I} p_{ijt} x_{ijt} + s_{jt} \left(d_{jt} - \sum_{i \in I} x_{ijt} \right) \right] \tag{13.36}$$

$$\text{s.t.} \sum_{i \in I} x_{ijt} \le d_{jt}, \forall j \in J, 1 \le t \le T, \tag{13.37}$$

$$\sum_{j \in J_{ik}} \sum_{t=1}^{T} x_{ijt} \le c_{ik} y_{ik}, \forall i \in I, 1 \le k \le |J_i|, \tag{13.38}$$

$$q_{ik} y_{ik} \le \sum_{j \in J_{ik}} \sum_{t=1}^{T} x_{ijt}, \forall i \in I, 1 \le k \le |J_i|, \tag{13.39}$$

$$x_{ijt} \in Z_+, \forall i \in I, j \in J, 1 \le t \le T, \tag{13.40}$$

$$y_{ik} \in \{0,1\}, \forall i \in I, 1 \le k \le |J_i|, \tag{13.41}$$

where x_{ijt} indicates the total number of containers assigned to carrier i for cargo of lane j and period t. For this problem, it has a linear programming relaxation where **x** and **y** are relaxed to take fractional values:

$$z_{MQC}^{LP} = \min \sum_{j \in J} \sum_{t=1}^{T} \left[\sum_{i \in I} p_{ijt} x_{ijt} + s_{jt} \left(d_{jt} - \sum_{i \in I} x_{ijt} \right) \right]$$

$$\text{s.t.}(37) - (39), \tag{13.42}$$

$$x_{ijt} \in R_+, \forall i \in I, j \in J, 1 \le t \le T, \tag{13.43}$$

$$0 \le y_{ik} \le 1, \forall i \in I, 1 \le k \le |J_i|. \tag{13.44}$$

The above linear programming relaxation model can be strengthened by the following valid constraints of (MQC IP), which are extended from Lim et al. (2006)

and Lim et al. (2008b). First, by (13.37) and (13.38), we have

$$x_{ijt} \leq d_{jt} y_{ik}, \forall i \in I, j \in J_{ik}, 1 \leq k \leq |J_i|, 1 \leq t \leq T.$$

Next, suppose that pairs (*i, k*) are sorted on a non-decreasing order of q_{ik}, and let $K_{\max}$ indicate the smallest position of the order, such that the sum of q_{ik} for the first $K_{\max}$ pairs of (*i, k*) exceeds D, where $D := \sum_{j \in J} \sum_{t=1}^{T} d_{jt}$. By (13.37) and (13.39), we have

$$\sum_{i \in I} \sum_{k=1}^{|J_i|} y_{ik} \leq K_{\max}.$$

Moreover, by extending the arguments from Lim et al. (2006, 2008b), we can show that the above two valid constraints both define facets of the convex hull of the integer programming model (MQC IP) under some mild conditions.

In addition to the linear programming relaxation, we can obtain a Lagrangian relaxation of model (MQC IP) by dualizing the demand constraint (13.37). Let $\mu_{jt} \geq 0$ indicate the Lagrangian multiplier, associated with constraint (13.37). Let $z^{LR1}_{MQC}(\mu)$ denote the optimal objective value for the following problem:

$$z^{LR1}_{MQC}(\mu) = \min \sum_{j \in J} \sum_{t=1}^{T} \sum_{i \in I} (p_{ijt} - s_{jt} + \mu_{jt}) x_{ijt} + \sum_{j \in J} \sum_{t=1}^{T} (s_{jt} - \mu_{jt}) d_{jt}$$

$$\text{s.t.}(38) - (41).$$

The above $z^{LR1}_{MQC}(\mu)$ can be decomposed by carriers $i \in I$ and lane groups in J_i into $\sum_{i \in I} |J_i|$ sub-problems, with each corresponding to a continuous knapsack problem, and thus it can be solved in polynomial time. Thus, one can apply a subgradient algorithm to maximize $z^{LR1}_{MQC}(\mu)$ over multipliers μ, so as to obtain a Lagrangian relaxation lower bound, denoted by z^{LR1}_{MQC}, for model (MQC IP).

Furthermore, by dualizing the capacity constraint (13.38) and the MQC constraints (13.39), we can derive a new Lagrangian relaxation of model (MQC IP) as follows, where π_{ik} and γ_{ik} indicate the associated Lagrangian multipliers.

$$z^{LR2}_{MQC}(\pi, \gamma) = \min \sum_{j \in J} \sum_{t=1}^{T} \sum_{i \in I} (p_{ijt} - s_{jt} + \pi_{ik} - \gamma_{ik}) x_{ijt} + \sum_{i \in I} (q_{ik} \gamma_{ik} - c_{ik} \pi_{ik}) y_{ik}$$

$$+ \sum_{j \in J} \sum_{t=1}^{T} s_{jt} d_{jt} \quad \text{s.t. } (37), (40), (41).$$

Based on the signs of $(q_{ik} \gamma_{ik} - c_{ik} \pi_{ik})$, we can determine values of $\mathbf{y}$, and then the remaining problems on $\mathbf{x}$ can be decomposed by (*j, t*) for $j \in J$ and $1 \leq t \leq T$ into $|J|T$ sub-problems, with each sub-problem equivalent to a continuous knapsack problem, and thus it can be solved in polynomial time. By applying a subgradient algorithm, we can thus maximize $z^{LR2}_{MQC}(\pi, \gamma)$ to obtain a lower bound on the optimal objective value of (MQC IP), denoted by z^{LR2}_{MQC}.

We can also derive an LP relaxation of the Dantzig-Wolfe reformulation of model (MQC IP) as follows. Let $\mathbf{x}^{ih}$for $h = 1,2,\ldots,H$ indicate all the feasible cargo allocations for carrier i which satisfy the capacity constraint (13.38) and the MQC constraint (13.39), and we denote the cost of each $\mathbf{x}^{ih}$ by $c^{ih} := \sum_{j\in J}\sum_{t=1}^{T}(p_{ijt} - s_{jt})x_{jt}^{ih}$. Let λ_{ih} indicate a binary variable that equals 1 if and only if cargo allocation $\mathbf{x}^{ih}$is assigned to carrier i. Thus, model (MQC IP) can be reformulated as follows:

$$\min \sum_{i\in I}\sum_{h=1}^{H} c^{ih}\lambda_{ih} \tag{13.45}$$

$$\text{s.t.} \sum_{h=1}^{H}\lambda_{ih} = 1, \forall i \in I, \tag{13.46}$$

$$\sum_{i\in I}\sum_{h=1}^{H}\lambda_{ih}x_{jt}^{ih} \le d_{jt}, \forall j \in J, 1 \le t \le T, \tag{13.47}$$

$$\lambda_{ih} \in \{0,1\}, \forall i \in I, 1 \le h \le H. \tag{13.48}$$

The linear programming relaxation of the above model, denoted by z_{MQC}^{LPDW}, can be solved by column generation, for which the pricing problem is equivalent to a continuous knapsack problem, and thus can be solved in polynomial time.

Proposition 1 below reveals that the four relaxations above are equally tight:

Proposition 1. $z_{MQC}^{LP} = z_{MQC}^{LR1} = z_{MQC}^{LPDW} = z_{MQC}^{LR2}$

Proof. The convex hull of feasible solutions to $z_{MQC}^{LR2}(\pi,\gamma)$ is the same as the convex hull of its linear programming relaxation, which implies that $z_{MQC}^{LP} = z_{MQC}^{LR2}$. The convex hull of feasible cargo allocations for each carrier i is the same as the convex hull of its linear programming relaxation, which implies that $z_{MQC}^{LP} = z_{MQC}^{LPDW}$. Finally, the convex hull of feasible solutions to $z_{MQC}^{LR1}(\mu)$ is the same as the convex hull of its linear programming relaxation, which implies that $z_{MQC}^{LP} = z_{MQC}^{LR1}$. Hence, the proposition is proved. □

The proposition above implies that the four relaxations mentioned above are equally tight. To derive tighter relaxations, one needs to introduce more valid constraints. For example, by (13.37) and (13.39) we can obtain the following valid constraint

$$\sum_{i\in I}\sum_{k=1}^{|J_i|} q_{ik}y_{ik} \le \sum_{j\in J}\sum_{t=1}^{T} d_{jt}. \tag{13.49}$$

By extending the argument in Lim et al. (2008b), it can be shown that by including (13.49) in $z_{MQC}^{LR1}(\mu)$, one can obtain a stronger Lagrangian relaxation, which can be transformed to a multiple-dimensional knapsack problem and solved by a dynamic programming algorithm. Thus, the resulting lower bound on the optimal objective value to (MQC IP) is tighter than z_{MQC}^{LP}.

Next, consider the problem with the MARC constraint, which can be formulated as follows:

$$(\textbf{MARC IP}) = \max \sum_{j \in J} \sum_{t=1}^{T} \sum_{i \in I} (s_{jt} - p_{ijt}) x_{ijt} \tag{13.50}$$

$$\text{s.t.} \sum_{i \in I} x_{itj} \leq d_{jt}, \forall j \in J, 1 \leq t \leq T, \tag{13.51}$$

$$\sum_{j \in J_{ik}} \sum_{t=1}^{T} x_{ijt} \leq c_{ik}, \forall i \in I, 1 \leq k \leq |J_i|, \tag{13.52}$$

$$\sum_{j \in J_{ik}} x_{ijt} \leq \alpha_{ikj} \left(\sum_{j \in J_{ik}} \sum_{t=1}^{T} x_{ijt} \right) + \xi, \forall i \in I, 1 \leq k \leq |J_i|, 1 \leq t \leq T, \tag{13.53}$$

$$x_{ijt} \in Z_+, \forall i \in I, j \in J, 1 \leq t \leq T. \tag{13.54}$$

It also has a linear programming relaxation by relaxing **x** to take fractional values:

$$(\textbf{MARC IP}) = \max \sum_{j \in J} \sum_{t=1}^{T} \sum_{i \in I} (s_{jt} - p_{ijt}) x_{ijt} \tag{13.55}$$

$$\text{s.t. } (51) - (53)$$

$$x_{ijt} \geq 0, \forall i \in I, j \in J, 1 \leq t \leq T. \tag{13.56}$$

The above linear programming relaxation model can also be strengthened by introducing some valid constraints of the model (MARC IP). For example, from (13.52) and (13.53), we have $\sum_{j \in J_{ik}} x_{ijt} \leq \alpha_{ikj} c_{ik} + \xi, \forall i \in I, 1 \leq k \leq |J_i|, 1 \leq t \leq T$. Thus,

$$\sum\nolimits_{j \in J_{ik}} x_{ijt} \leq \left\lfloor \alpha_{ikj} c_{ik} \right\rfloor + \xi, \forall i \in I, 1 \leq k \leq |J_i|, 1 \leq t \leq T. \tag{13.57}$$

We can establish the following theorem to show that under some mild conditions, the valid constraint (13.57) defines a facet of model (MARC IP), and thus is necessary.

Theorem 3. *If* $\left\lfloor \alpha_{ikj} c_{ik} \right\rfloor + \xi < \sum_{j \in J_{ik}} d_{jt}$ *and* $\sum_{t=1}^{T} \left(\left\lfloor \alpha_{ikj} c_{ik} \right\rfloor + \xi \right) < c_{ik}$*, then (*13.57) defines a facet of model (MARC IP).

Proof. To prove this theorem, we only need to show that if all feasible solutions to model (MARC IP) that satisfy (13.57) for some $i \in I$ and t with $1 \leq t \leq T$ at equality also satisfy

$$\sum_{i \in I} \sum_{j \in J} \sum_{t=1}^{T} a_{ijt} x_{ijt} \leq \theta, \tag{13.58}$$

at equal, then (13.58) is equivalent to (13.57).

First, since $\lfloor \alpha_{ikj} c_{ik} \rfloor + \xi < \sum_{j \in J_{ik}} d_{jt}$ and $\sum_{t=1}^{T} \left(\lfloor \alpha_{ikj} c_{ik} \rfloor + \xi \right) < c_{ik}$, there exists a feasible $\mathbf{x^1}$such that (13.57) is satisfied at equality for i, k and t, and such that constraints (13.51), (13.52) and (13.53) are all satisfied but not at equality. For any with i', j', t' with $i' \neq i$, or $j' \in J_{ik}$, or $t' \neq t$, consider $\mathbf{x^2}$, which is equal to $\mathbf{x^1}$except that

$$x^2_{i'j't'} = x^1_{i'j't'} + \varepsilon. \tag{13.59}$$

Thus, it can be seen that there exists $\varepsilon > 0$, such that $\mathbf{x^2}$ is feasible to model (MARC IP) and satisfies (13.57) for i, k, t at equal. Substituting $\mathbf{x^1}$ and $\mathbf{x^2}$ into (13.58) and subtracting one from the other results in $a_{i'j't'} = 0$.

Next, for any j and j' in J_{ik}, consider $\mathbf{x^3}$, which is equal to $\mathbf{x^1}$ except that

$$x^3_{ijt} = x^1_{ijt} + \varepsilon. \tag{13.60}$$

$$x^3_{ij't} = x^1_{ij't} - \varepsilon. \tag{13.61}$$

It can be seen that there exists $\varepsilon > 0$, such that $\mathbf{x^3}$ is feasible to model (MARC IP) and satisfies (13.57) for i, k, t at equal. Substituting $\mathbf{x^3}$and $\mathbf{x^1}$into (13.58) and subtracting one from the other results in $a_{ijt} = a_{ij't}$.

Thus, we can assume $a_{ijt} = a$ for $j \in J_{ik}$ and (13.58) can be represented as:

$$\sum_{j \in J_{ik}} a x_{ijt} \leq \theta. \tag{13.62}$$

Thus, since $\sum_{j \in J_{ik}} x^1_{ijt} = \lfloor \alpha_{ikj} c_{ik} \rfloor + \varepsilon$, we obtain that (13.58) is equivalent to (13.57).

Next, by dualizing the demand constraint (13.51), we can obtain a Lagrangian relaxation of model (MARC IP). Let $\mu_{jt} \geq 0$ indicate the Lagrangian multiplier, associated with constraint (13.51). Define $z^{LR}_{MARC}(\mu)$ as follows:

$$z^{LR}_{MARC}(\mu) = \max \sum_{j \in J} \sum_{t=1}^{T} \sum_{i \in I} (s_{jt} - p_{ijt} - \mu_{jt}) x_{ijt}$$

$$\text{s.t. } (52) - (54). \tag{13.63}$$

It can be seen that $z^{LR}_{MARC}(\mu)$ can be decomposed by carriers and lane groups into $\sum_{i \in I} |J_i|$ sub-problems, with each being equivalent to a single-carrier problem, and thus it can be solved in polynomial time (Lim et al. 2008a).

Moreover, we can also derive an LP relaxation of the Dantzig-Wolfe reformulation of model (MARC IP) as follows. Let $\mathbf{x}^{ih}$for $h = 1, 2, \ldots, H$ indicate all the feasible cargo allocations for carrier i, which satisfy the capacity constraint (13.52) and the MARC constraint (13.53), and we denote each saving by $c^{ih} := \sum_{j \in J} \sum_{t=1}^{T} (s_{jt} - p_{ijt}) x^{ih}_{jt}$. Let λ_{ih} indicate a binary variable that equals 1 if and only if cargo allocation $\mathbf{x}^{ih}$ is assigned to carrier i. Thus, model (MQC IP) can be reformulated as follows:

$$\max \sum_{i \in I} \sum_{h=1}^{H} c^{ih} \lambda_{ih} \tag{13.64}$$

$$\text{s.t.} \sum_{h=1}^{H} \lambda_{ih} = 1, \forall i \in I, \tag{13.65}$$

$$\sum_{i \in I} \sum_{h=1}^{H} \lambda_{ih} x_{jt}^{ih} \leq d_{jt}, \forall j \in J, 1 \leq t \leq T, \tag{13.66}$$

$$\lambda_{ih} \in \{0,1\}, \forall i \in I, 1 \leq h \leq H. \tag{13.67}$$

The linear programming relaxation of the above model, denoted by z_{MARC}^{LPDW}, can be solved by column generation, for which the pricing problem is equivalent to a single carrier problem, and thus can be solved in polynomial time.

The following proposition reveals the tightness of the three relaxations above:

Proposition 2. $z_{MARC}^{LP} \leq z_{MARC}^{LR} = z_{MARC}^{LPDW}$.

Proof. Noticing that both the Lagrangian dual and the pricing problem are equivalent to a single-carrier problem, we can obtain that $z_{MARC}^{LR} = z_{MARC}^{LPDW}$. Moreover, the convex hull of feasible solutions to the single-carrier problem is a superset of the convex hull of its linear programming relaxation, which implies that $z_{MARC}^{LP} \leq z_{MARC}^{LR} = z_{MARC}^{LPDW}$, completing the proof. □

Similarly, by dualizing the demand constraint, we can further derive Lagrangian relaxations, as well as the LP relaxation of the Dantzig-Wolfe reformulation, for problems with constraints on the number and preference of selected carriers, as well as on the fairness. Such relaxation techniques can also be applied to the (General MP) having *EDom* and *IDom* containing all the various constraints. It is of interest to investigate the tightness of these new relaxations as well as how to strengthen them. For example, for the problem with constraints on the number of selected carriers, we can derive a valid constraint directly from (13.11) as follows:

$$\sum_{b \in B_h} \sum_{k : J_{ik} \bigcap L_g \neq \varnothing} y_{ibk} \leq |B_h| \left| \left\{ k : J_{ik} \bigcap L_g \neq \varnothing \right\} \right| z_{ihg}, \forall i \in I, B_h \in BG, L_g \in LG. \tag{13.68}$$

For the problem with the fairness constraint, we can derive a valid constraint from (13.15) by replacing M with $\sum_{i \in I} \sum_{t=1}^{T} s_{jt} d_{bjt}$, since the total cost for shipping the cargo of each lane should not exceed the cost of shipping them all using the spot-market price.

13.5 Exact Algorithms

Since the TSPP, as well as most of its special cases, are often computationally intractable, it is of great interests to develop algorithms that can produce optimal solutions to relatively small sized instances of the TSPP in affordable running time.

To solve the model (General MP) for cases where (General MP)($\mathbf{y}$) can be solved to optimum efficiently, one can follow a branch-and-bound algorithm to search an optimal value of $\mathbf{y}$ that minimizes (General MP)($\mathbf{y}$). As it goes down the search tree, the branch-and-bound algorithm determines values of y_{ibk} one by one, and keeps the current best feasible solution denoted by $(\mathbf{x}^*, \mathbf{y}^*)$. At each node of the search tree, let Π_0 indicate the set of triples (i, b, k) with $y_{ibk} = 0$ and let Π_1 indicate the set of triples (i, b, k) with $y_{ibk} = 1$. Let $\Pi := \Pi_0 \cup \Pi_1$ indicate the set of triples (i, b, k) for determined y_{ibk}, and $\bar{\Pi}$ indicate the set of triples (i, b, k) for un-determined y_{ibk}. Hence (Π_0, Π_1) can be used to represent a partial solution.

Before assigning values to those un-determined y_{ibk}, the algorithm computes a lower bound on the best possible objective value that can be obtained by completing the current partial solution (Π_0, Π_1). This can be achieved by solving relaxations of the models on the remaining problems, such as the linear programming relaxation and the Lagrangian relaxation described in Sect. 4. If the obtained lower bound is not less than the objective value of the current best solution $(\mathbf{x}^*, \mathbf{y}^*)$, then the node can be pruned. Otherwise, the algorithm will select an un-determined y_{ibk} for $(i, b, k) \in \bar{\Pi}$, and assign y_{ibk} either 0 or 1, so as to generate two new nodes of the search tree. Given the partial solution of each new node, we can construct feasible solutions by various heuristics, which will be introduced later in Sect. 6. If the obtained solution has a better objective value than the current best solution $(\mathbf{x}^*, \mathbf{y}^*)$, the algorithm will update $(\mathbf{x}^*, \mathbf{y}^*)$. This branch-and-bound algorithm can be summarized in Algorithm 1.

Algorithm 1 The Branch-and-Bound Algorithm

```
Let NList represent the list of nodes of the search tree to be expanded,
  and set the initial value of NList to be the set that contains only the root node.
  Let UB indicate the objective value of the current best feasible solution
  $(\mathbf{x}^*, \mathbf{y}^*)$, and set the initial value of UB to be $\infty$.
while NList is not empty do
    Choose a node p in NList, exclude p from NList, and consider its
      associated partial solution $(\Pi_0, \Pi_1)$.
    Compute a lower bound LB on the best possible objective value that
      can be obtained by completing $(\Pi_0, \Pi_1)$.
    if $LB < UB$ then
        Let $\Pi := \Pi_0 \cup \Pi_1$
        for each $(i, b, k) \in \bar{\Pi}$ do
            for $v \in \{0, 1\}$ do
                Construct a new feasible solution $(\hat{\Pi}_0, \hat{\Pi}_1)$, where
                  $\hat{\Pi}_v := \Pi_v \cup \{(i,b,k)\}$ and $\hat{\Pi}_{1-v} := \Pi_{1-v}$.
                Construct a feasible solution from the new partial solution
                  $(\hat{\Pi}_0, \hat{\Pi}_1)$. If the feasible solution has a smaller objective value
                  than UB, then update UB and $(\mathbf{x}^*, \mathbf{y}^*)$.
                Add to NList the new node p that is associated with $(\hat{\Pi}_0, \hat{\Pi}_1)$.
            end for
        end for
    end if
end while
Return the current best solution $(\mathbf{x}^*, \mathbf{y}^*)$.
```

To enhance the branch and bound algorithm, we can strengthen the relaxations of the problem so as to obtain better lower bounds. For example, from the linear programming relaxation, we can obtain a fractional solution, which may not be feasible to the original model (General MP). In this case, an intuitive way to strengthen the relaxation is to introduce new valid constraints that can exclude the fractional solutions. This approach is usually referred to as a branch-and-cut algorithm (Lim et al. 2006).

The above branch-and-bound algorithm has been applied in the literature to solving the problem (MQC IP) (Lim et al. 2006). Given **y**, the model (MQC IP) can be reformulated as follows:

$$(\textbf{MQC IP})(y) = \min \sum_{j\in J}\sum_{t=1}^{T}\left[\sum_{i\in I} p_{ijt}x_{ijt} + s_{jt}\left(d_{jt} - \sum_{i\in I} x_{ijt}\right)\right] \tag{13.69}$$

$$\text{s.t.} \sum_{i\in I} x_{ijt} \le d_{jt}, \forall j \in J, 1 \le t \le T, \tag{13.70}$$

$$q_{ik}y_{ik} \le \sum_{j\in J_{ik}}\sum_{t=1}^{T} x_{ijt} \le c_{ik}y_{ik}, \forall i \in I, 1 \le k \le |J_i|, \tag{13.71}$$

$$x_{ijt} \in Z_+, \forall i \in I, j \in J, 1 \le t \le T, \tag{13.72}$$

which is equivalent to a min-cost network flow model, and thus can be solved efficiently. Moreover, when given a partial solution (Π_0, Π_1), the remaining problem can be formulated as follows:

$$(\textbf{MQC IP})(\Pi_0, \Pi_1) = \min \sum_{t=1}^{T}\sum_{j\in J} s_{jt}d_{jt} + \sum_{t=1}^{T}\sum_{(i,k)\in\overline{\Pi_0}}\sum_{j\in J_{ik}} (p_{ijt} - s_{jt})x_{ijt} \tag{13.73}$$

$$\text{s.t.} \sum_{(i,k)\in\overline{\Pi_0}, j\in J_{ik}} x_{ijt} \le d_{jt}, \forall j \in J, 1 \le t \le T, \tag{13.74}$$

$$q_{ik} \le \sum_{j\in J_{ik}}\sum_{t=1}^{T} x_{ijt} \le c_{ik}, \forall (i,k) \in \Pi_1, \tag{13.75}$$

$$q_{ik}y_{ik} \le \sum_{j\in J_{ik}}\sum_{t=1}^{T} x_{ijt} \le c_{ik}y_{ik}, \forall (i,k) \in \overline{\Pi} \tag{13.76}$$

$$x_{ijt} \in Z_+, \forall (i,k) \in \overline{\Pi_0}, j \in J_{ik}, 1 \le t \le T, \tag{13.77}$$

$$y_{ik} \in \{0,1\}, \forall (i,k) \in \overline{\Pi}. \tag{13.78}$$

Thus, relaxations of (MQC IP)(Π_0, Π_1) can provide valid lower bounds for the branch and bound algorithm. This model for the remaining problem can be further strengthened by including the valid constraints presented in Sect. 4.

Furthermore, the framework in Algorithm 1 can also be applied to problems with constraints on the number as well as on the shipper's preference for selected carriers. This is because these constraints are only associated with **y**, and thus, given **y**, the problems with these constraints are equivalent to the classical transportation problem, and can be solved efficiently in polynomial time.

For (General MP)(**y**), searching for an exact optimal solution is more complicated, since it needs to explore possible values for both **x** and **y**. However, in some cases, we may still be able to reduce search space by introducing some auxiliary variables. Consider the problem with the maximum-to-average-ratio commitment constraint (MARC IP). Define $v_{ik} := \sum_{j \in J_{ik}} \sum_{t=1}^{T} x_{ijt}$ as new auxiliary variables that represent the total volume of cargo assigned to carrier i for lanes in J_{ik} and for all the T periods. Given **v**, the model (MARC IP) can be reformulated as follows:

$$(\textbf{MARC IP})(v) = \max \sum_{i \in I} \sum_{j \in J} \sum_{t=1}^{T} (s_{jt} - p_{ijt}) x_{ijt} \tag{13.79}$$

$$\text{s.t.} \sum_{i \in I} x_{itj} \le d_{jt}, \forall j \in J, 1 \le t \le T, \tag{13.80}$$

$$x_{ijt} \in Z_+, \forall i \in I, j \in J, 1 \le t \le T, \tag{13.81}$$

$$\sum_{j \in J_{ik}} x_{ijt} \le \alpha_{ikj} v_{ik} + \xi, \forall i \in I, 1 \le k \le |J_i|, 1 \le t \le T, \tag{13.82}$$

which can be transformed to a min-cost network flow problem, and thus can be solved efficiently. Therefore, to solve (MARC IP), we can develop a branch and bound algorithm to find optimal values of **v** so as to maximize (MARC IP)(**v**).

However, since v_{ik} are not binary variables, the branch and bound algorithm for (MARC IP) needs to narrow the value ranges of v_{ik} iteratively. Thus, a partial solution needs to be represented here by a vector of pairs $(\underline{v}_{ik}, \bar{v}_{ik})$, where $\underline{v}_{ik}$and $\bar{v}_{ik}$ indicate the lower and the upper bound of each v_{ik}, respectively. At each decision node associated with the partial solution $\underline{v}_{ik}, \bar{v}_{ik}) : i \in I, 1 \le k \le |J_i|\}$, the algorithm first computes the lower bound *LB* on the best possible objective value that can be obtained by completing $\{(\underline{v}_{ik}, \bar{v}_{ik}) : i \in I, 1 \le k \le |J_i|\}$. This can be achieved by solving a relaxation of the integer programming model for given $\{(\underline{v}_{ik}, \bar{v}_{ik}) : i \in I, 1 \le k \le |J_i|\}$. Next, the algorithm selects any (i, k) with $\underline{v}_{ik} < \bar{v}_{ik}$, and computes the midpoint p of $[\underline{v}_{ik}, \bar{v}_{ik}]$, so that two new nodes can be generated with the range of v_{ik} being $[\underline{v}_{ik}, p]$ and $[p + 1, \bar{v}_{ik}]$, respectively. For each new node, we can construct feasible solutions to update the objective value *UB* of the existing best feasible solution.

13.6 Heuristic Algorithms

To tackle the TSPP and its special cases that are computationally intractable, one solution approach is to develop heuristic algorithms that can produce feasible solutions close to the optimum in affordable running time. Recall that for a minimization (or maximization) problem, a heuristic algorithm is said to be a ρ-approximation algorithm if, for every instance of the problem, the algorithm has a polynomial running time and returns a feasible solution that has an objective value at most ρ times the minimum objective value (or at least $1/\rho$ times the maximum objective value). The value of ρ is referred to as the approximation ratio of the algorithm.

The constant-ratio approximation algorithms known in the literature on transportation service procurement are mainly for the special case where only MQC constraints are taken into account. One such algorithm follows a greedy approach (Lim et al. 2006). In this greedy algorithm, two operators, *selection*(i, k) and *assignment*(i, k), are defined for the construction of feasible solutions. For each unassigned pair of carrier $i \in I$ and $1 \leq k \leq |J_i|$, the operator *selection*(i, k) selects carrier i and lanes in J_{ik}, and assigns carrier i the cheapest q_{ik} units of unassigned cargo of lanes in J_{ik} so as to satisfy the minimum quantity commitment constraint. For each assigned carrier $i \in I$ and $1 \leq k \leq |J_i|$, the operator *assignment*(i, k) assigns carrier i the cheapest unassigned cargo of lanes in J_{ik} for delivering. Based on these two operators, the algorithm constructs a feasible solution to the problem iteratively, and during each iteration, it applies the operator with the minimum average cost, until all the cargo has been assigned. Here, the average cost of *selection*(i, k) is measured by $\sum_{(j,t)\in A} (p_{ijt} - s_{jt})/q_{ik}$, where A is the set of q_{ik} cargo of lanes in J_{ik} newly assigned to carrier i, and the average cost of *assignment*(i, k) is measured by $(p_{ijt} - s_{jt})$.

We summarize the above greedy algorithm in Algorithm 2, which extends the one in Lim et al. (2006), as constraints presented in this chapter are more general, with time periods being taken into account. It can be seen that after each iteration of Algorithm 2, at least one of the three following events must happen: (i) a carrier i and lanes in J_{ik} are newly selected; (ii) the capacity of a selected carrier i is fully assigned for lanes in J_{ik}; (iii) the demand of lane j and period t is fully satisfied. This implies that the total number of iterations is $O(\sum_{i\in I} |J_i| + |J|T)$. Since each iteration has a polynomial running time, we obtain that Algorithm 2 runs in polynomial time. Moreover, by following a similar argument as in Lim et al. (2006), it can be shown that Algorithm 2 has an approximation ratio of b if all the carriers have unlimited capacity, if their minimum quantities q_i all equal a constant b, if they have only one lane group, and if the shipping price $(p_{ijt} -_{jt} s)$ forms a metric.

Algorithm 2 Greedy Algorithm

```
Set the selected set, $\Pi_1$, equal to empty, and set assigned quantity $\hat{d}_{jt}$ equal to zero.
while NOT all cargo has been assigned, i.e. there exists $(j, t)$ such that $\hat{d}_{jt} < d_{jt}$ do
    Choose an operator $\sigma$ with minimum cost among all selection(i, k) for $(i, k) \notin \Pi_1$ and assignment(i, k) for $(i, k) \in \Pi_1$, breaking ties by the quantity of newly assigned cargo.
    if $\sigma$ is selection(i, k) then
        Select carrier $i$ and lanes in $J_{ik}$ by setting $y_{ik} \leftarrow 1$ and $\Pi_1 \leftarrow \Pi_1 \cup \{(i, k)\}$.
        Let $A$ denote the multiset of $q_{ik}$ unassigned cargos $(j, t)$ that are of the $q_{ik}$ cheapest $(p_{ijt} - s_{jt})$ among all $j \in J_{ik}$ and $1 \le t \le T$ with $\hat{d}_{jt} < d_{jt}$.
        For each cargo $(j, t) \in A$, assign it to carrier $i$ for delivering, so that both $x_{ijt}$ and $\hat{d}_{jt}$ are increased by the number of copies of $(j, t)$ in $A$.
    else if $\sigma$ is assignment(i, k) then
        Let $(j, t)$ denote the undelivered cargo $(j, t)$ that minimizes the transportation cost $(p_{ijt} - s_{jt})$ among all $j \in J_{ik}$ and $1 \le t \le T$ with $\hat{d}_{jt} < d_{jt}$.
        Assign the cargo $(j, t)$ to carrier $i$ for delivering, so that both $x_{ijt}$ and $\hat{d}_{jt}$ are increased by $\min\{c_{ik} - \sum_{j \in J_{ik}} \sum_{t=1}^{T} x_{ijt}, (d_{jt} - \hat{d}_{jt})\}$
    end if
end while
Return (x, y) as an approximation solution.
```

When $T = |B| = 1$, $|J_i| = 1$ for $i \in I$, $c_{i1} = \infty$ for $i \in I$, and $q_{i1} = b$ for $i \in I$, the above problem with the MQC constraint is equivalent to a facility location problem, where the number of customers assigned to an open facility cannot be less than b. In addition to the above b-approximation algorithm, two bi-criteria algorithms are known in the literature (Guha et al. 2000; Karget and Minkoff 2000) which achieve constant approximation ratios with regards to the optimal total cost, but which violate the lower bound constraint by a constant factor. Svitkina (2010) has recently developed a constant approximation algorithm for this problem, by transforming it to a *capacity facility location problem*, where the number of customers assigned to an open facility cannot exceed a given capacity, and for which several constant-factor approximation algorithms are known. Its approximation ratio has further been improved by Ahmadian and Swamy (2013).

For the problem with the MARC constraint, a linear programming relaxation heuristic (LP heuristic) is known to have a good worst-case performance (Lim et al. 2008a). The basic idea of the algorithm is to use a fractional solution, obtained from the linear programming relaxation of the problem, so as to decompose the problem into a number of sub-problems such that each sub-problem consists of only a single carrier. Consider a special case of the problem that consists only of a single carrier i, and for each $1 \leq k \leq |J_i|$, we still use $v_{ik} := \sum_{j \in J_{ik}} \sum_{t=1}^{T} x_{ijt}$ to represent the total volume of cargo assigned to carrier i for lanes in J_{ik} and for all the T periods. The problem for carrier i and lanes in J_{ik}, denoted by $\mathbf{IP}_{ik}$ $(\mathbf{d}, \xi)$, can be formulated as follows:

$$\mathrm{IP}_{ik}(d, \xi) = \max_{0 \leq v_{ik} \leq c_{ik}} \mathbf{IP_{ik}}(\mathbf{d}, \xi, v_{ik}) \tag{13.83}$$

where

$$\mathbf{IP}_{ik}(\mathbf{d}, \xi, v_{ik}) = \max \sum_{j \in J_{ik}} \sum_{t=1}^{T} (s_{jt} - p_{ijt}) x_{ijt} \tag{13.84}$$

$$\text{s.t.}\, x_{itj} \leq d_{jt}, \forall j \in J_{ik}, 1 \leq t \leq T, \tag{13.85}$$

$$x_{ijt} \in Z_{+}, \forall j \in J_{ik}, 1 \leq t \leq T, \tag{13.86}$$

$$\sum_{j \in J_{ik}} x_{ijt} \leq \alpha_{ikt} v_{ik} + \xi, \forall 1 \leq t \leq T. \tag{13.87}$$

As shown earlier in Sect. 3, the model $\mathbf{IP}_{ik}$ $(\mathbf{d}, \xi, v_{ik})$ can be solved in polynomial time by a standard network flow algorithm or a simplified greedy algorithm.

Consider the fractional optimal solution $\hat{\mathbf{x}}$ to the linear programming relaxation of model (MARC IP), denoted by $\mathbf{LP}(\mathbf{d}, \xi)$. From $\hat{\mathbf{x}}$ we can construct an instance of the single carrier problem for each carrier $i \in I$ and lanes in J_{ik}, denoted by $\mathbf{IP}_{ik}$ $(\mathbf{d}^{(\mathbf{i})}, \xi)$, where $d_{jt}^{(i)} := \hat{x}_{ijt}$ for $j \in J_{ik}$ and $1 \leq t \leq T$. Let $\mathbf{x}$ denote the union of the obtained solutions to sub-problems $\mathbf{IP}_{ik}$ $(\mathbf{d}^{(\mathbf{i})}, \xi)$ for $i \in I$ and $1 \leq k \leq |J_i|$. Since $\sum_{i \in I} \hat{x}_{ijt} \leq d_{jt}$, it can be seen that $\mathbf{x}$ is feasible to (MARC IP). Moreover, Lim et al. (2008a) shows that such an LP based heuristic guarantees a worst case approximation ratio of $(\frac{\xi-1}{\xi})(\frac{\xi-1}{\xi-1+\sigma})(\frac{\tau-1}{\tau})$, where parameters σ and τ are defined as follows, and which depend on the instance of the problem:

$$\sigma = 1 + \max \left\{ |J_{ik}| \alpha_{ikj} T : i \in I, 1 \leq k \leq |J_i|, 1 \leq t \leq T \right\}, \tag{13.88}$$

$$\tau = \min \begin{cases} \dfrac{c_i}{|J_{ik}|T}, & \text{for } i \in I, 1 \leq k \leq |J_i|, \\ \dfrac{\xi}{|J_{ik}|}, & \text{for } i \in I, 1 \leq k \leq |J_i|, \\ \dfrac{d_{jt}}{|I|}, & \text{for } j \in J_{ik}, 1 \leq t \leq T, \text{and existing } i \text{ with } s_{jt} - p_{ijt} > 0. \end{cases} \tag{13.89}$$

It can be seen that such an approximation ratio is close to one when σ is small, and ξ or τ is large, which are often true in practice.

As for the problem with constraints on the number of selected carriers, this contains the k-median problem as a special case, for which a number of constant approximation algorithms are known (Arya et al. 2001; Li and Svensson 2013; Vazirani 2001). However, it still remains unknown as to whether or not these approximation algorithms can guarantee constant approximation algorithms for more general cases with BG containing multiple business units.

To develop heuristic algorithms for more general cases of the TSPP, we next introduce two approaches as follows, one based on rounding of the fractional solutions, and the other based on neighborhood search.

For the problems that can be formulated as integer programming models, we can first solve its linear programming relaxation and obtain a fractional solution denoted by $(\hat{\mathbf{x}}, \mathbf{y})$. If $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ contains only integer values, then a feasible solution is obtained. Otherwise, we can select one or more variables that have fractional values, and round them to integers. Fixing the values of these variables, we can obtain an integer programming model with a smaller scale, and its linear programming relaxation can be solved for the next iteration of rounding. As shown in Algorithm 3, this process is iterated until we obtain a feasible solution.

Algorithm 3 Linear Programming Rounding Heuristic

```
1: Let $X$ indicate the list of $x_{ibjt}$ whose values have been decided. Let $Y$ indicate
   the list of $y_{ibk}$ whose values have been decided. Set $X$ and $Y$ to initially be
   an empty set.
2: Solve a linear programming relaxation of the problem for the given values of
   variables in $X$ and $Y$. Denote the obtained fractional solution by $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$.
3: if no variables in $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ are fractional then
4:     Return $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$.
5: else
6:     Select $x_{ibjt}$ or $y_{ibk}$ having a fractional value, round it to an integer, and add
       it to $X$ or $Y$.
7:     Go to Step 2.
8: end if
```

For different specific problems or heuristics, we can use different ways to select and round variables that have fractional values in Algorithm 3. For example, the linear programming rounding heuristic has been applied to solving the model (MQC IP) (Lim et al. 2006). Since this model is equivalent to a min-cost network flow problem when $\mathbf{y}$ is given, the heuristic always selects y_i of the largest fractional value and rounds it to one. The numerical experiments have shown that such a linear programming rounding method outperforms some other heuristic methods.

For a problem that can be efficiently solved when values of $\mathbf{y}$ are given, we can explore near-optimal heuristic solutions by a neighborhood search approach.

For every feasible selection **y** of carriers, let $N(\mathbf{y})$ denote a subset of selections other than **y**, defined as the neighborhood of **y**. The neighborhood search approach iteratively moves the current feasible selection **y** to another feasible selection from its neighborhood $N(\mathbf{y})$, and returns the best feasible solution obtained after several iterations. We can summarize this approach as follows:

Algorithm 4 Neighborhood Search Heuristic

```
Let $(\mathbf{x}^*, \mathbf{y}^*)$ indicate the best feasible solution obtained.
while Stop condition is not satisfied do
    Choose a feasible selection $\mathbf{y}'$ from $N(\mathbf{y})$.
    Compute $\mathbf{x}'$ by solving (General MP)($\mathbf{y}'$).
    if $(\mathbf{x}', \mathbf{y}')$ is better than $(\mathbf{x}^*, \mathbf{y}^*)$ then
        Set $(\mathbf{x}^*, \mathbf{y}^*)$ to be $(\mathbf{x}', \mathbf{y}')$.
    end if
    If certain conditions are satisfied, then update $\mathbf{y}$ by $\mathbf{y}'$.
end while
```

For different specific problems or heuristics, we can adopt different neighbor-hoods, different ways of choosing new feasible selections, and different conditions to update the current feasible selection. For the problem with fairness constraints, the neighborhood search heuristic has been applied (Lim et al. 2012), where the neighborhood $N(\mathbf{y})$ is defined by two operators on **y**, including one that removes a selected carrier and another that inserts a new carrier. The heuristic algorithm chooses a new feasible selection from $N(\mathbf{y})$ by random picking, and updates the current selection **y** only when the new selection produces a better feasible solution. In order to avoid trapping in local optimal solutions, Lim et al. (2012) has further extended this neighborhood search heuristic to a Tabu search algorithm, so that certain selections that have been made in the neighborhood will be forbidden for several iterations. Numerical results have shown that this randomized Tabu search algorithm significantly outperforms commercial optimization solvers.

13.7 Future Research Directions

In this chapter, we have introduced a general optimization model for the transportation service procurement problem (TSPP), and have reviewed various existing solution methods for different variants and extensions of it. We have discussed existing results and developed some new results for the problem, including its tractability, relaxations, exact algorithms, approximation algorithms, and heuristic algorithms. From this, there are several themes that can be identified for future research.

When choosing carriers during the procurement process, shippers are concerned not only about the shipping cost but also about the shipping time, since a slow or unreliable shipping time may increase the shipper's production cost, or increase its risk of losing sales (Lu et al. 2014). Therefore, it would be of interest to also take into account transit times, and to factor this into the optimization model for the selection of carriers. However, as this may require making joint decisions with regard to transportation service procurement, production, and sales, the problem could be challenging.

In the existing literature on transportation service procurement, models and solution methods are mainly based on deterministic settings. However, due to the high volatility of the shipping market, there are significant uncertainties involved with both spot-market shipping rates and the actual shipping demands, and these can be taken into account in the future research. This will no doubt present new challenges when one develop optimization models and solution methods for this problem.

The volatility of the shipping market presents challenges not only for optimizing decision making over transportation service procurement, but also in coordinating the carriers and shippers. In practice, carriers are sometimes reluctant to purchase transportation services from carriers in advance, since they are concerned that the spot market price may suddenly drop. Thus, it would be interesting to design and study various forms of contracts that can facilitate the business between carriers and shippers, such as contracts that enable the sharing of risks and costs (An et al. 2014; Lee et al. 2014).

Furthermore, it would also be interesting to study the design of mechanisms for transportation service procurement. Since most problems faced in practice concern computational intractability, only heuristic algorithms are available for them. Therefore, Vickrey-based payment rules (Clarke 1971; Groves 1973; Vickrey 1961) are often not able to guarantee positive revenue for the shipper (Conitzer and Sandholm 2006; Parkes et al. 2001). Thus, in future work we can investigate effectiveness of different payment rules for situations that use heuristic algorithms (instead of exact algorithms) for the shipper to determine cargo allocations (Huang and Xu 2013; Xu and Huang 2013; Xu and Huang 2014).

Other interesting research directions may include the development of efficient solution methods for the general optimization model for the TSPP, as well as for those problems with more complicated cost structures, such as discounts that carriers may offer for multiple tiers of shipping volume (Qin et al. 2012).

Acknowledgement The work presented in this chapter was partly supported by the Hong Kong Polytechnic University under the research grant number G-YL96.

References

Ahmadian, S., & Swamy, C. (2013). Improved approximation guarantees for lower-bounded facility location. In *Approximation and online algorithms* (pp. 257–271). Springer: Berlin Heidelberg.

An, J., Lee, C.-Y., & Tang, C. S. (2014). Fractional price matching policies arising from the ocean freight service industry. Working paper.

Arya, V., Garg, N., Khandekar, R., Munagala, K., & Pandit, V. (2001). Local search heuristic for k-median and facility location problems. In *ACM Symposium on theory of computing*, pp. 21–29.

Bozkaya, B., Zhang, J., & Erkut, E. (2002). A genetic algorithm for the p-median problem. In Z. Drezner & H. Hamacher (Eds.), *Facility location: Applications and theory*. Berlin: Springer.

Caplice, C. (2003). Optimization-based procurement for transportation services. *Journal of Business Logistics, 24*(2), 109–128.

Caplice, C., & Sheffi, Y. (2005). Combinatorial auctions of truckload transportation. In P. Cramton, Y. Shoham, & R. Steinberg (Eds.), *Combinatorial auctions* (pp. 539–572). Cambridge: MIT Press.

Clarke, E. H. (1971). Multipart pricing of public goods. *Public Choice, 11,* 17–33.

Conitzer, V., & Sandholm, T. (2006). Failures of the VCG mechanism in combinatorial auctions and exchanges. In *Proceedings of 5th international joint conference on autonomous agents and multiagent systems*, Hakodate, Japan, pp. 521–528.

de Farias, I. R. (2001). A family of facets for the uncapacitated p-median polytope. *Operations Research Letters, 28*(4), 161–167.

Garey, M. R., & Johnson, D. S. (1983). *Computers and intractability: A guide to the theory of NP-completeness, chapter A3.1 covering, hitting, and splitting* (p. 221). New York: Freeman.

Groves, T. (1973). Incentives in teams. *Econometrica, 41,* 617–631.

Guha, S., Meyerson, A., & Munagala, K. (2000). Hierarchical placement and network design problems. *In proceedings of the 41th annual IEEE symposium on foundations of computer science*, pp. 603–612.

Hochbaum, D. S. (1982). Heuristics for the fixed cost median problem. *Mathematical Programming, 22,* 148–162.

Huang, G. Q., & Xu, S. X. (2013). Truthful multi-unit transportation procurement auctions for logistics e-marketplaces. *Transportation Research Part B: Methodological, 47,* 127–148.

Karget, D. R., & Minkoff, M. (2000). Building steiner trees with incomplete global knowledge. In *Proceedings of the 41th annual IEEE symposium on foundations of computer science*, pp. 613–623,

Lee, C.-Y., Liu, Q., & Zheng, S. (2014). A novel capacity allocation contract in the ocean freight industry. Working paper.

Li, S., & Svensson, O. (2013). Approximating k-median via pseudo-approximation. In *proceedings of the forty-fifth annual ACM symposium on theory of computing*, pp. 901–910.

Lim, A., Wang, F., & Xu, Z. (2006). A transportation problem with minimum quantity commitments. *Transportation Science, 40*(1), 117–129.

Lim, A., Rodrigues, B., & Xu, Z. (2008a). Transportation procurement with seasonally-varying shipper demand and volume guarantees. *Operations Research, 56*(3), 758–771.

Lim, A., Xu, Z., & Wang, F. (2008b). The bidding selection and assignment problem with minimum quantity commitment. *Journal of the Operational Research Society, 59*(5), 693–702.

Lim, A., Qin, H., & Xu, Z. (2012). The freight allocation problem with lane cost balancing constraint. *European Journal of Operational Research, 217*(1), 26–35.

Lorena, L. A. N., & Senne, E. L. F. (2004). A column generation approach to capacitated p-median problems. *Computers and Operations Research, 31*(6), 863–876.

Lu, T., Fransoo, J. C., & Lee, C.-Y. (2014). Carrier portfolio management in liner shipping. Working paper.

Parkes, D. C., Kalagnanam, J., & Eso, M. (2001). Achieving budget-balance with vickrey-based payment schemes in exchanges. In *Proceedings of 17th international joint conference on artificial intelligence, Seattle*, Washington, pp. 1161–1168.

Qin, H., Luo, M., Gao, X., & Lim, A. (2012). The freight allocation problem with all-units quantity-based discount: A heuristic algorithm. *Omega, 40*(4), 415–423.

Rolland, E., Schilling, D. A., & Current, J. R. (1997). An efficient tabu search procedure for the p-median problem. *European Journal of Operational Research, 96*(2), 329–342.

Sawhney, M. (2003). Forward thinking about reverse auctions. June 1, 2003 issue of CIO Magazine, http://www.cio.com/. Accessed 1 June 2014.

Senne, E. L. F., Lorena, L. A. N., & Pereira, M. A. (2005). A branch-and-price approach to p-median location problems. *Computers and Operations Research, 32*(6), 1655–1644.

Svitkina, Z. (2010). Lower-bounded facility location. *ACM Transactions on Algorithms (TALG), 6*(4), 69.

Vazirani, V. V. (2001). *Approximation algorithms*. Springer: Berlin Heidelberg New York.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance, 16*(1), 8–37.

Xu, Z. (2007). *Transportation procurement planning with side constraints*. PhD thesis, the Hong Kong University of Science and Technology.

Xu, S. X., & Huang, G. Q. (2013). Transportation service procurement in periodic sealed double auctions with stochastic demand and supply. *Transportation Research Part B: Methodological, 56,* 136–160.

Xu, S. X., & Huang, G. Q. (2014). Efficient auctions for distributed transportation procurement. *Transportation Research Part B: Methodological, 65,* 47–64.

Chapter 14
Ocean Transport and the Facilitation of Trade

Albert Veenstra

Abstract In this chapter, we describe transactions in international trade transaction, and we zoom in on the ocean transport part of that transaction. We use transaction cost economics as our theoretical framework. We highlighted a number of ocean transport related processes that generate uncertainty and costs in logistics chains: the use of the container in pre- and on-carriage, the release process of the container in port, and the formalities related to the supervision on vessels coming into ports in the country of destination. A first estimate of the uncertainties and costs that follow from these processes (delays, additional time required for supervision) reveals that the uncertainties far outweigh the additional costs. This holds especially because companies take into account a certain degree of uncertainty in their logistics planning, even though the probability is very low. We extend this discussion with the fundamental problem of the information quality of the ocean carrier documentation, the ship manifest, which has negative consequences for risk assessment by Customs in Europe.

We conclude that specific processes that are connected to ocean transportation do result in time loss, uncertainty and, to a lesser extent, in additional costs that impact the efficiency of logistics and supply chains. The transaction costs are predominantly generated as additional charges on the ocean transport bill, as a result of supervision, and as a result of the use of the container. Transaction cost theory predicts that if such frictions exist, there will be a tendency to move from a market relationship to a more hierarchical relationship between parties involved in the transaction. We observe that international transportation by sea is in an intermediate position between market and hierarchy and we identify at least two important developments in the Port of Rotterdam and its European hinterland that can be recognized as a further shift towards more hierarchy as a result of the frictions we identified.

A. Veenstra (✉)
School of Industrial Engineering and Innovation Sciences, Department of Operations, Planning, Accounting and Control, Eindhoven University of Technology, Eindhoven, The Netherlands
e-mail: albert.veenstra@tno.nl

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics*,
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_14

14.1 Introduction

It is common knowledge that ocean transport is the major carrier of international trade, at least in terms of weight of cargo. A large part of this international trade nowadays constitutes the movement of goods within company supply chains. If trade in intermediate goods is taken as a proxy for intra-supply chain movement of goods, The Organization for Economic and Cultural Development (OECD) estimates that 56 % of trade in goods is supply chain related trade (Miroudot et al. 2009). In this chapter, we concentrate on the transportation of containerized goods, for which this proportion is higher still.

The regulation for the international movement of goods attaches much importance to the Bill of Lading. This document is at the same time a receipt that goods are in the custody of a ship's captain, proof of a contract for transportation, and a document of title for the owner of the goods. For every consignment that is presented to a ship's captain in port, the party that presents the cargo receives such a Bill of Lading, and with this Bill of Lading that can be endorsed to a buyer, the holder can reclaim the cargo in the port of destination.

In general trade that is not supply chain related, the basic uncertainty that buyers and sellers do not know and trust each other, has resulted in an arrangement where the Bill of Lading is also an important trigger for payment. Once the goods are formally out of the hands of the seller, the buyer can pay the seller, even though the goods may not have arrived at the buyer's premises. When moving goods in supply chains, this same mechanism is often used, because tax regulations in many countries require companies within the same group to do so-called arms-length trading (see for instance Miller and Matta 2008). This means they have to act as if the sister-company or subsidiary is an unrelated company.

The mechanism of the Bill of Lading thus plays an important in supporting trade as well as the movement of goods in global supply chains. The importance of the ship manifest, as a collection of bills of lading, goes further. When moving goods across borders, government agencies, such as Customs, become involved in charging taxes and supervising the impact of the goods on the national economy which they represent. In many countries in the World, the ship's manifest is an important document for this type of supervision. It is used as a basis for risk assessment, or, alternatively, as a means to verify which imported or exported goods actually moved across the border.

In this chapter, we link this mechanistic view on ocean transport with trade facilitation and supply chain efficiency. We argue that disruptions in the physical and administrative operations of ocean shipping are a barrier to trade in the sense that they cause delay, uncertainty and additional costs. In addition, we will illustrate that the basic documentation for ocean transport that is now used for supervision is suboptimal in terms of the quality of data, and that again causes some delays and uncertainties at the supply chain level. In this chapter we will concentrate on transportation of containers, because the value of time for goods in containers is generally higher, and frictions in operations have generally more negative impact.

This chapter will address the consequences of the current, complex ocean container transportation mechanism for the facilitation of trade, and, since a large part of trade is supply chain related, for global supply chains. The methodological framework we will employ to characterize these consequences is transaction cost economics. Transaction cost economics was introduced by Williamson (1981). This branch of economics studies transactions in and between firms. In standard economic theory, transactions are often assumed to happen instantaneous. In reality, of course, they do not. Executing transactions takes time and effort. As a result of this, economic subjects, such as firms therefore economize on transaction costs, and this has impact on their economic decision making. This chapter aims to describe and analyze the "ocean shipping transaction" in some detail, in order to understand what impact the execution of this transaction may have on the supply chain and logistics operations of which it is a part.

This chapter is organized as follows. We first elaborate on transaction cost theory applied to ocean transportation. We then present an overview of ocean transport and the international movement of containerized goods through ocean shipping. We briefly summarize the main consequences, and attempt to measure their magnitude. The subsequent section deals with international supervision and consequence of the use of the ship's manifest for supervision by customs. In the final section, we formulate the implications for international trade, and a related research agenda for ocean transport.

14.2 Transaction Cost Economics for Ocean Container Shipping

In international trade, buyers and sellers of goods find each other, and conduct a commercial transaction. In this transaction, important variables such as the exact description of the good, the amount of goods, the commercial value of the goods and production and delivery deadlines are agreed. An important part of the transaction is also the arrangement for the transfer of ownership (and the resulting payment), and transportation of the goods.

For centuries, the arrangements on change of ownership, payment and transportation are connected. The reason for this is that in international trade, buyers and sellers may not know or trust each other at all. The execution of the commercial transaction in terms of payment and transfer of the goods can then become a difficult problem: should the buyer pay first, and receive the goods later, or should the seller send the goods first, and get paid later? This conundrum was solved by connecting payment to the movement of the goods through a third party. This third party is a specialized carrier of goods, who takes care of the movement of goods from the seller to the buyer's premises (or some destination designated by the buyer). As soon as this party takes control of the goods, proof of this will be sent to both buyer and seller, and payment can take place.

In trade practice, several standard arrangements have been formulated that are known as Incoterms. These are standard trade terms that are maintained by the

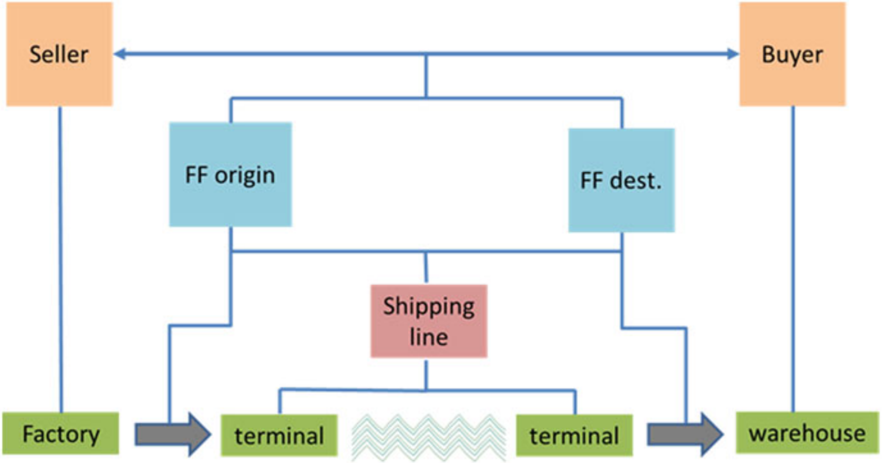


Fig. 14.1 General representation of a trade transaction (*FF* stands for Freight forwarder. We use this as a generic term for a combined forwarder, customs agent and local transport operator function)

International Chamber of Commerce[1]. The incoterms differ in terms of the location in the transport chain between buyer and seller where specific responsibilities for the goods are transferred: ownership, transportation and insurance of the goods. Standard locations of transfer are: the seller's premises, the port of loading, the port of discharge, or the buyer's premises. See for an interesting discussion on the interaction between Incoterms and logistics planning, Kumar (2010).

The mechanism to move goods by sea is centuries old. These mechanisms are codified in several international conventions, such as the Hamburg Rules, the Hague Visby Rules and more recently, the Rotterdam Rules. In quite a number of countries, the Hamburg Rules are still the basis for legislation. This convention stems from 1924. For a general source on these conventions, see Murray et al. (2012).

In this regulatory framework, the proof for the transport party to take control of the goods is called the Bill of Lading, as was already introduced above. Furthermore, this bill of lading sets off the payment cycle between buyer and seller and their banks, and it is proof for the seller to claim the goods from the transport operator.

But the framework of a general trade transaction is more complicated, since it contains more parties than the buyer, seller and ocean carrier. Observe the following picture (Fig. 14.1).

The Figure makes a distinction between the commercial seller and buyer, and the manufacturer (factory) and receiver (warehouse), which, more often than not, are separate companies. The party that sends the goods (termed the consignor) is

---

[1] http://www.iccwbo.org/products-and-services/trade-facilitation/incoterms-2010/.

the factory, and the party that receives the goods (the consignee) is often the party running the warehouse in the country of destination.

As soon as the buyer and seller have divided responsibilities, they instruct their freight forwarders to make the necessary arrangements. For the incoterm Free on Board (FOB), the freight forwarder at destination will manage the ocean transport booking, and it will instruct the freight forwarder at origin accordingly. In many cases, the freight forwarders prefer not to reveal their customers, for fear of direct competition from the shipping line. They obscure their clients' names by offering their clients a so-called House Bill of Lading. This Bill of Lading will state as consignor the factory, but as the carrier, the freight forwarder. They then contract separately with the ocean carrier. The Bill of Lading from the Ocean Carrier, which is then referred to as the Master Bill of Lading, will then list as consignor the freight forwarder, and possibly as consignee the freight forwarder at destination. This is then the information that goes into the ship manifest. This way of working has big repercussions for supervision, as we will see later in this chapter.

The ocean carrier contracts with the terminals. Nor the buyer and seller, nor the freight forwarders at either side, have much influence on the choice of ports and the choice of terminals. They also have no direct contracting relationship with the terminals.

From a Transaction Cost Economics (TCE) perspective, this complete international trade transaction seems to be almost completely market-based, in the sense that the links between the various parties are all contracting relationships, and not internal firm procurement arrangements. In the terminology of TCE, this reflects the low degree of site, asset and human specificity (Williamson 1981, p. 555 ff.). Site specificity refers to so-called core technology that is crucial to the transaction and that parties in a transaction prefer to co-own. However, in ocean shipping, the core technology can be partitioned in ships, loading units and cargo handling technology that is sufficiently standardized to be separable across parties. Asset specificity refers to the degree to which assets are dedicated to the transaction. In ocean container transport, assets are not specific to a single transaction, but can facilitate many transactions at the same time, and none of the transactions require a specific ship. There may be a connection to the transaction and a single container, but again, there no specific individual container is required for a transaction. We will focus on the role of the container further below, and we will see that the container generates much of the frictions that cause delays and costs. Finally, the human specificity refers to specific knowledge that is requires for the transaction. Again, this is not the case in the sense that specific individuals need to be involved. There is specific knowledge, but this knowledge is shared among many individuals, and as such individual knowledge is not specific to any transaction.

The benefits of market relationships in a transaction are that scale economies in assets can be fully utilized, uncorrelated demand can be aggregated and bureaucratic hazards related to internal purchasing processes are avoided. Drawbacks are suboptimization in parts of the transaction, costly conflict resolution, and poor access to relevant information for the settlement of disputes or measurement of performance (Williamson 1981).

Williamson's theory predicts that under uncertainty and with increasing specificity of especially assets (ships, containers, handling equipment), the benefits of market relationships disappear, and governance costs in the transaction increase. This will eventually result in a move towards a more hierarchical relationship, whereby the transaction becomes an internal purchasing transaction. Williamson identifies an intermediate stage, which he calls semi-specific transaction governance (Williamson 1979, p. 247). This is a stage in which the governance of the contractual relationship is not an intra-firm arrangement yet, but it is more specific, more personal, one could say, than a market relationship. Looking at the international trade transaction in this way, it is clear that most relationships within the transaction are in fact of this intermediate nature. The relationships between parties are mostly long term relationships, and the execution of elements of the transaction (for instance the arrival of a ship in port) requires planning and coordination between transaction parties, and with external, parties as well. For this, parties have to know each other.

Against the background of TCE, it is interesting to investigate the various elements of the international trade transaction in more detail, to see if there are tendencies towards increasing transaction costs, and if these tendencies could be an early sign for further hierarchical integration within the international trade transaction.

We interpret this transaction cost approach as a method to assess the level of difficulty of performing an international trade transaction. TCE provides us with a dynamic approach for this assessment that consists of the measurement and explanation of the source of transaction costs, and the expected changes in the relationships between parties in the international trade transaction to mitigate increasing transaction costs.

14.3 The Ocean Shipping Sub-Transaction and Trade Facilitation

In this chapter, we focus on the ocean transport part of the international trade transaction. We leave a similar analysis of the international sale and purchase transaction for further research. We concentrate on the relationship between the buyer/seller and the ocean carrier. The relationship between the ocean carrier and the terminals are outside the scope of our analysis.

The contracting of the shipping line is what we call the "ocean shipping transaction". Of course, physical ocean transportation is the core of this transaction. However, for the containerized goods, other important elements in the transaction are:

1. Pre-carriage: the ocean carrier can be asked to pick the container in some inland location, and bring it to the port of loading,
2. Formalities that the ocean carrier has to fulfill to bring cargo into a country.
3. On-carriage: the ocean carrier can be asked to deliver the container to some inland destination,

4. Use of the container as a loading unit for the goods,
5. Release of the container from the port, in case end carriage is done by the buyer,

We will discuss each of these transaction elements in detail.

14.3.1 Pre-Carriage

The business of ocean carriers is to sail ships from one port to another. This means that goods to be carried need to be presented to them in the port of loading. However, for containerized goods, the ocean carrier becomes involved in the process earlier, because an empty container has to be provided to the seller. This is a standard part of the transport booking process, since the ocean carrier is usually the owner of the containers. The seller therefore requires a location to pick up the empty container from, or may ask the ocean carrier to deliver an empty container to its premises for loading. In the latter case, the ocean carrier will hire a local truck company to pick up the empty container, deliver it to the seller for loading of the goods, and pick it up again for delivery to the port. In practice, however, this does not occur that often, because sellers can usually arrange local transport easier than the shipping line.

Given that the ships of the ocean carrier sail in a schedule, it will communicate a deadline to the seller before which the containers with the goods have to be in port. This deadline is called the cut-off date. Containers that arrive after that date cannot be loaded on the original ship, but will have to be re-booked to the next ship.

Even when a container arrives in the terminal on time, it may still be delayed and miss the ship. The main reason for this is that export customs may want to scan of physically inspect the container before it is loaded on board of the ship. In such a situation, the container is blocked for loading by customs and this can easily several days. The cut-off date will then be missed.

14.3.2 Formalities in the Country of Arrival

Coastal ports are usually country border as well. Most countries in the world have the basic arrangement that ships can drop off cargo in a special customs controlled zone in a country, after which on-carriage can take place if certain conditions are met. Ocean carriers have to inform the proper authority what they are bringing into the port, and they do this by means of the ship's manifest. To bring the goods into the country's territory, a further declaration to customs is usually required either of the intention to move the goods to another customs controlled location, or of the intention to bring the goods in free circulation[2].

---

[2] There are many other possibilities, such as temporary import, or temporary import for the purpose of processing, but we do not consider these arrangements here.

Countries differ not so much in this basic arrangement, but in the order and timing in which they require this information. In Europe, the ship's manifest is submitted first, and followed by declarations later, but in Singapore and Hong Kong, the importer requests an import license first, which is then checked against the ship's manifest. In the US, and in many other countries and areas nowadays, preliminary manifest information is requested before the ship departs from the port of origin. In the US, additional information on all relevant parties in the chain is also requested at this time (the so-called $10 + 2$ requirement[3]).

The customs filing requirements before commencement of loading of the ship are performed by the ocean carrier, but the information they need is supplied by either the seller, or his freight forwarder. Given there is a deadline for the carrier to provide this information, this deadline is part of the booking requirements for the clients of the ocean carrier. Most carriers also charge a fee for this filing, of US$ 25[4].

Apart from the ship's manifest, any ship entering a port has to submit a host of other documents: lists of personnel on board, of all non-commercial goods on board, technical certificates, and so on.

Upon arrival of the goods, the unloaded cargo, or cargo to be loaded is in the container terminal, which is a controlled customs zone. The registration of goods in this zone is not a responsibility of the ocean carrier, but of the terminal operator. The terminal is also responsible for managing the release of the container to leave this customs controlled zone and be transported into the country, both on behalf of customs and on behalf of the ocean carrier.

In many countries, certainly in Europe, the ship manifest is the basis for a first risk assessment by customs. There are four possible outcomes of this analysis:

1. The container was considered no risk. It can be moved out of the port if customs is properly informed about the destination and status of the goods,
2. Customs requires additional information to assess the risk. This usually means that the ultimate sender and/or receiver need to be revealed to customs, or the value of the good has to be proven with an invoice,
3. Customs requires the container to be scanned,
4. Customs requires the container to be physically inspected.

The involvement of customs thus introduces uncertainty in the process of moving goods into the country, as well as a possible delay if additional activities are required to satisfy customs.

14.3.3 On-Carriage

The buyer may approach the ocean carrier to solve their problem of moving the goods out of the port to the buyer's premises. This process occurs frequently, and therefore

---

[3] See http://www.cbp.gov/border-security/ports-entry/cargo-security/importer-security-filing-102.

[4] See for instance www.maerskline.com.

acquired a dedicated name: carrier haulage. This is the end haulage that is arranged by the (ocean) carrier, *under the same transport conditions that apply to the ocean transport leg*. This last part is relevant, because it is an important extension of the legal framework that governs ocean transport. Air transport has a similar extension.

The use of a carrier haulage arrangement makes it possible for the ocean carrier to gain control of transportation into the hinterland of a port, in such a way that the connection with the ocean transport schedule is optimized, contracting time for the on-carriage leg is negligible, and control over the flow of containers (which the carrier would like to receive back empty as soon as possible) is enhanced.

This is the dominant mechanism for transport to the United States of America, where on-carriage is mainly conducted by rail under ocean B/L condition. In Europe, carrier haulage was also dominant, until the early 1990s. After that period, on-carriage was mostly arranged by transport operators on behalf of the buyers, which is termed merchant haulage. The split carrier/merchant haulage in Europe is now said to be 30 %/70 % (NEA 2010). One important reason for this is that the liability of the transport operator goes much further when trucking takes place under a trucking contract compared to trucking that takes place under the ocean bill of lading conditions. Ocean transport has a rather low liability cap of about 2 SDR per kilo, while the trucking liability if capped at 8 SDR per kilo. Owners of high value goods will therefore prefer to use their own trucking operators instead of trucking under the ocean bill of lading.

14.3.4 The Container as a Loading Unit

The container is not a very valuable item in itself. A new standard 40 ft container may cost around $ 2500–2800. However, the ocean carriers own substantial numbers of containers. A large shipping company, such as Maersk Lines owns up to 3.8 mln containers[5]. Their entire fleet of containers thus represents a significant investment.

In addition, containers generate revenue if they are filled with cargo for a paying customer. To optimize the revenue of this important asset, ocean carriers attempt to return empty containers to important loading areas as soon as possible. To achieve this, all ocean carriers have devised a penalty system to induce their customers to pick up the containers upon arrival as quickly as possible, and to bring back the empty container as quickly as possible. If a full container is picked up late, it is said to be in “demurrage”, and a daily penalty is due. If an empty container is delivered back late, it is said to be in “detention” (at the customer) and a daily penalty is due. As a courtesy to the customer, some days are allowed before the demurrage period starts, and some days are allowed before the detention period starts. These numbers of days, as well as the penalties are negotiable. Especially for large customers, much

---

[5] www.mymaerskline.com, visited dd. 21 May 2014.

more free days and much lower penalties usually apply compared to the average client.

Apart from the asset management perspective, the ocean carriers have another reason why they have to keep track of containers. Tax authorities usually consider the container as a good in its own right, that needs to be declared to customs, and for which value added tax must be paid upon entry into a country. If the good is in fact packaging material, and the import is only temporary, duties and taxes may be deferred, or set to zero. In return, however, the ocean carriers have to keep track of the containers and how long they have been stationed in the country.

All this leads to some interaction between the ocean carrier and the buyer and/or his transport operator in especially the port of arrival, about who will pick-up of the container, the exact date of arrival of the container (for the calculation of the demurrage period), and the location where the empty container needs to be delivered back—the so-called empty depot. The ocean carrier issues a pin-code with which the legitimate party can picks up the container from the port. Sometimes, some items on the transport bill need to be paid first, before the ocean carrier issues these crucial information items (pin code, pick up location in port, empty depot).

These arrangements introduce restrictions on the transport operations in the hinterland: transport needs to take place in the narrow time window between the demurrage and detention deadlines. If these deadlines are violated, the buyer incurs a penalty.

14.3.5 Release of the Container in Port

There are various reasons why containers cannot leave a port without some form of control: customs needs to release the container after assessing risks, the carrier needs to be sure the hinterland transport operator is the right party, and known where the empty container needs to go, and the terminal is responsible for verifying these conditions.

As a result of this, the terminals have an elaborate identification procedure for hinterland transport operators, in which they verify who they are, if they have the right pin codes for the containers, if the containers have the right status to be moved and if the required documentation for this status are available. If all these requirements are met, the container can be released to the transport operator.

Currently, many terminals are automating this identification process with biometric identity cards, and digital documentation. This is, however, a slow process, especially for export cargo. A terminal can anticipate on the release requirements if it knows the container, and there is some time to inform the responsible party of the release status. This holds for import containers, which are unloaded from a vessel. These containers usually spend some time in a terminal, and this time is enough to communicate with the outside world about the status, missing documents and so on. For export containers, any gaps in the requirements will only be identified upon arrival of the transport operator at the terminal. This means the driver will have to correct the problem before the terminal will accept the container for loading on a ship, with the risk that it misses the ship on which it was intended to be loaded.

We will discuss these uncertainties and costs in more detail in the next section, where we will identify some common mitigating measures, and estimates of the delay and cost items.

14.4 A First Cost Estimate

In this section, we provide a first estimate of the uncertainty and cost items for the following transaction elements we have identified above:

- Ocean transport: schedule reliability (delay), and additional cost items
- Pre-carriage: no relevant items to discuss
- Formalities: details related to pre-departure declarations, and pre-arrival manifest declarations,
- On-carriage: no relevant items to discuss,
- Container: overview of demurrage and detention by carriers,
- Release: impact of customs control activities.

We will not discuss pre- and on-carriage in detail, because there are no relevant delay or cost items to discuss.

Our estimates are based on a case study by Veenstra et al. (2013) for the Dutch Advisory Council for Regulatory Pressure concerning administrative costs related to government and self-regulation of companies involved in the international movements of goods. In this study, detailed cost and time data was collected for the items listed above. We will express all costs as much as possible in time and euro per container, calculated over all containers entering the Port of Rotterdam.

In terms of the general validity of this data, we consider that the Port of Rotterdam is the largest port in Europe. Its throughput constitutes a substantial share of European trade. In addition, much of the relevant regulation is European regulation or general ocean carrier operating practices, which means that much of the data has a broader validity than this specific case. As such, we present our estimates of additional costs and time as estimates that are valid at least for all cargo transported to and from Europe.

Schedule Reliability and Additional Costs Several authors have reported on ocean carrier schedule reliability. Vernimmen et al. (2007) report average schedule performance based on a survey performed by Drewry Shipping Consultants in 2006: of 65 shipping lines investigated, 15 had a schedule reliability of 60 % of more, 12 had a schedule reliability between 50 and 60 %, 38 carriers had a schedule reliability below 50 %. More than 40 % of vessels arrive one or more days late. The distribution of on-time/late arrival, according to Vernimmen et al. (ibid, p. 194) is: 52 % of vessels are on time, 21 % arrives more than 1 day but less than 2 days late, 8 % arrives more than 2 but less than 3 days late, 14 % arrives more than 3 days late and the remaining 5 % was more than 1 day early, where on-time arrival is defined by Drewry as arrival on the scheduled day *plus/minus 1 day*. Chung and Chiang (2011) report on later surveys by Drewry (2007–2009). They also report very low numbers of perfect schedule reliability: only a handful of 60 carriers have a reliability of 90 % or more. These numbers translate in an average schedule deviation of 1–1.5 day.

The ocean carriers maintain a complicated tariff structure for their service. There is a base rate for the carriage of containers, which is based on published tariffs, but negotiable for large customers. An average rate for a container shipped on various important routes has an order of magnitude of around US$ 2000 for a 40 ft container. In addition to this base rate, they charge a number of additional fees: special services fees (for reefers, special care, non-standard containers), bunker adjustment factor (BAF), currency adjustment factor (CAF), port congestion surcharge, peak season surcharge, terminal handling charges, war risks, ISPS surcharges, winter surcharges, piracy risk, dangerous goods and refrigerated goods special treatment charges and document fees (see for an overview and analysis of the relationship of the BAF and underlying costs, Cariou and Wolf 2006). Security charges are around US $ 10–20, BAF/CAFs are usually several hundred US$ per TUE. All in all, the surcharges can add US$ 1000 or more to the freight bill. In general, these costs are not considered to be additional costs. Many parties consider these costs part of the transport bill. There is a tendency, however, of shippers to misdeclare the nature of the goods to avoid paying certain charges. This occurs frequently for especially dangerous goods.

Formalities- ENS The Entry Summary Declaration for Security for Europe (ENS) is part of the Modernized Customs Code that was adopted on 23 April 2008 as Regulation (EU) 450/2008[6]. The ENS has to be lodged 24 h before commencement of loading of the ship in the port of origin, and has to be sent to the first port of call in the European Union territory.

The general arrangement for Europe is that the ocean carriers collect the ENS data, and send it as one batch to Customs in Europe. The data needs to be presented to the ocean carrier by the sellers or their representatives, the freight forwarder.

The freight forwarder needs to be instructed and has to compile the relevant information. In Veenstra et al. (2013), the time required to instruct the freight forwarder, as well as the time to send the data to the ocean carrier was estimated to be 25 min. Based on the number of bill of ladings for containers coming into the Port of Rotterdam, and the hourly tariff for a local freight forwarder, this translates in a cost of € 3.50 per container. In addition, $ 25 or about € 19 is charged per bill of lading by the carriers for submitting the ENS. Given there are on average 2 containers on a bill of lading[7], this amounts to a further € 9.50 per container.

The consequences of the ENS can be:

1. Risk type A: A no load decision from the Customs authority in Europe. In this case the container may not be loaded for transport to Europe.
2. Risk type B: A decision to intercept at the first port of call in Europe,

---

[6] Official Journal L 145 4.6.2008, p. 1.

[7] Private communication with Dutch Customs.

3. Risk type C: A decision to intercept at the port of discharge in Europe.

If these risks do not apply, there will be no interference of Customs with the cargo as a result of the ENS declaration.

There is no general information on the number of decisions in the above categories. There is some anecdotal evidence that the Risk type A is very rarely identified, as well as the risk type B. Risk type C is more common, but the probability is a single digit number.

Formalities-SAL The European Customs regulation stipulates that 72–24 h before arrival of the vessel, the ship's manifest has to be sent to customs. This is called the summary declaration at unloading, or SAL in the Netherlands. This document is then used for risk assessment. If the receiving company is AEO, in principle, the data from the ENS declaration can be re-used for the SAL declaration. It is not easy, however, to communicate to Customs that receiving companies are AEO, because the SAL declaration does not contain a free field to enter an AEO registration number.

Overall, the work involved in compiling the SAL falls to the ocean carrier, and is a matter of copying manifest lines for the containers to be unloaded one by one into the SAL declaration page. This copying is done by hand, because the space for cargo descriptions per line in the declaration in the SAL are limited and some copying, pasting and summarizing has to be done to end up with satisfactory and more or less complete cargo descriptions in the SAL. In Veenstra et al. (2013), the costs per container for submitting the SAL were estimated, but they were negligible (less than € 1 per container).

What is relevant is the outcome of the risk assessment. There can be four possibilities (Veenstra et al., ibid.):

1. The container has not been targeted, and can be released if follow-up activities are known to customs,
2. Additional information is required. This occurs for 15–20 % of incoming containers, and usually means Customs requires an invoice or house bill of lading to verify the actual consignee and consignor,
3. The container needs to be scanned. This occurs in about 2 % of the cases, and the resulting additional costs are € 275,
4. The container needs to be physically inspected. This occurs in about 1 % of the cases (largely the same cases as the scanned containers), and the resulting additional costs are € 1200.

For the Netherlands, the administrative efforts are included in the total fee of the freight forwarder, but our estimate is that this costs about € 10 per container[8]. The scanning and physical check adds another € 26 per container (Veenstra et al. 2013).

---

[8] The total fee for a freight forwarder for administrative handling of a container in the port of destination is about € 35–55, and about € 25 of this amount is involved in regular activities. We estimate that of the remainder € 10 is involved in dealing with administrative issues related to Customs checks, and another € 5 for logistics preparation for scanning and inspections (see Veenstra et al. 2013).

Table 14.1 Overview of demurrage and detention regimes. (Source: Storm 2011)

Days	Demurrage working days		Detention calender days		Combined calender days	
Mode	Truck	Train/barge	Truck	Train/barge	Truck	Train/barge
Maersk	3	3	3	5		
MSC					3[a]	7[a]
CMA CGM					7	7
COSCO					7	7
Hapag Lloyd	5	5	2[b]	6[b]		
Evergreen[c]	3	3	2[a]	6[a]		
APL	5[a]	5[a]	3	6		
CSCL	3	3	2[a]	10		
Hanjin Shipping	4	4	5[a]	5*		
MOL	3[a]	3[a]	3	5		
NYK Line	4	4	6	6		
Hamburg Süd					7	7
OOCL	3	3	2[a]	6[a]		
CSAV	4[a]	4[a]	6	6		
Yang Ming	4	4	3[a]	3[a]		

[a] Opposite (calendar or working day)
[b] Calendar days, but first weekend/holiday excluded
[c] Different days per trade lane; this data for Asia European trade lane

Container Demurrage and Detention All ocean carriers maintain a demurrage and detention regime. The regime contains free days, after which the fees for demurrage or detention apply. The regimes for various ocean carriers are listed in the Table 14.1 below. Observe that there are quite some differences between the regimes: some carriers have separate demurrage and detention regimes, and some have a combined regime. Furthermore, the regimes make a distinction between hinterland modes (truck, and barge/rail), which hardly transfers to the free days. These regimes therefore tend to favor trucking as the hinterland mode.

In addition to the days, there are also fees. These are depicted in Table 14.2. Note that here, there are considerable differences between the fees. Some carriers heavily penalize late pick up of return of boxes (for instance Maersk Line), and others charge very low fees (for instance MSC).

The estimate of the impact of demurrage and detention regimes is difficult for a number of reasons:

- The days and fees are negotiable. It is therefore impossible to assess what rate apply to the general population of containers coming into Europe (or any other area in the World),

Table 14.2 Overview of demurrage and detention fees. (Source: Storm 2011)

stages	Demurrage costs		Detention costs		Combined costs	
	1st stage	2nd stage	1st stage	2nd stage	1st stage	2nd stage
Maersk (€)	45 (1–7)	75 (8 >)	35 (1–7)	35 (8 >)		
MSC (€)					10 (1–7)	20 (8 >)
CMA CGM (€)					12 (1–7)	20 (15–30)
COSCO (€)					30 (1–6)	60 (7 >)
Hapag Lloyd (€)	22 (1–5)	34 (6 >)	25 (1–3)	50 (4 >)		
Evergreen (€)	13.61 (1–4)	20 (8–20)	3.63 (1–2)	7.26 (4 >)		
APL (€)	16 (1–5)	33(6–45)	10 (1–3)	20 (4–21)		
CSCL (€)	15 (1–2)	30 (3 >)	10 (1–3)	20 (4 >)		
Hanjin Shipping (€)	37 (1–5)	75 (6–45)	23 (1–20)	45(21 >)		
MOL (€)	25 (1–5)	45 (6 >)	15 (1–3)	30 (4 >)		
NYK Line (€)	27.50 (1–5)	44 (6 >)	15 (1–10)	35 (11 >)		
Hamburg Süd (€)					20 (1–4)	40 (5 >)
OOCL (€)	40 (1–10)	50 (11–30)	25 (1–3)	50 (4 >)		
CSAV (€)	45 (1–7)	50 (8–14)	30 (1 >)			
Yang Ming (€)	15 (1–5)	30 (6 >)	5 (1–3)	10 (4 >)		

- The exact payable fee depends on the measurement of the number of days of containers spent in port, and in the hinterland. This data is not easily accessible. While time spent in port could be obtained from container terminals, data of time spent in the hinterland, and when exactly the container was stripped is generally unavailable.

Storm (2011) presents a first estimate of the total incurred demurrage and detention fees for cargo coming into the Port of Rotterdam of about € 25 mln, but with very wide boundaries. This amount would add € 12 per container. In practice, it is quite common to use the available free time to the maximum. See for instance Ypsilantis et al. (2014). As a result of this, we can take the average number of days for demurrage (4 days) and the average number of days for detention (4 days) and compare it to the minimum required time, which in many cases is only 1 or 2 day. The demurrage and detention mechanisms can therefore be said to cause a general delay in hinterland transport.

Container Release The container release contains three stages: customs release, carrier release and, finally, the terminal release. The last stage is a purely administrative process that relies on the releases of customs and the carrier. For the customs release, a follow up activity needs to be declared to customs. This can be full import, transit to another customs controlled location, or one of many other possibilities, such as temporary import, processing, etc. Customs can only be informed about this release if the container is really present in the terminal. This means that a freight forwarder who needs to lodge a follow up declaration spends considerable time in verifying the physical presence of the container, and any eventual inspection blocks on the container.

In addition, the freight forwarder has to communicate with the carrier about the release of the container. The carrier will demand payment for certain items on the transport bill, provide the empty depot location, and the PIN code for the container. The freight forwarder will record this information, make the payment, and arrange transport.

Together these activities do not take much time, but they need to be done at a specific time, that is triggered by the physical presence of the container in the terminal. This means that there is some time involved in polling the container terminal and the port community system's websites, to verify the presence of the container, and then time spent to fix any remaining release problems. The value of time for the freight forwarder in this process was estimated in RSM (2010) to be € 5–25 euro per container. Delay caused by this process is the time between the physical presence of the container and the first possible time to pick up the container. This period is determined by several factors: some terminals maintain a delay between the loading time and the reporting time of the presence of the container on their website. Many freight forwarders take as a proxy for the presence of the container the departure time of the ship. Given these factors, the real delay caused by the release process can be neglected.

Summary Below, we provide an overview of the consequences of ocean shipping transaction in terms of lead time, uncertainty and additional costs from the perspective of a logistics chain (Table 14.3).

From the table, it appears that the combination of ocean transport, formalities and release of the container generates considerable uncertainty, as well as additional costs for the buyer. The delays for the ship arrival and customs inspection processes can be added up and amount in total tot a maximum of 11.5 days[9]. The probability of this maximum delay is about 1 % (the probability of physical inspection), but many buyers build in slack to compensate for this eventuality. This slack is sometimes considerably more than the few days of delay we found above: we found a Dutch

---

[9] 1.5 days average ship arrival delay, a maximum average delay due to physical inspection of 4 days and 6 additional days for demurrage and detention.

Table 14.3 Overview of ocean transport transaction impacts on logistics

	Base duration and costs	Due date	Uncertainty	Additional costs
Ocean transport	Far East—Europe route is about 30–35 days; transport bill: US$ 2000	–	1.5 days delay	charges add US$ 1000 to transport bill per container
Formalities: ENS	Submission and instruction takes about 25 min	24 h before departure of the ship	Inspection decision 1–2 % of containers	US$ 25 per Bill of Lading, or € 9.50 per container
Formalities: SAL	Submission takes about 30 s per line on the declaration	72–24 h before ship arrival	Related to customs control: 2 day delay for scanning, 3–5 days delay for physical inspection	Scanning: € 215 per container Physical inspection: € 960 per container Freight forwarders' handling fee: € 10 per container
Container	–	Average free time for demurrage is 3–5 days, and for detention 2–6 days	Demurrage and detention causes about 6 additional days in hinterland transport	Penalties for demurrage and detention add about € 12 per container
Release	Verification of release status: several hours to several days	Demurrage free time	No additional delay, since this process is carried out in the demurrage free time	Costs for information gathering is € 5–25 per container

BAF bunker adjustment factor (fee for high fuel costs), *CAF* currency adjustment factor (compensation for currency losses of the carrier), *THC* terminal handing charges—the carrier's charge for the cargo handling costs in the terminal

retailer that calculated 1 week of slack for picking up containers in Rotterdam and bringing them to their warehouse, 1 week for possible deconsolidation, 1 week for customs inspection and release and 1 week for unloading and storage in the warehouse. Saving 1 day of time in the interval between arrival at a port of destination and arrival in the inland destination can be worth a lot for a buyer. In the Port of Rotterdam, with an annual amount of 1.8 mln containers entering the country, and assuming the average value of a container is about € 100,000 and a return on capital of 7 %, 1 day represents a working capital of € 33.8 mln. Especially for importers of high value containers, this kind of saving of working capital can be substantial.

In terms of costs, there are some major components:

- The basis transport bill; € 2000 basic transport costs and € 1000 charges per container,
- Costs related to scanning and inspection; € 215 and € 960 respectively,
- Various fees and penalties; € 46.5 per container.

The costs for scanning and inspection again relate only to the 2 or 1 % containers that are selected. All in all, and given the benchmark total transport bill of about € 3000 per container, in which the additional charges are included, the additional costs are not so substantial. It is clear that a physical inspection adds considerable costs, but this applies only to 1 % of containers. In addition, and contrary to the factor time discussed above, companies do not seem to anticipate these costs.

14.5 International Supervision

A final element in the discussion on trade facilitation is the supervision by customs. We have already considered the mechanisms of supervision—the ENS and SAL declarations and the administrative, scanning and physical controls -. In this section, we look into the quality of data that customs agencies have to work with if they base their risk assessment on the classic ocean transport documentation.

This line of thinking was triggered by several disastrous events in international ocean shipping in recent years. The most striking case is the fire on the MSC Flaminia, where cargo under deck exploded on 14 July 2012, and the grounding of the MSC Napoli on 18 January 2007. In the first case, apparently explosive goods were stored under deck, which is generally prohibited. The reason was that the goods were misdeclared on the manifest[10]. For the MSC Napoli, the cause for the grounding was severe listing of the ship, due to misleading of the containers. The accident report by the Marine Accident Investigation Branch (2008) states that of the 660 containers on deck, about 20 % has a declared weight on the manifest that differed from than 3 t from the actual weight. All in all, the total weight of these containers was 312 t more than the registered weight on the manifest, with one container weighing 20 t more than was recorded on the manifest. In addition, the report looked into the declared positions of the containers by the loading terminal, and their actual positions on the ship. 7 % of these containers were in the wrong place.

It follows from these incidents that the ocean carrier manifest is not a very accurate document, and that it may not generate the quality data that Customs would like to use for their risk analysis (Hesketh 2010). We also identify several steps in the ocean transport chain where information loss may occur:

1. The ENS declaration information needs to be provided in a very early stage (before commencing the loading operation of the ship), which means that in practice, some information—box count, number of consignments in the container, total weight of the container—may not be known yet,
2. The transfer from the bill of lading into the manifest may result in loss of information, since the manifest usually retains less information on goods descriptions,

---

[10] See for instance http://shippingnewsandviews.wordpress.com/2012/09/10/a-question-of-misdeclaration-msc-flaminia/.

3. The transfer from the manifest into the SAL declaration will result in a loss of information, because the goods descriptions on the manifest may not fit into the SAL declaration form.

Apart from apparent sloppiness of shippers and perhaps even wilful misdeclaration of goods to avoid some of the additional charges, there is also a legal reason for this lack of information. The Hague-Visby rules only require the shipper to provide the ocean carrier with basic data such as numbers of boxes, quantity and weight and a superficial description of the goods. The shipper is often hesitant to declare the full value (or provide a detailed description) of the goods to the carrier, because that will make the carrier liable for the entire cost in event of damage or loss (Hesketh, ibid.). The shipper usually covers the risk by taking out insurance himself.

Finally, we have seen in Sect. 3 that the role and function of the freight forwarders also results in loss of information, because they shield their clients from the ocean carrier by listing the freight forwarders on both sides as the consignor and consignee. This results in much interaction between Customs agencies and freight forwarders to gather information on the true consignor and consignee, because that will help Customs' assessment of the risk of that transaction.

The shortcomings of the ship manifest for supervision are beginning to be recognized in international trade circles. In fact, in a number of countries, the manifest is no longer the main or first document for risk assessment: Singapore and Hong Kong use an import/export licensing system, and the USA has replaced the manifest declaration by the $10+2$ regulation. European Customs supervision is, however, still firmly based on the manifest of the ocean carriers.

The consequences of the poor informational content of the ship manifest are not so easy to assess. First of all, Customs will demand additional information, which occurs for 15–20 % of incoming containers. This leads to an administrative burden for companies[11]. A more serious, and more difficult problem to quantify, is that Customs may target the wrong containers for scanning and physical inspection. There is, as far as we know, no public study on the two main errors: selecting containers for inspection that should not have been selected, and not selecting containers that should have been inspected.

14.6 Implications for Global Supply Chains

In this chapter, we provided an overview of international trade transaction, and we zoomed in on the ocean transport part of that transaction. We highlighted a number of ocean transport related processes that generate uncertainty and costs in logistics

---

[11] An off the cuff calculation of this burden is: about 1 mln incoming B/L's, of which 20 % requires additional information, that takes both the customs official and the freight forwarder about 30–45 min to process. Given an average hourly rate of € 30 for both the customs official and the freight forwarder, this amounts to an annual cost of about € 9 mln for the Port of Rotterdam alone.

chains: the use of the container in pre- and on-carriage, the release process of the container in port, and the formalities related to the supervision on vessels coming into ports in the country of destination. A first estimate of the uncertainties and costs reveals that the uncertainties that follow from these processes (delays, additional time required for supervision) far outweigh the additional costs. This holds especially because companies take into account a certain degree of uncertainty in their logistics planning, even though the probability is very low. We extend this discussion with the fundamental problem of the information quality of the ocean carrier documentation, the ship manifest, which has negative consequences for risk assessment by Customs in Europe.

We can conclude that specific processes that are connected to ocean transportation do result in time loss, uncertainty and, to a lesser extent, in additional costs that impact the efficiency of logistics and supply chains. The transaction costs are predominantly generated as additional charges on the ocean transport bill, as a result of supervision, and as a result of the use of the container. Transaction cost theory predicts that if such frictions exist, there will be a tendency to move from a market relationship to a more hierarchical relationship. We already observed that international transportation by sea is in an intermediate position between market and hierarchy. We would therefore expect to see some real hierarchical integration as a result of the frictions we identified.

This type of development can be observed in practice. A first case is the development of a new ocean carriage product by Maersk Line called “Inland CY”. This is a re-invention of the carrier haulage service, with the difference that the carrier does not deliver to the final destination, but to some intermediate location called the inland container yard. This product entails the development of a network of selected inland locations and connections between the seaport and these locations by barge and/or rail. With the choice of intermediate points, Maersk realizes the required level of consolidation of cargo to use barges and trains. This translates into CO2 savings. The direct connection to inland locations under the ocean B/L also means that a large number of the frictions related to the use of the container and the container release are avoided. This leads to a reduction of costs and delay. Finally, because of the limited number of inland destinations, which will also be an empty container depot, there are advantages for shippers in the availability of empty containers and the quick return of empty containers in the hinterland. Maersk develops this network together with a terminal and transport partner, BCTN. Therefore, this is not a full shift towards the hierarchy that TCE predicts, but it is a shift from the original situation towards a more exclusive partnership.

A similar development was pioneered by container terminal operator ECT in the Port of Rotterdam with the European Gateway Service[12]. This is a similar development of a network of terminals and connections that are prioritized in the seaport vis-à-vis other operators. The core of this network is formed by ECT owned terminals, but other terminal operators are gradually integrated in the network. The transport

---

[12] www.europeangatewayservices.com.

links are also based on exclusive partnerships. The interesting development is that in this case, the ocean terminal has taken the lead in this partial network integration, whereas it is the ocean carrier and the shipper who incur the greater burden as a result of the frictions. The reason for this is that the competitive environment in the Port of Rotterdam will change drastically in the near future, from a near monopolistic situation with ECT as the incumbent, to an oligopoly with three large terminal operators working alongside each other. ECT sees this network integration as a way to gain a competitive advantage. EGS thus also offers benefits in terms of avoiding frictions in the Port of Rotterdam for the container release and on-carriage.

These two initiatives focus largely on the operational frictions related to the containers. Frictions related to supervision are much more difficult to get rid of. One proposal that has been made is to develop a so-called extended gate of the seaport terminals in the sense that the customs release, but also all inspection processes, if applicable, could be deferred to an inland terminal. For inspection processes, this is currently not allowed. For the release, there is a solution that is based on the network operator providing the follow-up information for the containers directly to customs. The carrier can do this based on the B/L conditions, while the terminal operator must obtain instructions from their clients. Here there is a clear advantage for the ocean carrier. Further innovative use of European Customs regulation is being developed to entirely remove the customs related frictions. This may take several years, as well as some changes in the practical guidelines of the European Union Customs Code.

Given the considerable time-related frictions we found, another development that will facilitate trade is the actual reduction of the lead time of containers in the transport chain. This can be done in several ways:

1. Structurally reducing the delay of ships in port. Maersk Line has been doing this through a global program called Reliability. The result of this has been the service Daily Maersk. Maersk consistently scores high on the schedule reliability ranking of ocean liner companies.
2. Reducing the time containers spend in port. This can be done by improved communication on container availability and release status. Especially container availability information can help if the actual information on availability is pushed to interested parties in real time.
3. Improved information provision by customs on the release status of the container. Customs, in many ports, only actively communicates the inspection status. In addition, the release status could also be communicated in an early stage. This will avoid trucking companies attempting to pick up containers that are not fully released yet.

In conclusion, we have identified a number of frictions in the ocean transport transaction. These frictions currently create delays in logistics and supply chains, that translate into additional stock keeping and slack in logistics planning. Removing these frictions is the basis for several partial network integration initiatives that can be observed in practice in Europe. We expect that initiatives focused on reducing the time containers spend in port will provide a further contribution to trade facilitation of international ocean transport.

References

Cariou, P., & Wolff, F. C. (2006). An analysis of bunker adjustment factors and freight rates in the Europe/Far East market (2000–2004). *Maritime Economics & Logistics, 8*(2), 187–201.

Chung, C. C., & Chiang, C. H. (2011). The critical factors: An evaluation of schedule reliability in liner shipping. *International Journal of Operations Research*, *8*(4), 3–9.

Drewry Shipping Consultants. (2007–2009). Carrier performance insight. Quarterly report.

Hesketh, D. (2010). Weaknesses in the supply chain: Who packed the box. *World Customs Journal, 4*(2), 3–20.

Kumar, S. (2010). Logistics routing flexibility and lower freight costs through use of incoterms. *Transportation Journal,* 48–56.

Miller, T., & Matta, R. (2008). A global supply chain profit maximization and transfer pricing model. *Journal of Business Logistics, 29*(1), 175–199.

Miroudot, S., Lanz, R., & Ragoussis, A. (2009). Trade in intermediate goods and services. OECD Trade Policy Papers, No 93, OECD Publishing.

Murray, C., Holloway, D., & Timson-Hunt, D. (2012). Schmitthoff: The law and practice of international trade (12th ed.). Sweet & Maxwell.

NEA. (2010). The Netherlands as one logistics network in 2015 (in Dutch). Report commissioned by ABN AMRO Bank.

RSM. (2010). Synthesis of the group reports about ICT integration at freight forwarders in the port of rotterdam. Report commissioned by Syntens, in cooperation with Fenex. Nederland Digitaal in Verbinding, 23 November 2010.

Storm, R. (2011). Controlling container demurrage and detention through information sharing. Master thesis, Rotterdam School of Management, December 2011.

Veenstra, A. W., de Putter, J., Warbout, M., & Peters, P. (2013). Reduction of regulatory pressure in the logistics section for import and export (In Dutch). TNO report R11072 20130936, commissioned by ACTAL.

Vernimmen, B., Dullaert, W., & Engelen, S. (2007). Schedule unreliability in liner shipping: Origins and consequences for the hinterland supply chain. *Maritime Economics & Logistics, 9*(3), 193–213.

Williamson, O. E. (1979). Transaction-cost economics: The governance of contractual relations. *Journal of Law & Economics, 22,* 233.

Williamson, O. E. (1981). The economics of organization: The transaction cost approach. *American Journal of Sociology,* 548–577.

Ypsilantis, P., Zuidwijk, R., & van Dalen, J. (2014). Dwell time of containers in container terminals: The shippers' effect. Paper presented at the ILS 2014 conference, Hamburg.

Chapter 15
Modelling Global Container Freight Transport Demand

Lóránt A. Tavasszy, Olga Ivanova and Ronald Aprilyanto Halim

Abstract The objective of this chapter is to discuss methods and techniques for a quantitative and descriptive analysis of future container transport demand at a global level. Information on future container transport flows is useful for various purposes. It is instrumental for the assessment of returns of investments in network infrastructure and fleets, the prediction of environmental impacts of transport and the analysis of success of governmental policies about maritime markets and hinterland transport systems. As the future development of global freight flows is unknown and quite uncertain, models are used to define plausible and consistent scenarios of the future performance of the sector.

Models of global container transport demand can follow the generic architecture available for freight transport modelling. We describe the methods and techniques available by reviewing the literature with a specific focus on global level freight modelling and treat the subject in two main parts. One part involves modelling the demand for movement between regions, i.e. the outcome of the processes of production, consumption and trade. The second part involves the modelling of demand for transport services by mode and route of transport, including the demand for maritime and inland port services. In both parts we find that surprisingly little research has been conducted specifically for descriptive models of global container movements.

Future work can focus on the linkages between container transport and supply chain management. This may include a better understanding of the contribution of shippers' preferences to observed shipping choices. Also, future developments in geographic restructuring of supply chains because of changes in manufacturing locations or distribution structures, could be looked into. Finally, as global, integrative models do not yet exist, combining new trade and transport network models in a consistent way should provide new tools for long term forecasting and policy analysis.

L. A. Tavasszy (✉) · O. Ivanova · R. A. Halim
TNO and Delft University of Technology, Deft, The Netherlands
e-mail: lori.tavasszy@tno.nl

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics*,
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_15

15.1 Introduction

Container flows have been growing fast during the past decades. The development of trade of the last decades shows that it is increasingly decoupled from the economy. Container transshipment growth has by far outrun trade growth, with annual growth rates between 5 % and 15 % during the years before and after the recent economic crisis (Fig. 15.1). This double decoupling indicates that there are many underlying mechanisms at work by which, eventually, container transport demand can be predicted.

The objective of this chapter is to discuss methods and techniques for a quantitative and descriptive analysis of future container transport demand. Information on future container transport flows is useful for various purposes. It is instrumental for the assessment of returns of investments in network infrastructure and fleets, the prediction of environmental impacts of transport and the analysis of success of governmental policies about maritime markets and hinterland transport systems. As the future development of global freight flows is unknown and quite uncertain, models are used to define plausible and consistent scenarios of the future performance of the sector. Whereas normative models are geared towards optimization of parts of the freight transport system, descriptive models aim to project a future state of the system which is in line with observable behavioral patterns of the industry and in a broader sense, the global economy.

Freight transport demand is defined here as the aggregate result of complex and interrelated choices of stakeholders in the maritime supply chain, leading to a need for transport of containers door-to-door between a shipper (who obtains and stuffs a container) and a receiver (who strips the container at its arrival, returning it to its

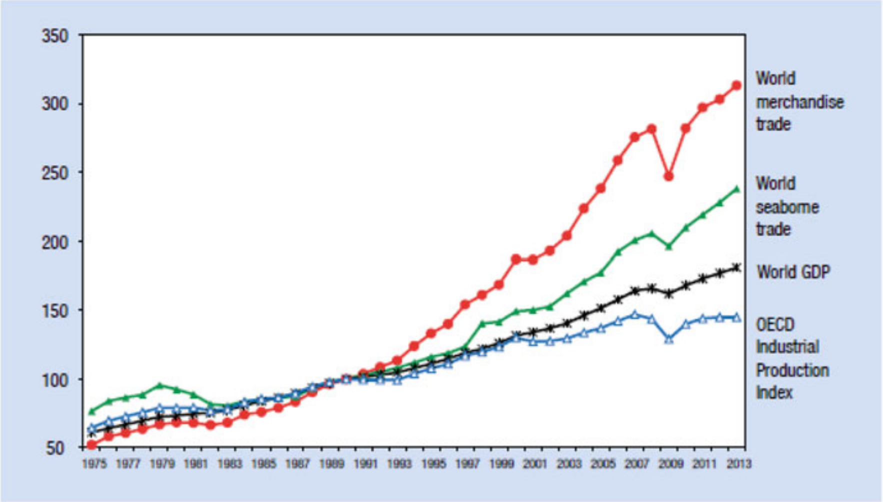


Fig. 15.1 Decoupling of trade and GDP. (1990 = 100; Source: UN Review of Maritime Transport 2013)

owner). Key agents in this set of choices are the shipper and the forwarder. Typically, demand from our perspective is measured in Twenty Foot Equivalent container Units (TEU's) per year between regions of the world, along a certain route and by certain modes of transport. In contrast to demand, the supply of transport services is provided by maritime and inland carriers and ports. They anticipate on and respond to demand by offering services of a certain quality and a certain price between origins and destinations of freight. In this chapter we do not deal with the formation of maritime networks and services. Instead we limit ourselves to modelling demand for transport.

Container freight demand is strongly dynamic as it is derived from developments in the global economic, logistical and technological environment. Over the long term, container demand has grown steadily and has been led by many different influences in, amongst others (Tavasszy et al. 2003, Rodrigue 2006):

- Continuing population increase and economic growth in various countries of the world,
- Increased specialization of production and mass-individualization of consumption
- Globalization of production networks, due to a decrease of barriers for trade and transport,
- Increasing modularization and associated scale economies in shipping
- Formation of new and complex network structures, creating hubs and corridors.

Current methodologies for forecasting freight transport demand differ with respect to their overall goal. Most of the existing modelling tools are designed for a what-if type of analysis that compares the business as usual situation with some alternative scenario. Examples of the type of questions that can be answered by these tools include the effects of transport costs changes, removing or reduction of monetary and non-monetary trade barriers as well as relocation of production capacities on freight transport flows. At the basis of these analyses lies the exploration of freight transport flows in the long-term future, given economic development. Prediction of such long-term developments requires the knowledge of how world-wide specialization and production patterns are going to develop in the future, usually available in qualitative scenarios. These qualitative scenarios are created for different visions of the world in the future and may vary a lot with respect to their outcomes. We refer to MIT (2011), DHL (2012) and Mazzarino (2012) for generic scenarios on freight transport and to Lloyd's register et al. (2013) for global maritime container transportoriented explorations.

Recent scenario studies demonstrate that it is not only regional and sectoral economic growth, that will govern the future volumes of container flows. Besides the usual basis of GDP predictions, assumptions about the geographic characteristics of trade and logistics are equally important. We provide one example of a scenario study that illustrates this point, around long term forecasts for container transshipment at the Port of Rotterdam. This is a relatively well-documented forecast, both in terms of the results (Port Authority Rotterdam 2011) as well as the methodology followed (De Langen et al. 2012). The official forecast shows a range of throughput between 21 and 41 mln. TEU (numbers rounded to two digits for convenience) in the year 2040, with present numbers around 12 mln. TEU per year. Despite the intensive, expert supported process that was followed for these forecasts, the scenarios are strongly

Table 15.1 Two groups of scenario's for the Port of Rotterdam in 2040

Source	Scenario	Throughput 2040 (MTEU)
Port Authority forecast Varying GDP growth (Port of Rotterdam 2011)	Low Growth	21
	European Trend	34
	Global Economy	41
Alternative forecasts Equal GDP growth (Van Diepen et al. 2012)	Local production	10
	Regional orientation	16
	Global scale	27

MTEU = Millions of TEU's, where *TEU* = Twenty foot Equivalent Unit, a standard unit for measuring volumes of containers.

growth-driven, with a clear hierarchy from low to high GDP growth. Around that time an alternative forecast was presented by Van Diepen et al. (2012) which focused on uncertainties in the geography of trade, logistics and transport. More importantly, for demonstration purposes, it kept GDP growth equal between scenarios. The key variable factors for the scenarios included location of manufacturing, export orientation, speed of mass-individualization of consumption, global transport infrastructure and transport pricing; leading to the following three alternative futures:

- Global scale: Continuing trend of globalization. 'The world is flat', with global factories serving a global market.
- Regional orientation: Emphasis on trade within continents, with clusters emerging per region or continent.
- Local production: Local supply chains with trade occurring predominantly within countries and between neighbors (in extremis: 3D printing). Transport over longer distances is dominated by raw materials

The scenarios of Van Diepen, despite sharing one and the same level of GDP growth, still showed as much variation between alternative futures (roughly a factor 2), as the Port Authority's own scenarios. This is interesting, as it implies that container transport demand is influenced by many different factors (especially in critical combinations) than just growth. Table 15.1 shows the two contrasting sets of scenarios for throughput of the port of Rotterdam in 2040.

Two points deserve an additional comment. Firstly, the studies are reasonably consistent, in the sense that the GDP growth assumed in Van Diepen et al. (2012) is comparable to the European Trend scenario. The difference of 7 MTEU is significant, but can be attributed to specific differences in assumptions concerning, amongst others, the development of Eurasian rail transport and the increase in service quality of Mediterranean ports. In this sense, the alternative scenarios can be interpreted as "variations around the mean" of the Port's scenarios, albeit with a downward bias. Secondly, we note that it is not only the total volumes of freight that differed between these studies. The composition of the flows in terms of commodity groups, and the modes of transport in the hinterland of Rotterdam turned out to differ substantially.

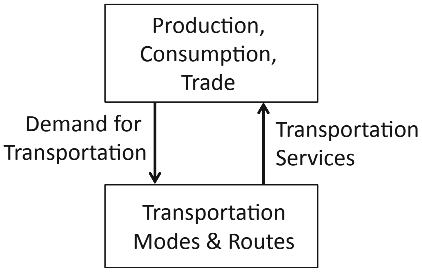


Fig. 15.2 Complementary models of the transport system

In other words, independent of total volumes, there could be major changes in the added value of the port's activities and the use of the hinterland infrastructure.

This example shows that the relevant issues for future container transport have to be derived from a complex system of assumptions, that goes far beyond worldwide growth alone.

Further in this chapter, we discuss a number of models that allow comprehensive scenarios of global trade and container transport to be developed and quantified. We review the state of the art in demand modeling for global container flows, from the perspective of the providers *and* the users of transport services. It is the users who decide upon the creation of trade relations, the logistical organization and transport contracts with service providers. Complementarily to recent foresight oriented texts on containerization (e.g. Notteboom and Rodrigue 2009) we take this demand oriented perspective as it allows more systematic scenario building. We discuss models that operationalize the derived nature of freight demand, within the limits allowed by publicly available data. In addition, we show how this can be translated into transportation oriented models of the global container network.

Broadly speaking, the models needed to predict future demand divide into two categories:

- Spatial models of regional growth and interregional trade, and
- Models that predict transportation flows by specific modes and routes.

The two types of models are complementary, as the results of the first will feed into the second, as a spatial demand for movement (Fig. 15.2).

In between these two, a number of translations of flow formats is necessary to link the very different accounting systems of these model types. As we will discuss further in the chapter, also a new type of model is emerging at their intersection that describes the spatial organization of global and continental supply chains. We include a brief discussion of these models as well.

The chapter is built up as follows. Section 2 describes the spatial price equilibrium models and its variations, used for prediction of trade patterns. Section 3 describes the steps needed to arrive at flows specific by mode and route of transport (i.e. including the maritime and hinterland ports of transshipment). Section 4 describes new directions of development that aim to improve the bridge between the two main components, by including the market of logistics service providers. Section 5 concludes our chapter with a summary of the modeling steps and ideas for future work.

15.2 Global Trade and Economic Development

The fact that freight transport has become more efficient and cheaper has contributed to integration of not only goods but also capital market by increasing the amount and geographical coverage of Foreign Direct Investments (FDI). International freight flows are determined by an increasingly dynamic spatial distribution of production and consumption of various goods. The patterns of international freight flows follow closely the patterns of international trade flows. Some large freight transit countries such as Singapore, Panama and the Netherlands represent an exception from this rule. Non-monetary barriers such as cultural differences, difficulty of administrative procedures at the countries' borders have been proven to be one of the most important factors that currently make free trade movements more difficult. All these factors are important for the prediction of international trade and international freight transport flows.

In this section we review models of global trade. We pay attention to two dimensions that are particularly relevant for modelling future container transport demand, and where the current literature is limited. Firstly, we see that the trade literature usually considers transport costs as exogenous variables or approximates them by the geographical distance between countries. This assumption is usually made in the absence of detailed data on the transportation costs. In addition, much of the present trade-related research focuses on trade factors other than transport infrastructure and logistics. These include monetary and non-monetary trade barriers (e.g. custom procedures, level of corruption), differences in history, culture and institutions between the countries. Nevertheless, the question which attribute of the transport network to use as resistance factor, remains important. Costs of transport are broadly known not be linear with distance and to include other factors than out-of-pocket transport costs, such as the value of time lost in transport. In models of global trade and container transport demand, it is important to provide a rich framework that accounts for the effect of as many relevant factors as possible. The system of interest can be pictured through the following figure, indicating consistency in quantities and prices within and between spatially separated product markets[1] (Fig. 15.3):

A second aspect concerns heterogeneity. Much of the empirical literature on trade modelling ignores differences between various commodities. An aggregated analysis

---

[1] Note that this entails a simple, static view on the system. Due to several reasons (imperfect knowledge and anticipating capability, non-zero response times, delayed responses due to inertia, etcetera) spatial equilibrium will probably never be reached. Nevertheless, this model provides an explanation of freight transport and its relation to the global economy that is tractable with currently available aggregate data (see e.g. Harker 1985 for an early formulation).

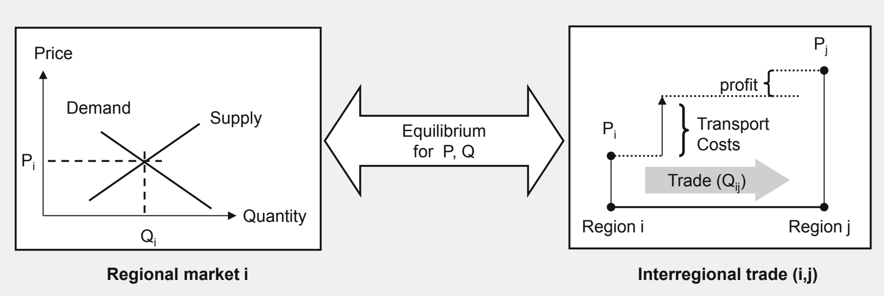


Fig. 15.3 Spatial price equilibrium

does not allow one to look at particular commodities and diminishes the usefulness of the econometric approach for the purposes of predicting freight transport flows. The focus of researchers on the aggregated trade flows is explained by the large heterogeneity in the results of econometric estimates on the detailed trade data by type of commodity. However, especially with a view to the predictive purpose of models, it is important to recognize the importance of heterogeneity, as shown by the growing individualization of consumption, increasing industrialization and dematerialization of economies. Therefore, models that allow insight into not just supply chains but into global networks of sectors and, in addition, are able to predict how effects of investments or policies will propagate through these sectors, are preferable over those that assume rigid structures.

During the past decades, various methodological approaches have been developed for forecasting and analysis. The most important ones for container transport flow forecasting include gravity models, disaggregate models of trade, Input-Output (IO) models and Spatial Computable General Equilibrium (SCGE) models[2]. We discuss these below.

The first and most widely used model of spatial interaction is the gravity model. According to the gravity model the trade flows between two countries or regions is the function of the level of economic activity (usually approximated by Gross Domestic Product or sectoral output) and the geographical distance between these countries, as a measure of trade resistance. The latter has a negative impact upon the flow of goods and can include both monetary costs such as transportation and logistics costs, costs of capital as well as non-monetary costs such as the costs of corruption, costs related to complicated administrative procedures and language and cultural barriers. This administrative dimension is especially important for the developing countries and the process required to reduce these costs is called 'trade facilitation'. Reduction of either monetary or non-monetary costs of trade leads to increase in the flow of

---

[2] A relatively new approach is the application of Systems Dynamics modelling, following the Club of Rome's world model. We do not discuss this stream of work here as we limit ourselves to models which allow a detailed spatial analysis at global scale and are based on widely accepted economic theories.

trade between two countries or regions. Gravity models of international trade have been widely used in the last decades to analyze the effects of changes in monetary and non-monetary trade barriers on trade flows.

In the majority of trade literature transport costs have been considered as exogenous variables and approximated by the geographical distance between the geographical locations. This assumption has been largely justified by the absence of reliable data on the transportation costs as well as the focus of the present trade-related research on trade factors other than transport infrastructure and logistics.

Most of the existing empirical literature on the gravity model focuses on aggregated trade flows data and ignores the differences between various commodities and region-pairs in terms of the impact of transport costs on trade patterns and volumes. This aggregated approach does not allow for making policy-relevant conclusions related to particular industries and commodities and diminishes the usefulness of the econometric analysis.

The focus of researchers on the aggregated trade flows might be explained by the large heterogeneity in the results of econometric estimates on the detailed trade data by commodity type in combination with a large share of non-significant parameters in case of regressions for detailed commodity groups. Another challenge is the presence of a significant number of zero trade flows in the detailed data which requires the use of more sophisticated econometric estimation techniques. Instead of using ordinary least squares (OLS) for the estimation a researcher should then make use of more advanced estimation models (Poisson, Tobit or Heckman's two stage estimation models) in order to properly estimate the gravity model in the presence of zero trade flows.

Processes of spatial distribution can also be modelled at the disaggregate level, of the individual decision maker. Micro-economic theory also provides a representation of individual behavior of consumers and producers. A model widely used for prediction of flows (mostly for passenger flows) is the discrete choice model. The foundation of the discrete choice models have been laid down by Daniel McFadden. These models describe and predict individual choices between two and more discrete alternatives. In a spatial trade context, discrete choice models can be used to represent the behavior of the trader who decides upon the origin of the goods that he wants to purchase in order to satisfy the demand of his customers in another destination region, where prices are higher. His choice is determined by the prices of goods in various locations, the costs of transporting these goods between two locations and other observed and unobserved characteristics of the county or the region of origin. These other characteristics can include the various types of location attributes or (non-) monetary trade costs. Discrete choice models estimate the probability that an individual trader will choose to purchase goods from a particular location and sell it at another location. Disaggregate models of trade are rare. Only recently, research has emerged into such models, emphasizing the importance of disaggregate analysis (Anderson and Yoto 2010). To our knowledge, however, there are no disaggregate, empirical models of global trade, probably because of the prohibitive data needs of such models (one notable exception being Srinivasan and Archana (2009) which is, however, limited to Indian manufacturers).

There is an interesting linkage between discrete choice and gravity modelling at the aggregate level that deserves to be mentioned. Erlander and Stewart (1990) provide an elegant reasoning why the aggregate logit model provides a theoretical explanation of the gravity model, effectively linking the two frameworks. We repeat the main components of their proof below.

If the utility of trade U_{ij} between regions i and j with product prices P_i and P_j respectively and generalized transport costs C_{ij} can be defined as:

$$U_{ij} = P_j - P_i - C_{ij} \tag{15.1}$$

And the probability of trading between i and j can be denoted according to the logit model as

$$Pr\{ij\} = \frac{\exp(U_{ij})}{\sum_{ij} U_{ij}} \tag{15.2}$$

And the volume of trade T_{ij} on relations (i, j) is given by multiplying the probabilities by the total flow T in the system of all (i, j),

it can be shown that the following gravity model equation holds

$$T_{ij} = O_{ij} D_{ij} exp(-C_{ij}) \tag{15.3}$$

Where O_i and D_j are the region-specific constants that can be estimated econometrically and exp(-Cij) represents the deterrence function. In summary, the simple aggregate logit model of trade and the gravity model are mathematically equivalent. A more formal derivation from micro-economics can be found in Bergstrand et al. (2013). Recent applications of this model to global trade can be found in De Benedictis and Tajoli (2011) and Hausman et al. (2013).

Still, this aggregate model does not describe the interactions between sectors of industry of the world. The models used for this purpose include models of intersectoral exchanges or Input-Output (IO) models, combined forms of gravity and IO approaches (multiregional IO or MRIO) and integrated approaches: the Spatial Computable General Equilibrium (SCGE) models. As with the gravity model, these models include an explicit or implicit representation of micro-economic decisions of households and firms, as well as a representation of equilibrium on the markets for goods and factors of production that results in the equilibrium prices.

The IO model describes the monetary flows of commodities and services between the economic agents in the following way:

$$X_m = \sum_n X_{mn} + Y_m \tag{15.4}$$

Where $X_m$is the total output of sector m that is defined as the sum of intermediate uses of its products by all sectors in the economy x_{mn}plus the final consumption of the households, Y_m. The direct purchases of the sectors can be expressed via the use of Leontief technical coefficients a_{mn} as

$$X_m = a_{mn} X_n \tag{15.5}$$

The initial equation can be than rewritten in a classical way as follows:

$$X_m = \sum\nolimits_n a_{mn} X_n + Y_m \tag{15.6}$$

The work on the multi-regional IO modelling was started by Chenery, Isard and Moses in the 1950s. They have introduced inter-regional trade coefficients that allow for calculation of inter-regional trade but did not specify their functional forms. The standard multi-regional IO models have fixed inter-regional trade coefficients and hence do not reflect how the changes in regional prices and inter-regional trade and transport costs affect trade patterns. Newer types of multi-regional IO models include a representation of variable trade coefficients such that the inter-regional trade flows respond to the changes in monetary and non-monetary trade barriers and freight transport costs.

One can make use of the discrete choice model in order to capture the behavior of the sectors with respect to the choice of the origin for their intermediate inputs in a cost minimizing way. The discrete choice models such as logit and nested logit can be estimated econometrically and integrated within the general structure of the multi-regional IO models, to arrive at the multi-regional Input-Output models with variable trade coefficients for forecasting of freight transport flows. The development of the discrete choice theory of McFadden has resulted in the development of the specific functional forms for the trade coefficients that are based on the random utility functions. As seen above, one can use the gravity model to account for the dependence of trade flows on transport costs.

Despite the fact that the multi-regional IO models with variable trade coefficients take into account the impact of transport costs and trade barriers on trade flows, they do not account for changes in prices and incomes as opposed to the Spatial Computable General Equilibrium (SCGE) models. Computable general equilibrium models (or CGE models) offer a framework to take into account the full details of economic interactions between producers and consumers on various markets. Here, prices and quantities of the goods exchanged are kept consistent between the two sub-systems of production/consumption and trade. The SCGE model is a micro-economics based macro-model, since it includes both the representation of microeconomic behavior of economic agents such as utility maximization and profit maximization as well as the representation of equilibrium on all markets in the economy. The theoretical foundations of SCGE models go back to the theory of Walras. The core of this theory is the market equilibrium where the "invisible hand" acts an auctioneer and keeps demand equal to supply on all the markets in the economy. Below we provide a verbal outline of the main principles of these models and discuss some empirical models. For a summary of the mathematical SCGE model specification and a detailed discussion of the empirical applications we refer to Ivanova (2014).

From the firm perspective, production and trade decisions lie at the basis of the SCGE model. These are modeled using the Constant Elasticity of Substitution (CES) function[3]. According to the CES function, the elasticity of substitution measures the percentage change in factor proportions due to a percentage change in the marginal rate of technical substitution. The CES function has nice mathematical properties which make it attractive to include in complex modelling tools such as CGE models. The work of Paul Krugman on the development of the New Economic Geography (NEG) theory provides the basis for explanation of the mechanisms of agglomeration and dispersion forces and the role of freight transportation costs in formation of spatial patterns of production and consumption. His work on the foundations of New Trade Theory (NTT) has laid the basis for the explanation of international trade patters as the result of product differentiation and agglomeration forces in a model of monopolistic competition. The CES function is used to capture these phenomena and has become the workhorse of the NEG and NTT theories. CGE models include complex, non-linear mathematics. This allows them to model (dis)economies of scale, external economies of spatial clusters of activity, continuous substitution between capital, labour, energy and material inputs in the case of firms, and between different consumption goods in the case of households.

The design of a SCGE model requires several steps. First, the structure of the general model is determined. Then, a particular functional form has to be chosen for the production and demand functions. Apart from CES, sometimes also Cobb-Douglas and Linear Expenditure System (LES) specifications are selected. Finally, the parameter values for the functional forms must be derived. Ideally, all the parameters in the SCGE model may be econometrically estimated, using simultaneous equation estimation methods that take into account the overall model structure. However, given the size of SCGE models, the required sophistication of techniques, the identification problems and the lack of data, this procedure is considered infeasible (Gunning and Keyzer 1995). Therefore, the most commonly used procedure to determine the parameter values is calibration (Mansur et al. 1981). The calibration procedure implies that the parameters of the model are identified on the basis of a single observation of the economy. The economy under consideration is assumed to be in equilibrium, a so-called reference equilibrium or benchmark equilibrium. In practice, the benchmark equilibrium or benchmark data set is a Social Accounting Matrix, constructed from national accounts or other governmental data sources. The calibration procedure ensures that the parameters of the model are specified in such a way that the model will reproduce the initial data set as an equilibrium solution. Once the parameters are calibrated, the model is complete and different policy changes can be simulated. The parameter values are crucial in determining the results of the policy simulations.

During the last decades, several operational SCGE models have been developed for the analysis of questions related to international or interregional trade. Some examples of well-known SCGE models with representation of regions of the world and disaggregation on the level of regions include EXIOMOD (a global version of

---

[3] From a micro-economic theoretical perspective, the CES function is closely related to the logit model. Anderson et al. (1987) demonstrate that it is a special case of the logit model.

RAEM, see Ivanova et al. 2007), GEM-E-3 (Capros et al. 1997) and the Global Trade Analysis Project (GTAP) model (Walmsley et al. 2012).

The applications of SCGE models include regional infrastructure investments, transport policy related issues such as road taxes or road charging and regional planning (housing, labour market policy, land use). Although the application of these models is gaining ground, their complexity creates an additional burden for the tasks of data analysis, estimation and communication of results. These are seen as important drawbacks to perform policy analysis using modelling. At the same time, models that provide only a partial view of world trade, like the gravity or the IO model, or combined forms of these, are known for their structural limitations and may provide outcomes that are inconsistent with theory. The researcher will have to weigh the pros and cons of these methods to provide a fitting solution to each individual modelling challenge.

In the next section, we turn towards the modelling of impacts of trade on transport flows. Before being able to assign flows to networks, we need to change the units of calculation from those typical for trade analysis to those of use for transport calculations. These transformations include:

- From sectoral to commodity group classifications
- From monetary value (typically, dollars) to weight units (metric tonnes)
- From weight units to loading units (containers or TEU's)

As the flows in tons need to be converted in containers, transport statistics such as the Eurostat or OECD database can be used. The area coverage of Eurostat is of course less complete than the international trade database, as it covers only transports between Europe and the rest of the world, but information on the exchanges is available in tons, unitized tons and TEU's. This database allows to calculate *ratios of unitization* (total cargo tons potentially unitized divided by the total cargo tons) and *ratios of density* (unitized tons divided by TEU). These two ratios can be calculated per type of goods and OD pair on the basis of transport statistics. The ratios are applied directly to the trade database. In case some ratios are lacking, the averaged ratio per OD and/or type of goods can be used. In order to include empty containers in the input data, databases are available which make it possible to extract the number of empty and full TEU for each OD-pair. A return percentage of empty containers can thus be obtained. As before, the percentage can be applied per OD-pair and type of goods.

We note that these transformation ratios are not necessarily constant over time. They may change in the future as a result of changes in the industry (dematerialization as a result of the introduction of product-service systems) or even the organization of transport flows (reduction of empty container flows as a result of co-operation between carriers).

15.3 Use of Transport Services

Transport is commissioned by shippers or forwarders to carriers, who plan and execute transport activities, either for the entire transport chain or for parts of it, be it the movement by maritime or inland modes transport, or the transfer amongst

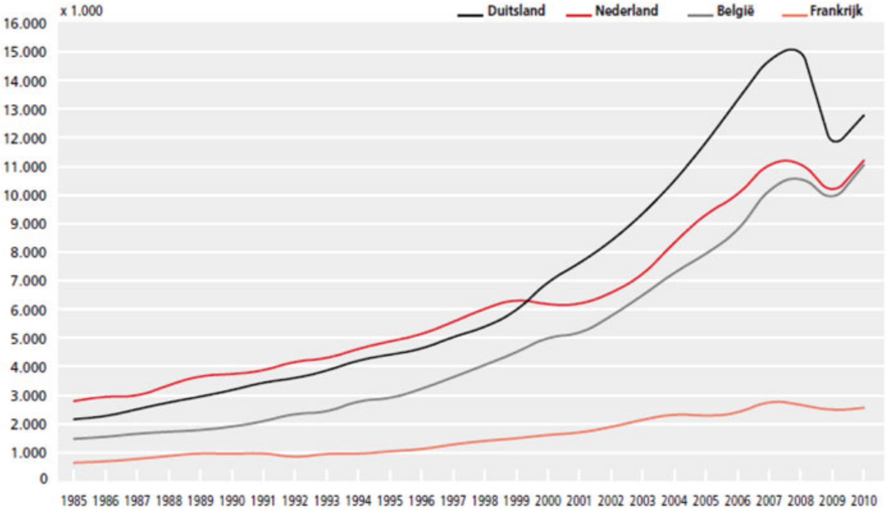


Fig. 15.4 Container transshipment per country (TEU). (Source: Port of Rotterdam)

these modes. The second stage in transport demand modelling is to determine the sequence of activities within the transport chain, i.e. establish by which mode and route of transport the flow between first origin and final destination of the shipment will be carried out. In global container networks, these choices involve the following:

- Route choice includes the path followed by the landside modes and locations of transfer between modes. This includes the choice of maritime ports and dryports in the hinterland.
- Maritime and landside modes of transport are complementary, but may also compete in the case of land bridges (e.g. the Trans-Siberian corridor). In many port hinterlands, there is a clear competition between modes of transport.

An interesting case that illustrates the need to develop a deeper understanding of freight routing in connection with spatial and economic development includes the shift in total container transshipment of neighboring countries in Europe. Next to the well-known fact of Rotterdam being the largest container port in Europe, perhaps less well known is that already since 2000, Germany has overtaken the Netherlands as largest transshipping country in Europe (Fig. 15.4).

One important reason for this, as we expect, is that the growth of the hinterland in the past decade has not followed the existing geographical pattern but has moved eastward, following the development of Central and Eastern European countries.

Two approaches for network choice can be distinguished in the literature: a separate treatment of mode and route choice and an integrated one using supernetworks. We discuss these below.

Firstly, mode and route choice can be modelled separately. For these cases, there is a rich literature on mode choice available. This literature, however, almost exclusively deals with models for inland transport of containers with surface transport modes. Notable exceptions include Lewis (1994) who develops detailed logistics cost functions and Tsuboi et al. (2010), Kato et al. (2012) and Yang et al. (2013) who model the choice of mode itself. Typically, the models are based on generalized costs (include transport costs and time). Insofar as mode choice is treated in the hinterland, models specifically for container transport are rare (see Feo-Valero et al. 2011 for a recent review). Little research is available on descriptive models that discuss competition with sea transport. Brooks et al. 2011 study competition between land and short-sea mode. Using disaggregate random utility, discrete choice (logit or probit) choice modelling. Hummels and Schaur (2012) specify and develop an empirical model based on US imports data that considers air and sea transport, at a very detailed level of disaggregation. An interesting relation between the mode choice and the trade literature lies in an explanatory factor behind mode choice decisions: the value of transport time. In general, the more valuable goods are, the more likely they are to be shipped by air transport. For the categories of freight where the choice of sea or air is not trivial (Kato et al. 2012), the value of time in the supply chain is an important explanatory variable for both the geographical patterns of trade and the choice of mode. This value of time parameter can be specified as linearly dependent on value density of goods. As such it allows differentiation between goods types.

As a second approach, the choice for modes and routes can be modelled together in a connected multimodal network, also called a supernetwork (Fig. 15.5). Again, there have been various models for freight transport in supernetworks (see e.g. Jourquin and Beuthe 1996; Tavasszy et al. 1998; de Jong and Ben-Akiva 2007; Yamada et al. 2011; Zhang et al. 2013), but none of these authors have treated global container transport flows. The essential difference with the conventional mode/route choice supernetwork model is the addition of port choice. In the remainder of this section we discuss an approach that combines the choice of mode and route, including port choice, into one model for global container transport flows.

Academic literature on port choice identifies a multitude of service-related and cost factors that influence the decisions made by shipping lines and shippers. These factors relate primarily to port infrastructure, the accessibility over land and via the sea, the geographical location vis-à-vis the immediate and extended hinterland and the main shipping lanes (centrality and intermediacy), port efficiency, port connectivity, reliability, capacity, frequency and costs of inland transport services, quality and costs of auxiliary services (such as pilotage, towage, customs, etc.), efficiency and costs of port management and administration (e.g. port dues), the availability, quality and costs of logistic value-added activities (e.g. warehousing) and the availability, quality and costs of port community systems, see e.g. Murphy et al. (1992); Murphy and Daley (1994); Malchow and Kanafani (2001); Tiwari et al. (2003); Nir et al. (2003); Song and Yeo (2004); Lirn et al. (2004); Guy and Urli (2006) and Yeo et al. (2013). Most quantitative research on port selection (see e.g. Chou (2009); Tongzon (2009); De Langen (2007); Malchow and Kanafani (2004) uses disaggregate behavioral analysis, which limits the geographical scope of application of models, due to the high

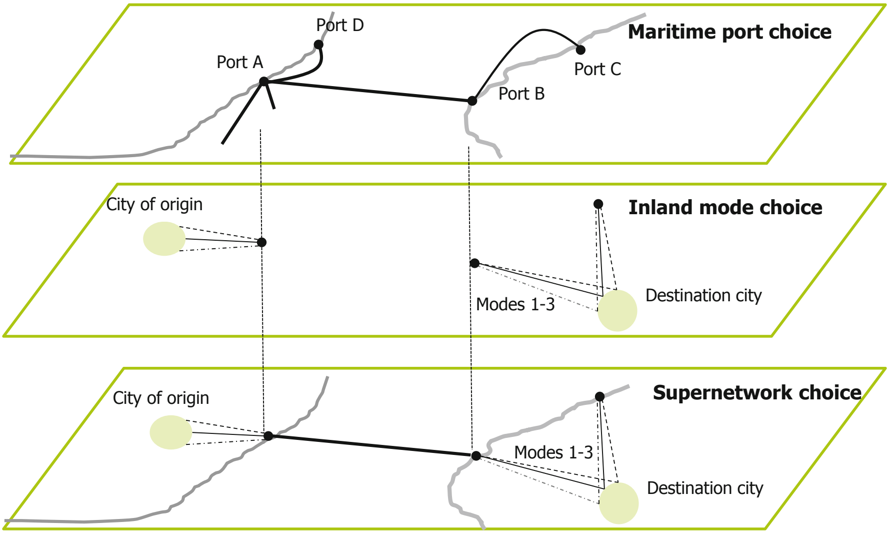


Fig. 15.5 Supernetwork choice in the context of global container transport demand

costs of data acquisition involved. Aggregate models could in principle be applied at a global level, if the specification of the model is such that data needs do not become prohibitive. The transportation literature presents a limited number of aggregate models for the routing of seaborne freight, e.g. Tang et al. (2008); Giannopoulos et al. (2007); Jula and Leachman (2011); Frémont (2007); Veldman and Bückmann, (2003); Zondag et al. (2010); Tavasszy et al. (2011) and Meng et al. (2013). Most of these include explicitly the role of hinterland modes and routes. Nevertheless, they all have a regional basis. Only Tavasszy et al. (2011) develop an empirically tested global scale assignment model for all containerized world trade flows. We summarize this model below and refer the interested readerto the original paper for details.

We assume a network with hinterland transport networks of different modes and maritime shipping lines that provide door-to-door connectivity between world regions. The performance of this comprehensive network is measured in transport costs and transport times. Jointly, these lines create a set of alternative routes between origins and destinations, that include direct lines where possible as well as indirect routes that include transshipment, using multiple ports and transport lines. The route choice model describes the selection of optimal routes, from this set of alternatives, between origin and destination regions (i.e. at the level of O/D pairs). Furthermore, we assume that route choices are made by profit maximizing shippers who have knowledge of the main routing alternatives over land and sea, for goods traded between two countries.

The basis for the model is an aggregate logit route choice model, where

- flow volumes are calculated by assigning yearly O/D volumes (in tons or TEU/year) to the alternative routes by their calculated probabilities,
- alternative routes include land and sea modes as well as ports of transshipment,

- routes are enumerated in advance using an enumeration algorithm (k-shortest path),
- possible route overlap is taken into account in the cost function (path size logit),
- choice probabilities depend on the route specific generalized costs, and
- generalized costs depend on distance, tariff, transport time and value of time.

The choice probability of a route from a set of alternative routes (the choice set) is calculated as:

$$Pr\{r\} = \frac{\exp\left(-\mu(C_r + lnS_r)\right.}{\sum_{h\in CS}\exp\left(-\mu(C_h + lnS_h)\right.} \tag{15.7}$$

where:

$Pr\{r\}$	the choice probability of route r
C	generalized costs
CS	the choice set
h	path indicator
μ	logit scale parameter

With the path size overlap variable defined as

$$S_r = \sum_{a\in\Gamma_r}\left(\frac{Z_a}{Z_r}\right)\frac{1}{N_{ah}} \tag{15.8}$$

where:

a	link in route r
$\mathrm{S_r}$	degree of path overlap
Γ_r	set of links in route r
z_a	length of link a
$\mathrm{z_r}$	length of route r
N_{ah}	number of times link a is found in alternative routes

The generalized cost function is given by the following equation:

$$C_r = \sum_{p\in\Gamma_r} A_p + \sum_{l\in\Gamma_r} C_l + \propto \left(\sum_{p\in\Gamma_r} T_p + \sum_{l\in\Gamma_r} t_l\right) \tag{15.9}$$

where:

C_r	costs of route r
p	ports used by the route
l	links used by the route
A_p	total cost of transshipment at port p
c_l	total cost of transportation over link l
T_p	time spent during transshipment at port p
t_l	time spent during transportation over link l
α	value of transport time (USD/day/ton)

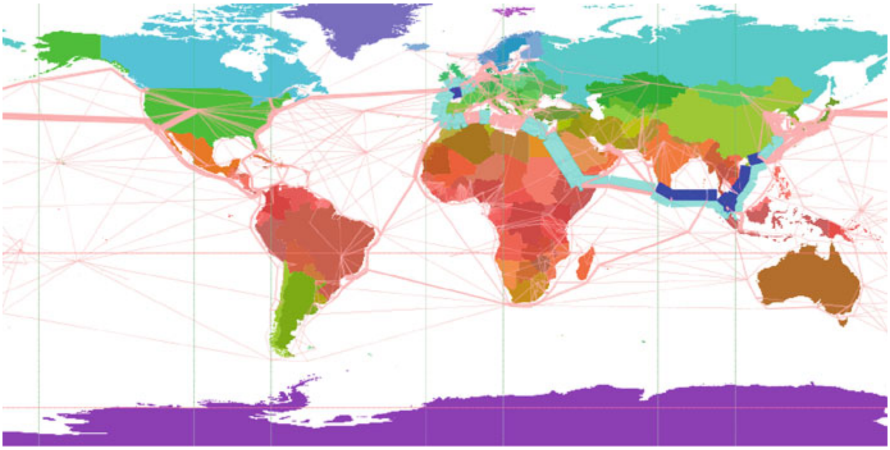


Fig. 15.6 Result of a global freight supernetwork flow assignment

A number of simplifications were made in this model. Firstly, congestion is not endogenous. We assume that investments in port capacity can keep track with the growth of flows to an extent that transport times do not increase to a significantly higher degree than those of competing ports. In case congestion is expected at a certain port, this can be modeled exogenously as part of a scenario, or via a simple capacity restraint modification. Secondly, the total cost of transshipment A_p includes non-observable port use costs and now has to be estimated. An alternative to this approach, and subject of future research, would have been to include observable cost elements and estimate mark-ups per port. Thirdly, the value of time α is not observable directly and has to be estimated or obtained from external sources. There is a fair amount of experience with estimation of values of time for land modes (see De Jong 2014) but very little for sea transport (see Feo-Valero 2011 for a recent survey). A particular research gap (and this still holds also for land modes) is to estimate distributions of preferences, which allow to model a spread over alternative routes that differ in terms of their price or speed orientation. Fourthly, shipping lines in this model are exogenous, while it can be expected that in the future they will adapt to demand. Despite these simplifications, to our knowledge the model is more sophisticated than any aggregate network model currently in use at the global level and has shown to produce satisfactory empirical results in terms of the fit of calculated port transshipment values with observations. Figure 15.6 provides a snapshot of the global container network assignment.

Finally, we note that also the inland transport links are also characterized by their transport time and prices, allowing us to distinguish between different modes of transport. Typically, there will be a trade-off between road, rail and waterways as alternative hinterland modes, in terms of speed and costs. The share of hinterland modes will be determined by the distribution of the preferences of the user of the service, with respect to time and costs. Expensive, sensitive or perishable products will not allow long transit times and their owners will be willing to pay for speed. Owners of low value, bulk goods will generally prefer slower and cheaper modes of transport that allow in-transit inventory. The application in Van Diepen et al. (2012) of this model illustrates the possibilities to develop logistics oriented scenarios for forecasting purposes. Assuming high values of time will typically create flows that favor maritime legs that are relatively short, even implying changes in port choice. In addition, rail and waterways will be less attractive. Changes in the composition of goods flows may thus be an important driver for port choice in the future.

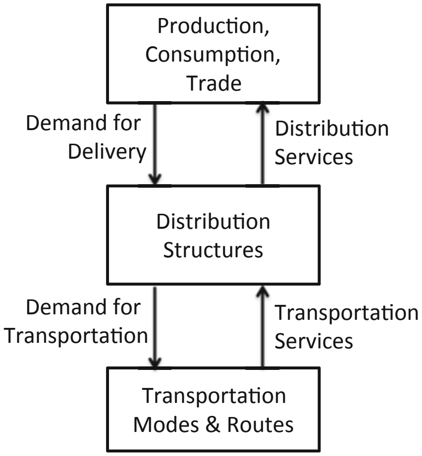
15.4 Future Developments

From the above, we conclude that global container freight demand modelling still requires further development. The following developments for the models discussed above are being studied:

- Estimation of detailed trade models in terms of sectors and commodities.
- Improved connections between trade and transport models.
- Stability in time of the transformation ratios between trade and transport models.
- Definition of global scenario models that include trade and transport.
- Dynamic models that depart from the assumption of equilibrium based economics.

A new development which extends the boundaries of the current models and is relevant for the global trade in consumer goods, concerns the decision of intermediate inventories in the supply chain. As trade in manufactured goods takes place over longer and longer distances, the associated inventories of ready products increases as well. Companies find it difficult to respond to changes in consumer demand as their supply chains become longer. Moreover, as the time pressure is high for on-demand deliveries, the cost of transport for products made to order will be very high in the case of global supply chains. Intermediate inventories can help to reduce the costs of transport as they create a decoupling point, where transport costs are relatively low upstream. Eventually, where lead times between manufacturing and consumption need to be reduced, manufacturing can be moved more closely to the regions of consumption (see e.g. the signalling report of Boston Consulting Group 2006). Modelling locations of intermediate inventories can be important for several reasons (Tavasszy et al. 2012). Firstly, the location of warehouses determines the routing of freight flows. Assuming that goods flow directly between trading partners may result in errors in prediction of the usage of ports and other infrastructure. Secondly, the function of distribution centres is to keep total logistics costs as low as possible given the service requirements, making use of the trade-off between transportation and inventory costs. Transportation costs only provide part of the picture of trading costs between regions. Knowing the distribution structures allows us to predict the costs of doing trade—and hence, trade itself—more accurately. This leads us to a third issue. In times when transport costs are becoming higher (e.g. due to increasing oil prices), companies may readjust their structures, so as not to

Fig. 15.7 Distribution structures determining global freight flows



experience the full impact of the increase in transport costs on their operations. Here, inventories do not only act as buffer for demand uncertainty, but also for uncertainty in resource costs. The implication is that the elasticity of the system will be lower for price changes than expected using a conventional freight model.

The view on demand as pictured in Fig. 15.2 is thereby extended with an additional step, where the market of logistics service providers is added in between the markets of exchange of goods and that of transportation services. The structures that are now added in between the regions generating trade flows can be termed as distribution structures (Fig. 15.7).

An example of such a model has recently been developed for the European continent. The networks can comprise of an EDC (European Distribution Centre), multiple RDCs (Regional Distribution Centre) and a BDC (Bulk distribution Centre) within a local distribution network (Fig. 15.8). In a network with an EDC, products are shipped directly from the production facility or ports to the EDC (usually in full containers), and then further distributed to customers in different places in Europe. This configuration is cost-efficient as it produces economy of scale from consolidating the inbound transport and the fact that it only uses one DC.

In addition to the growth of the market in the EU has grown in the past few years (caused by the integration of new member countries), delivery lead times of goods has been shortened considerably. For some big companies, supplying the demand in the whole Europe from a centralized distribution centre has become impossible or inefficient. To maintain efficiency, they have switched their distribution network configuration from having a centralized DC to multiple regional DCs (RDCs). They typically have 3 or 4 RDCs spread out over the European Market to serve a particular region or market area. Each of these RDCs is supplied by either a centralized DC or overseas production facilities. Clearly, port choice will be influenced by the choice of location for such a large scale distribution centres. As consumer regions will grow in Central and Eastern Europe, distribution centres will move eastwards as well, which will go at the cost of the Western-European ports.

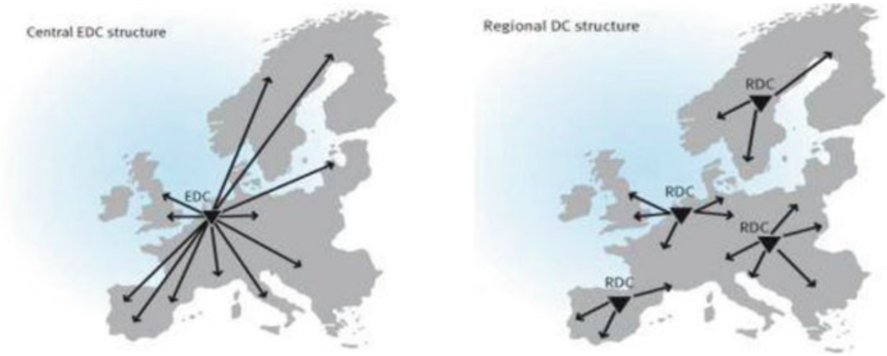


Fig. 15.8 Centralized (*left*) and decentralized (*right*) distribution centre structures (source: HIDC, 2014)

There are only a few descriptive freight logistics models that are built based on an optimization model, see Maurer (2008) and Friedrich et al. (2014). These models determine the number, level, locations of the distribution centres (DC) and the allocation of the DC's to the consumption points. The models describe respectively the distribution systems of the food retailing sector in the UK and Germany. In addition, an aggregate logit choice model for distribution structures was developed by Davydenko (see Davydenko et al. 2014) which takes into account transport and inventory costs in the choice model. The model has been calibrated with observed freight flows in Europe (using transport OD data). An optimization model was developed as part of a global freight simulation framework (Halim et al. 2012) that calculates the probability that regions will host distribution centres for specific flows towards regions within Europe. This model determines distribution structures in port hinterlands based on a well-known logistic problem called the Network Design Problem (NDP). The NDP deals with the determination of distribution centre location and the routing of the flows of the transported materials in such a way the total logistics cost are minimized and minimum service requirements can be met. The model is designed using a similar modelling principle as presented in Friedrich et al. (2014), with a specification that allows estimation of unknown parameters using observed flow data.

Figure 15.9 shows a result of this model for the layout of distribution centres.

This model is being developed further in an empirically validated, descriptive model setting. Especially for those goods that require short lead times between order and delivery, and hence rely on the presence of distribution centres, these models will be an important addition to the conventional freight models.

Fig. 15.9 model result of one layout of DC structures within Europe

15.5 Summary and Concluding Remarks

Models of global container transport demand can follow the architecture available for freight modelling. We describe the basic methods and techniques available for these models by reviewing the literature in this area. Surprisingly little academic literature is available that treats modelling approaches specifically geared towards global container transport demand, despite specificities like:

- Influence of transport and non-transport barriers
- Heterogeneity in commodities and impact on choice
- Competition between sea and air
- Port choice and routing determined at global level
- Dependence of global supply chains.

In this chapter we treat the subject in two main parts. One part involves modelling the demand for movement between regions, i.e. the outcome of the processes of production, consumption and trade. The second part involves the modelling of demand for transport services by mode and route of transport, including the demand for maritime and inland port services.

The global trade literature is extensive and presents different empirical models, that allow trade forecasting and policy analysis at a global level. The literature on modelling the requirements with respect to services is less dense, however, when it concerns maritime container transport. Most of the modelling has developed in the area of inland (unimodal and multimodal) transportation. Some work on mode choice is available (for our purposes: mostly between air and sea), partly emanating from the global trade literature. Network choice will first and foremost focus on the choice of port of call. Here, there is ample literature, but again, little that has been shown to be applicable at the global level.

Future work may include the exploitation of linkages with supply chain management. This may include a better understanding of the contribution of shippers' preferences to observed shipping choices. Also, future developments in geographic restructuring of supply chains because of changes in manufacturing locations or distribution structures, could be looked into. Finally, as global, integrative models do not yet exist, combining new trade and transport network models may provide exciting insights for long term forecasting and policy analysis.

References

Anderson, J. E., & Yoto, V. (2010). The changing incidence of geography. *The American Economic Review, 100*(5), 2157–2186.

Anderson, S. P., De Palma, A., & Jacques-Francois, T. (1987). The CES is a discrete choice model? *Economics Letters, 24*(2), 139–140.

Bergstrand, J. H., Egger, P., & Larch, M. (2013). Gravity redux: Estimation of gravity-equation coefficients, elasticities of substitution, and general equilibrium comparative statics under asymmetric bilateral trade costs. *Journal of International Economics, 89,* 110121.

Brooks, M. R., Puckett, S. M., Hensher, D. A., & Sammons, A. (2011). Understanding mode choice decisions: A study of Australian freight shippers, ITLS working paper ITLS-WP-11-20. Sydney: University of Sydney.

Capros, P., Georgakopoulos, T., Filippoupolitis, A., Kotsomiti, S., & Atsaves, G. (1997). The GEM-E3 model for the European Union: Reference manual National Technical University of Athens and others.

Chou, C. C. (2009). An empirical study on port choice behaviors of shippers in a multiple-port region. *Marine Technology Society Journal, 43*(3), 71–77.

Davydenko, I., Thissen, M., & Tavasszy, L. (2014). Translation of trade flows into transport flows in European logistics facility choice model. In Transportation Research Board 93rd Annual Meeting (No. 14-2847).

De Benedictis, L., & Tajoli, L. (2011). The world trade network. *World Economy, 34,* 1417–1454.

De Jong, G. (2014). Freight service valuation and elasticities. In L. A. Tavasszy & G. de Jong (Eds.), *Modelling freight transport*. London: Elsevier.

De Langen, P. W. (2007). Port competition and selection in contestable hinterlands; the case of Austria. *European Journal of Transport and Infrastructure Research, 7*(1), 1–14.

De Langen, P. W., Van Meijeren, J., & Tavasszy, L. A. (2012). Combining models and commodity chain research for making long-term projections of port throughput: An application to the Hamburg-Le Havre range. *European Journal of Transport & Infrastructure Research, 12*(3), 310–331.

DHL. (2012). Delivering tomorrow—Dialogue on future trends. www.delivering-tomorrow.com/. Accessed 27 Oct 2014.

Erlander, S., & Stewart, N. F. (1990). The gravity model in transportation analysis. Utrecht: VSP.

Feo-Valero, M., García-Menéndez, L., & Garrido-Hidalgo, R. (2011). Valuing freight transport time using transport demand modeling: A bibliographical review. *Transport Reviews, 31*(5), 625–651.

Fremont, A. (2007). Global maritime networks: The case of Maersk. *Journal of Transport Geography, 15*(6), 431–442.

Friedrich, H., Tavasszy, L. A., & Davydenko, I. (2014). Distribution structures. In L. A. Tavasszy, & G. de Jong (Eds.), *Modelling freight transport*. London: Elsevier.

Giannopoulos, G., Aifadopoulou, G., & Torok, A. (2007). *A port choice model for the transsshipment of containers in eastern Mediterranean*. Athens: Hellenic Institute of Transport.

Gunning, J. W., & Keyzer, M. A. (1995). Applied general equilibrium models for policy analysis. *Handbook of development economics, 3*(1), 2025–2107.

Guy, E., & Urli, B. (2006). Port selection and multicriteria analysis: An application to the Montreal-New York alternative. *Maritime Economics & Logistics, 8*(2), 169–186.

Halim, R. A., Tavasszy, L. A., & Seck, M. D. (2012). *Modeling the global freight transportation system: A multi-level modeling perspective*. Proceedings of the 2012 Winter Simulation Conference (WSC), pp. 1–13.

Hummels, D., & Schaur, G. (2012). Time as a trade barrier (No. w17758). National Bureau of Economic Research.

Hausman, W. H., Lee, H. L., & Subramanian, U. (2013) The impact of logistics performance on trade. *Production and Operations Management, 22*(2), 236–252.

HIDC. (2014). http://www.ndl.nl/services/faq/what-is-the-difference-between-an-edc-an-rdc-and-a-bdc/. Accessed 5 Feb 2014.

Ivanova, O. (2014). Modelling inter-regional freight demand with inputoutput, gravity and SCGE methodologies. In L. A. Tavasszy & G. de Jong (eds.)*, Modelling freight transport*. London: Elsevier.

Ivanova, O., Heyndrickx, C., Spitaels, K., Tavasszy, L. A., Manshanden, W., & Snelder, M. (2007). RAEM: version 3.0. Leuven: TML. www.tmleuven.be/project/raem/RAEMFinalreport.pdf. Accessed 27 Oct 2014.

Jong, G. de, & Ben-Akiva, M. (2007). A micro-simulation model of shipment size and transport chain choice. *Transportation Research B, 41*(9), 950–965.

Jourquin, B., & Beuthe, M. (1996). Transportation policy analysis with a geographic information system: The virtual network of freight transportation in Europe. *Transportation Research Part C, 4*(6), 359371.

Jula, P., & Leachman, R. C. (2011). Long-and short-run supply-chain optimization models for the allocation and congestion management of containerized imports from Asia to the United States. *Transportation Research Part E: Logistics and Transportation Review, 47*(5), 593–608.

Kato, T., Hanaoka, S., Chin, A. T., & Kawasaki, T. (2012). Investigating sea shift in international freight transport: A case between Southeast Asia and the US. www.titech.ac.jp. Accessed 27 Oct 2014.

Lewis, D. B (1994). *Freight mode choice: Air transport versus ocean transport in the 1990's, FTL Report 94-9*. Cambridge: MIT.

Lirn, T. C., Thanopoulou, H. A., Beynon, M. J., & Beresford, A. K. C. (2004). An application of AHP on transhipment port selection: A global perspective. *Maritime Economics and Logistics, 6,* 70–91.

Lloyd's Register, QinetiQ, & University of Strathclyde. (2013). *Global marine trends 2030*. London: Lloyd's Register Group Limited.

Malchow, M., & Kanafani, A. (2001). A disaggregate analysis of factors influencing port selection. *Maritime Policy & Management, 28*(3), 265–277.

Malchow, M. B., & Kanafani, A. (2004). A disaggregate analysis of port selection. *Transportation Research E, 40,* 317–337.

Mansur, A., Whalley, J. S., & Shoven, J. B. (1981). *Numerical specification of applied general equilibrium models: Estimation, calibration, and data. Applied general equilibrium analysis. 3*. Conference on Applied General Equilibrium Analysis, San Diego.

Mazzarino, M. (2012). Strategic scenarios of global logistics: What lies ahead for Europe? *European Transport Research Review, 4,* 1–18.

Maurer, H. (2008). *Development of an Integrated Model for Estimating Emissions from Freight Transport*, PhD. Dissertation. Leeds: University of Leeds, Institute for Transport Studies.

Meng, Q., Wang, X., & Miao, L. (2013). *Boundary estimation of probabilistic port hinterland for intermodal freight transportation operations, Selected proceedings of the WCTR*. London: WCTRS.

MIT. (2011). Future freight flows study. http://ctl.mit.edu/research/futurefreightflows/scenarios. Accessed 27 Oct 2014.

Murphy, P., & Daley, J. (1994). A comparative analysis of port selection factors. *Transportation Journal, 3,* 15–21.

Murphy, P., Daley, J., & Dalenberg, D. (1992). Port selection criteria: An application of a transportation research framework. *Logistics and Transportation Review, 28,* 237–255.

Nir, A.-S., Lin, K., & Liang, G.-S. (2003). Port choice behaviour—From the perspective of the shipper. *Maritime Policy and Management, 30,* 165–173.

Notteboom, T., & Rodrigue, J.-P. (2009). The future of containerization: Perspectives from maritime and inland freight distribution. *Geojournal, 74*(1), 7–22.

Port of Rotterdam. (2011). *Port vision 2030.* Rotterdam: Port of Rotterdam.

Rodrigue, J.-P. (2006). Challenging the derived transport demand thesis: Issues in freight distribution. *Environment & Planning A, 38*(8), 1449–1462.

Song, D. W., & Yeo, K. T. (2004). A competitive analysis of Chinese container ports using the analytic hierarchy process. *Maritime Economics and Logistics, 6,* 34–52.

Srinivasan, T. N., & Archana, V. (2009). *India in the global and regional trade: Determinants of Aggregate and Bilateral Trade Flows and Firms' Decision to Export.* ICRIER Working Paper 232.

Tang, L. C., Low, J. M. W., & Lam, S. W. (2008). Understanding port choice behavior—A network perspective, Networks and Spatial Economics.

Tavasszy, L. A., & G. de Jong (2014). *Modelling freight transport*. London: Elsevier.

Tavasszy, L. A., Minderhoud, M., Perrin, J.-F., & Notteboom, T. (2011). A strategic network choice model for global container flows: Specification, estimation and application. *Journal of Transport Geography, 19*(6), 1163–1172

Tavasszy, L. A., Ruijgrok, C. J., Thissen, M. J. P. M. (2003). Emerging global logistics networks: Implications for transport systems and policies, Growth and Change: *A Journal of Urban and Regional Policy, 34*(4), 456–472.

Tavasszy, L. A., Smeenk, B., & Ruijgrok, C. J. (1998). A DSS for modelling logistics chains in freight transport systems analysis. *International Transactions in Operational Research, 5*(6), 447–459.

Tavasszy, L. A., Ruijgrok, K., & Davydenko, I. (2012). Incorporating logistics in freight transport demand models: State-of-the-art and research opportunities. *Transport Reviews, 32*(2), 203–219.

Tiwari, P., Itoh, H., & Doi, M. (2003). Shippers' port and carrier selection behavior in China: A discrete choice analysis. *Maritime Economics and Logistics, 5,* 23–39.

Tongzon, J. L. (2009). Port choice and freight forwarders. *Transportation research Part E, 45,* 186–195.

Tsuboi, T., Hyodo, T., & Wakita, T. (2010). International Marine/Air cargo modal split model considering logistics cost. *Transport Policy Studies' Review, 12*(4), 32–41.

Van Diepen, L. A., & Tavasszy, K. C. (2012). Predicting the effects of global economic scenarios on long term container port throughput, 2nd Asian Logistics Round Table, Vancouver.

Veldman, S. J., & Bückmann, E. H. (2003). A model on container port competition: An application for the west European container hub-ports. *Maritime Economics & Logistics, 5*(1), 3–22.

Walmsley, T., Aguiar, A., & Narayanan B. (2012). *Introduction to the Global Trade Analysis Project and the GTAP Data Base, GTAP Working Paper No. 67.* West Lafayette: Purdue University.

Yamada, T., Imai, K., Nakamura, T., & Taniguchi, E. (2011). A supply chain-transport supernetwork equilibrium model with the behaviour of freight carriers. *Transportation Research E, 47*(6), 887907.

Yang, D., Ong, G. P., & Chin, T. H. (2013). An exploratory study on the effect of trade data aggregation on international freight mode choice. *Maritime Policy & Management, 41*(3), 212–223.

Yeo, G. T., Ng, A. K., Lee, P. T. W., & Yang, Z. (2013). Modelling port choice in an uncertain environment. *Maritime Policy & Management, 41*(3), 251–267.

Zhang, M., Wiegmans, B., & Tavasszy, L. (2013). Optimization of multimodal networks including environmental costs: A model and findings for transport policy. *Computers in Industry, 64*(2), 136–145.

Zondag, B., Bucci, P., Gützkow, P., & de Jong, G. (2010). Port competition modeling including maritime, port, and hinterland characteristics. *Maritime Policy & Management, 37*(3), 179–194.

Chapter 16
Competition and Co-operation in Maritime Logistics Operations

Eon-Seong Lee and Dong-Wook Song

Abstract This chapter aims to examine environmental challenges that maritime logistics operators have recently faced with, and to investigate strategic ways that maritime logistics operators can effectively manage competition and co-operation with their rivals for better responding to those challenges and thus achieving their strategic goals, i.e. maximisation of maritime logistics value. In order to address the aforementioned issues, this chapter adopts social network embeddedness and knowledge management perspectives. Based on the literature, a theoretical framework is established to show the positive relationship between co-opetitive networks, knowledge acquisition and maritime logistics value. A comprehensive survey to existing literature reveals that a high level of co-operation in a co-opetitive network (i.e., high numbers of and strong relationships between maritime logistics operators) facilitates knowledge acquisition, and competition promotes the positive impact of co-operation on knowledge acquisition. The acquired knowledge helps to improve maritime logistics value. This outcome will certainly provide maritime operators with a strategic insight into the identification of determinants and/or sources for competitive advantage and greater organisational performance from inter-organisational coordination and knowledge-based perspectives.

16.1 Introduction

Over the last decade, the maritime transport industry has experienced a variety of environmental challenges, such as the following: changes in trade patterns; larger-sized vessels; intensive competition; port privatisation; intermodality; and finally the

---

E. -S. Lee (✉)
Department of Maritime Logistics and Management, Australian Maritime College,
Univerisity of Tasmania, Launceston, Australia
Tel.: + 61 3 6324 9882
e-mail: e.lee@amc.edu.au

D. -W. Song
Transport Research Institute, Edinburgh Napier University, Merchiston Campus, Edinburgh, UK
Tel.: + 44 (0)131 455 2553
e-mail: D.Song@napier.ac.uk

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_16

global expansion of maritime transport operations. In order to effectively respond to such environmental challenges, maritime transport operators have been forced to be effectively integrated into the global logistics system by offering not only a sea transport service but also additional logistics services such as warehousing, material handling, inventory, and packaging (Panayides 2006). The maritime transport system which fulfils such demands is referred to as a maritime logistics system. The maritime logistics system consists of major maritime transport operators such as shipping lines, port/terminal operators and freight forwarders, and the operators are required to improve their logistics value by providing their service in a more efficient and effective manner. This can be achieved when they offer quicker, responsive, flexible and reliable services with a lower price. As the maximisation of the maritime logistics value may help to improve the entire logistics performance as well as to refine the operators' sustainable competitiveness, it has become one of the most significant strategic goals maritime operators want to achieve.

Having acknowledged the above trend, the following research issue arises: how can the maritime logistics operators achieve these strategic goals? This chapter aims to thoroughly examine how the maritime logistics operators can improve their maritime logistics value. For this, this chapter applies inter-organisational co-opetition and knowledge management strategies, and systematically addresses how those strategies can help maritime operators to maximise their maritime logistics value. More specifically, this chapter initially examines how maritime logistics operators co-operate with their rivals and then gain knowledge-based advantages from such co-opetitive relationships. This chapter also analyses whether the acquired knowledge may facilitate the improvement of operational efficiency and service effectiveness of the maritime logistics system.

In the earlier part of this chapter, the concept of maritime logistics and maritime logistics value is introduced, and its strategic significance in the current business environment is identified. Literatures in co-opetitive relationships and knowledge acquisition are reviewed, and a theoretical framework exploring the relationship between co-opetitive networks, knowledge acquisition and maritime logistics value will then be developed. Finally, managerial implications are suggested based on the empirical findings.

16.2 Strategic Goal in Maritime Logistics Operations

In order to maximize logistics performance, logistics activities which are globally dispersed such as global sourcing, warehousing, transportation and packaging should be operated in a highly integrated manner (Waters 2003; O'Leary-Kelly and Flores 2002). As maritime transport operators are key components of the global logistics system, they must also, as an integrated entity of the logistics system, be well connected with other logistics functions (Bowersox 1978; O'Leary-Kelly and Flores 2002; Panayides 2006; Roh et al. 2007; Panayides and Song 2008). The maritime transportation system thus no longer an independent entity which pursues its own interest, but must be an integral component which provides integrated logistic services, which themselves are involved in the ocean carriage from planning to delivering of goods and information to final customers. This process has been referred to as a maritime logistics system (Panayides 2006; Lee and Song 2010).

There are three key parts of a maritime logistics operation: shipping, port/terminal operation, and freight forwarding. The value of maritime logistics services can be improved when the maritime operators' activities are well co-ordinated as a single team and provide a more efficient and effective service to their customers (Lai et al. 2002; O'Leary-Kelly and Flores 2002). In this sense, the maritime logistics value can be referred to as operation efficiency and service effectiveness of the maritime logistics system (Lee and Song 2010; Song and Lee 2012), and it can be measured by the extent to which maritime operators can provide their service with lower costs and quicker time, and also by how the operators deliver the goods in a more flexible, responsive and reliable manner (Baudin 2004; Lai et al 2002).

Maritime logistics value is very crucial for maritime logistics operators because improving operational efficiency and service effectiveness may help maritime operators to successfully satisfy their customers, and as a result it may promote the greater organisational performance of the operators. Furthermore, as all of the activities of in a global logistics are inter-linked with each other, and as individual activities may mutually affect other functions within the system, the improvement of maritime logistics value may contribute to the enhancement of the performance of the entire logistics system as a whole. Thus, maximising maritime logistics value has recently become a significant strategic goal that maritime logistics operators need to achieve. The next section discusses the dynamically changing business environments in the maritime logistics industry, which should be considered in order for maritime logistics operators to achieve this strategic goal.

16.3 Business Environments of the Maritime Logistics Industry

Over the last decade, the maritime logistics industry has experienced environmental challenges. The challenges are presented as follows: firstly, containerisation has necessitated a larger vessel size. Shipping lines now vigorously compete to ensure vessel size as large as they can manage, in order to gain advantages of economies of scale and attract powerful shippers with a large amount of products to be shipped (Fremont 2007). For example, Maersk Line adopts the EMMA MAERSK, which has a capacity of up to 11,000 TEU, and operates the ultra large container vessel for their Asia—Europe line. Other leading shipping lines, such as CMA CGM, or Hanjin Shipping, have also operated large sized vessels over 8000 TEU.

Secondly, intermodality has recently been an increased amount of attention. Intermodality is defined as "integrated transportation systems consisting of two or more modes" (Haasis 2008, p. 269). Most shippers normally arrange two or more forms of transport modes in order to ensure that their goods are efficiently delivered to the final destination. Port is an inter-mediate entity that connects different modes of transport such as sea, road and rail. In order to offer a quick and efficient service,

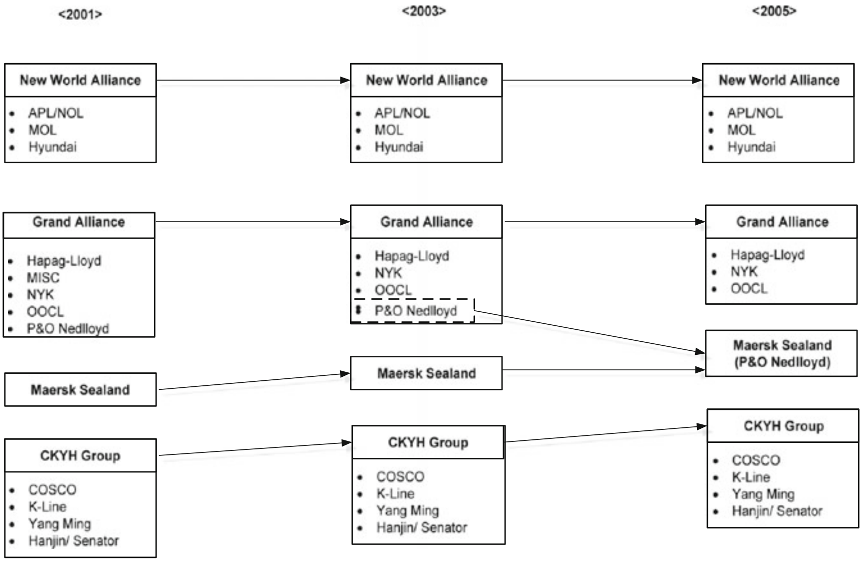


Fig. 16.1 The participation of shipping lines in strategic alliances. (Source: Busan Port Authority 2007, p. 27)

maritime operators are forced to put together all possible transportation modes and to effectively coordinate with other modes of transport (Marlow and Paixao 2003). For example, ports should ensure that cargoes are smoothly and safely connected into road or rail modes, and from there delivered to their final destination (Song 2003). Today, efficient inter-modality is extremely important for maritime operators to fulfil their customer needs and to improve the performance of the entire logistics system.

Thirdly, alliances and integration of shipping lines have accelerated. Due to the aforementioned environmental challenges such as the larger vessel size and inter-modality, shipping liners must look for ways to collectively respond to the large enterprises with huge sized vessels and to offer the most flexible services by collaborating with other liners. This effort has boosted the alliances among shipping lines. Evidence of three large strategic alliances and one merger and acquisition are described in Fig. 16.1. Global strategic alliances are the most popular form of shipping lines' co-operation. The Grand Alliance consists of Hapag-Lloyd, Nippon Yusen Kaisha (NYK) and Orient Overseas Container Line (OOCL), who plan to extend their co-operative relationship up to 2017. Cosco/K-Line/Yangming/Hanjin Alliance and New World Alliance are also one of the most popular strategic alliances in the shipping industry. In 2005, Maersk Sealand acquired P&O Nedlloyd, which was a member of the Grand Alliance (Lee 2010).

On the other hand, vertical integration of shipping lines has also proven to be a popular strategy. Vertical integration occurs when a firm acquires another firm operating in the inbound or outbound logistics chain for the acquiring firm's products

Table 16.1 Shipping lines' involvement in port terminals. (Source: Busan Port Authority 2007, p. 28)

Shipping Line	Port Terminals
Maersk	Hong Kong, Kaohsiung, Yokohama, Rotterdam, etc.
Maersk line	Oakland, Long Beach, New York/New Jersey, etc.
Evergreen	Los Angeles, Tacoma, etc.
COSCO	Hong Kong, Shekou, etc.
NOL/APL	Karachi, Los Angeles, Oakland, etc.
OOCL	Kaohsiung, Vancouver BC, etc.
Hyundai shipping	Long Beach, Busan, Gwangyang, Kaohsiung, etc.
Hanjin shipping	Long Beach, Busan, Gwangyang, Kaohsiung, Seattle, Chicago, Tokyo, Osaka, etc.

or services. For example, shipping lines have expanded their businesses into port operations through vertical integration. Table 16.1 shows examples of this vertical integration by shipping lines, who have all recently expanded the port terminal operation by establishing dedicated, globally situated terminals. The vertical integration has thus allowed shipping lines to have priority to use their own terminals, to offer a wider range of services, to provide shippers with a more stable service, and also to maintain control over shipments (Lee 2010).

Fourthly, port privatisation is also one of the most significant challenges in the maritime logistics industry. Recently, as many port organisations have become managed by private companies, port competition has become more intensive (World Bank 2006). On the other hand, thanks to port privatisation, port terminal operators have more opportunities to easily enter new foreign port operation markets (Slack and Fremont 2005).

Fifthly, port terminal operators are globally expanding their business. Table 16.2 gives a brief summary of the global expansion of global port terminal operators. The purpose of such expansion of port terminal operators is to increase their competitive influence globally by broadening their business scale and scope, and to gain valuable resources in foreign markets (Notteboom 2004; Notteboom and Winklemans 2001; Slack and Fremont 2005).

Finally, competition among maritime operators is getting increasingly intense. As this fierce competition among maritime logistics operators has become one of the most serious challenges in the industry, this specific trend and the current situations are thoroughly discussed in the next section. The next section also examines how the maritime logistics operators can strategically administrate the intensive competition among their rivals in order to overcome the environmental challenges and to improve maritime logistics value. This will be discussed from the co-opetition strategy perspective, i.e. co-operation among competitors.

Table 16.2 Global Expansions of Port Terminal Operators. (Source: Compiled from the Ministry of Land, Transport and Maritime Affairs in Korea (2008))

Global Port Terminal Operators	Europe	North America	East and North Asia
HPH	Felixstowe, Rotterdam (ECT, Delta, ECT Home, Hanno), Thamesport, Harwich, Gdynia		Hong Kong (HIT, Cosco-Hit, Asia Port, Rivertrade), Shanghai (SCT, SPICT), Yantian, Juizhou, Nanhai, Shantou, Jiangmen, Gaolan, Xiamen, Ningbo, Guangdong, Shanghai Mingdong, Busan, Kwangyang (HKT, KIT)
PSA	Antwerp, Zeebrugge, Genoa, Venice, Shines		Singapore, Dalian, Nantong, Fuzhou, Guangzhou, Tiacang, Incheon, Hibiki
Eurogate	Bremerhaven, Hamburg, La Spezia, Giaio Tauro, Lisbon, North Sea Terminal, MSC Gate, Livorno, Salerno, Contentori Ravenna, CICT Porto, Rjieka		
SSA		Los Angeles, Long Beach, New Orleáns, Oakland, Portland, Seattle	
Cosco	Antwerp, Naples (Molo Bausan)	Long Beach	Hong Kong (Cosco-HIT), Dalian (DPC, DPCT), Qingdao (QCIT, QQCT), Shanghai (SPICT, SCT), Zhangjiagang, Yantian (YICT), Yingkou, Yangzhou Yuanyang, Quanzhou, Tianjin, Nanjing, Zhenjiang Jinyuan, Taicang
DPW	Southampton, Tilbury, Antwerp, Le Havre, Germersheim, Constantza, Marseille-Fos	Vancouver	Yantai, Shekou, Hong Kong, Tianjin, Qingdao, Pusan, Vostochnyy
APMT	Aarhus, Rotterdam, Antwerp, Bremerhaven, Dunkirk, Giaio Tauro, Constanza	Tacoma, Oakland, Los Angeles, New York, Baltimore, New Orleans. Portsmouth, Charleston, Jacksonville, Hampton, Port Everglades, Miami, Houston	Kobe, Yokohama, Dalian, Qingdao, Shanghai, Guangzhou, Kaohsiung, Yantian

16.4 Competition and Co-operation in Maritime Logistics Operation

This section deals with the specific function and role of the maritime logistics operators, and examines how the maritime logistics operators compete with each other and how they strategically co-operate with their rivals in order to both respond to this fierce competition and improve their maritime logistics value. This section also examines knowledge acquisition advantages, which can be gained from their co-opetitive relationships. Finally, the role of knowledge acquisition in improving the maritime logistics value will then be discussed. A theoretical framework which shows the relationship between co-opetitive network embeddedness, knowledge acquisition and maritime logistics value will be then developed.

16.4.1 Shipping Lines

Over the last decade, worldwide container shipping volumes have steadily increased. Figure 16.2 shows the growth of international containerised shipping volume from 2001 to 2013. As seen in the Fig. 16.2 (UNCTAD) (2013), the international trade derived from containerisation have rapidly increased since 2001, and over 70 per cent of the value of world international seaborne trade is being moved in containers.

Table 16.3 shows a global ranking of major shipping lines, based on the total TEU capacity deployed by the named carrier. The top ranked firm, APMM Group, is one of the leading liner shipping firms in the world, serving customers all over the globe, including Asia, Europe, Africa, Oceania, Central/South America, etc. Other liners ranked in a top class also provide a wider variety of shipping routes around the world. Whilst a small number of large enterprises move over the world with a large shipping capacity of their own, a great number of small and medium sized shipping lines focuses more on specialised shipping routes (Panayides and Gray 1999).

Today, the customers of shipping lines have become more demanding. They expect shipping lines to offer more frequent and flexible shipping schedules, reliable and safe sailings, and quicker service with lower freight rates (Notteboom 2006). In order to satisfy the customers, shipping liners must therefore establish good relationships with their customers as well as shipping industry stakeholders such as other carriers, suppliers, manufacturers and final customers. They are also required to adopt leading information technology and related systems to satisfy their customer demands and to respond to dynamically changing environments (Panayides and Gray 1999).

Therefore, the competition in a shipping industry becomes much more intense. For example, the competition among shipping lines which have similar shipping routes is extremely tough. Shipping lines vary in sizes—e.g. large, medium and small firms—and, the competition among firms of a similar size tends to be more intense (Panayides and Gray 1999). In order to respond to this intense competition, shipping

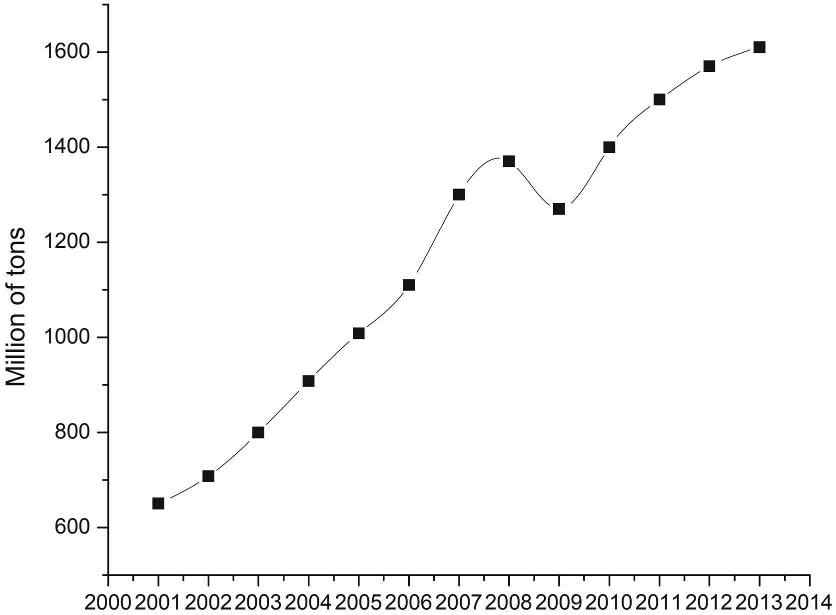


Fig. 16.2 Growth of international containerised shipping volume, 2001–2013. (Source: Modified from UNCTAD (2013, p. 23))

Table 16.3 World major shipping lines (as of 2013). (Source: Modified from Containerisation International (2013a))

Shipping lines	Rank	TEU
APMM Group	1	2,521,494
MSC	2	2,121,030
CMA CGM	3	1,453,463
Evergreen Line	4	734,975
COSCON	5	731,866
Hapag-Lloyd	6	654,919
APL	7	603,133
CSCL	8	571,683
Hanjin	9	566,734
MOL	10	527,300
OOLC	11	478,110
Hamburg Sud	12	448,257
NYK Line	13	410,417
Yang Ming	14	392,999
HMM	15	379,714

lines have proactively collaborated with their competitors. For example, shipping conferences, vessel sharing agreements and strategic alliances have increased, so as to avoid mutually destructive competition and allow shipping lines to protect their business (Frankel 1982; Brooks 2000). As shipping alliances have been discussed in the previous section, this section details further shipping conferences and vessel sharing agreements under the term of shipping collaboration.

Shipping liners which move similar shipping routes have collaborated by joining shipping conferences. The shipping conferences have enabled shipping lines to fix the freight rates in order to reduce competition among themselves and protect their business. The Transatlantic Agreement (TAA) is a good example of shipping conferences. TAA, which was made in 1993, allows shipping lines which participate in the agreement to control the price rates and capacity of cargoes as well as other business conditions of shipping services in the North Atlantic (Heaver et al. 2000). However, there has recently been a movement to regulate against such shipping conferences. For example, in 2008 the Competitiveness Council of the European Union (EU) decided to stop admitting a liner shipping conference in Europe (Korean Fair Trade Commission 2009). On the other hand, vessel sharing agreements have also been popular between shipping liners. The purpose of the vessel sharing agreement is to fix the amount of vessel capacity and share the carriers' slot per trip between shipping lines. For example, two shipping lines who are partners for vessel sharing along the same route, fix the vessel capacity to efficiently share their cargoes in order to maintain optimised use of the vessels, offer various time schedules in a more flexible way and deliver the cargoes on time (Lei et al. 2008).

Yet despite the above efforts to survive in the industry, shipping lines are still confronted with a lot of strategic tasks in flexibly responding to the volatile demands of their customers, and in maximising their profits under the dynamically changing business environments. They also need to design and implement optimal strategies in order to gain competitive advantages and enhance maritime logistics value as a key component in maritime logistics.

16.4.2 Port Terminal Operation

Ports are the interface between sea and land, and areas for berthing or anchoring ships and allowing for the transfer of goods from ship to land or ship to ship (Alderton 1999). Port terminal operators in a port provide a cargo handling service as well as various logistics services such as warehousing and packaging services (Panayides 2006; Roh et al. 2007). Table 16.4 summarises major global port terminal operators. Seeing the market structure of the port industry, port terminal operations around the world are dominated by a small number of global port terminal operators (Slack and Fremont 2005). For instance, the top five operators controlled about 60 % of the global container-handling activity (Containerisation International 2013b).

In recent times, port terminal operators are required to be effectively integrated into the entire logistics system. This view that port terminal operators are an integrated entity of the global logistics system has been referred to as port logistics (Roh

Table 16.4 Leading port terminal operators (as of 2012). (Source: Modified from Containerisation International (2013b))

International port terminal operator	Rank	Total Million TEU
HPH	1	74.3
APM Terminals	2	66.2
PSA International	3	59.7
Cosco Pacific	4	56.3
DP World	5	54.5

et al. 2007), as illustrated in Fig. 16.3. They subdivide port logistics into two flows: physical flows and information flows. The upper part of the figure shows a physical flow of moving cargoes through port terminal, e.g. the port entry system, stevedore system, transit system, storage system and linkage system. All relevant information moves with the physical flow. Each sub-system is interlinked according to the cargo flow in the port logistics process (Roh et al. 2007).

Competition between port terminal operators has become very fierce. This is mainly due to the advent of vigorous privatisation in ports across the world. The competition among port terminal operators has also been overheated by their customers, i.e. shipping lines, who have become bigger and more powerful through their respective collaborations. These powerful shipping groups are demanding much more favourable service charges and operational conditions from port terminal operators; these demands, in turn, have become a huge threat to port terminal operators (Notteboom 2004). The increased power of these customers has also forced port terminal operators to establish new, large terminals and invest huge amounts of money into information systems and modern communication technologies, in order that they may effectively handle the huge amount of cargoes moved by these larger-sized vessels (Shang and Marlow 2005). This threat has caused port terminal operators to react aggressively, through collaboration with port terminal operators located worldwide. For example, PSA and HPH are involved in joint ventures for the mutual interests of securing their business. In China, HPH now co-operates with Shanghai Port Container Co. Ltd. by investing in their 50/50 joint venture (De Souza et al. 2003).

16.4.3 Freight Forwarders

Freight forwarders are intermediate entities in global logistics, which connect shippers and shipping lines and facilitate cross-border trade (Murphy and Daley 2001). Figure 16.4 describes the operations of global freight forwarders. Freight forwarders provide a great number of various services with shippers (i.e. exporter and importer), which include customs authorities in both the country of origin and country of destination. Table 16.5 shows the top twenty global ocean freight forwarders ranked by TEU volumes.

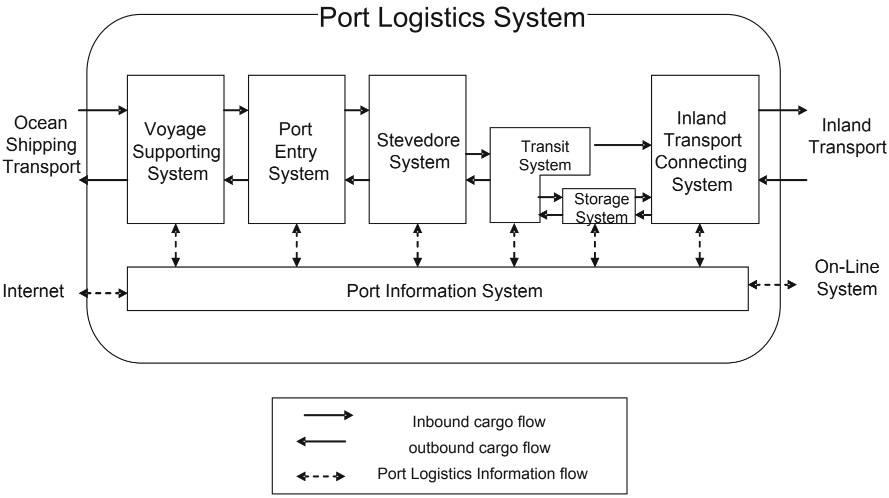


Fig. 16.3 Port logistics system. (Source: Roh et al. 2007, p. 289)

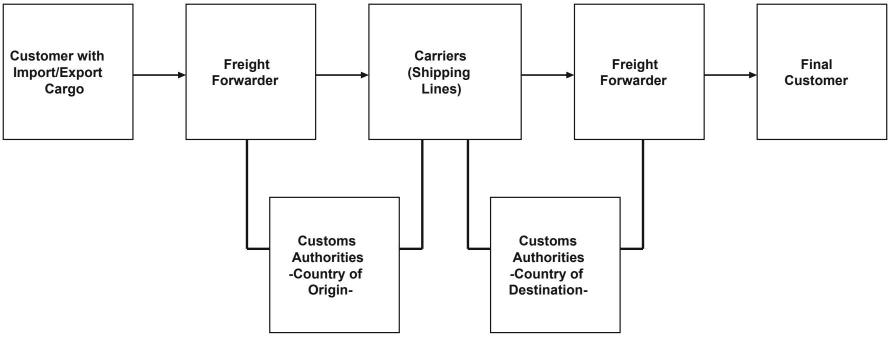


Fig. 16.4 Structure of the freight forwarding industry. (Source: Bernal et al. 2002, p. 240)

A rapid increase in international trade volumes has facilitated the constant growth of the freight forwarding industry. Freight forwarders vary in type and size, from smaller and more specialised firms who deal with particular types of goods or operate within particular areas, to bigger firms who can cover huge ranges of goods and geographical areas in their forwarding services. By providing the above crucial maritime logistics services with shippers, freight forwarders play a critical role in moving raw materials to finished products (Bernal et al. 2002).

The specific services provided by freight forwarders may include planning the most appropriate route for a shipment, based on nature of the goods, cost, transit time and security; arranging payment of freight and other charges on behalf of the shippers; preparing documentation issues, such as bills of lading (B/L), or any documents required for customs clearance or insurance requirements (Coyle et al. 1999; Bernal

Table 16.5 Top 20 global ocean freight forwarder ranked by TEU Volume (2005). (Source: Global Logistics and Supply Chain Magazine (2006))

Freight Forwarder	Rank	TEU
Kuehne + Nagel	1	1,600,000
DHL Danzas	2	1,200,000
Shenker	3	890,000
Panalpina	4	842,000
BDP International	5	800,000
Excel	6	717,000
UPS SCS	7	660,000
Expeditors	8	643,300
NYK Logistics	9	619,000
ABX Logistics	10	500,000
Kerry Logistics/EAS	11	405,000
Kintetsu Worldwide Express	12	311,000
UTi	13	252,000
Nippon Express	14	250,000
TNT Logistics	15	230,000
Hecny/Global Link	16	160,000
Wolf D Barth	17	121,000
Round-The-World-Logistics	18	110,000
Phoenix International Freight	19	101,000
Top Ocean	20	100,000

et al. 2002). In recent time, freight forwarders have been required to provide multiple logistical services, such as custom-house agency, inventory management, appropriate packing, warehousing, tracking, inland transportation, and expediting shipments and offering recommendations about the most suitable shipping routes, rather than simply offering forwarding services (Murphy and Daley 2001).

Due to the complex customer demands outlined above, the competition among freight forwarders has become increasingly tough. Big players who can meet the customer needs are getting larger, while small players are either struggling to survive or being liquidated from the industry (Rushton and Walker 2007). This industrial trend has changed the industry structure of freight forwarders in a way that has forced freight forwarders to proactively engage in inter-organisational co-operation among the forwarding operators. This co-operation takes place in the forms of merger and acquisitions (M&A) or strategic alliances, for the aim of providing a more agile and flexible service (Bradley et al. 1999).

The functions and role of maritime operators within the context of global logistics have been discussed in this section. The research reveals that in order to survive the intense competition in the maritime industry, maritime logistics operators have become proactively engaged in various types of co-operative relationships with their rivals, with the ultimate aim of pursuing common interests. These relationships are referred to as a co-opetitive network, in which companies are multiply connected by co-operating with their competitors (Brandenburger and Nalebuff 1995; Gnyawali and Madhavan 2001; Luo 2004; Tsai 2002). The effectiveness of the co-opetition strategy in acquiring useful information and knowledge from their competitors and then improving maritime logistics value will be detailed further in the following section.

16.5 Co-opetitive Networks, Knowledge Acquisition and Maritime Logistics Value

Maritime logistics operators can gain comprehensive knowledge-based advantages from the co-opetitive networks, which in turn may help to improve maritime logistics value. Co-opetitive network is referred to as a set of multiply linked relationship in which competition and co-operation simultaneously occur between competitors (Luo 2004; Tsai 2002). This section elaborates on how maritime logistic operators can acquire useful knowledge throughout the co-opetitive networks within which they are embedded, and how the acquired knowledge can facilitate the improvement of maritime logistics value.

16.5.1 Co-opetitive Networks and Knowledge Acquisition

Knowledge acquisition occurs when an organisation acquires a useful information or know-how from its transacting partners. Existing literature has addressed that co-opetitive networks may bring network entities a number of knowledge-acquisition advantages through the synthetic mechanism between competition- and co-operation-based relationships (Tsai 2002). For example, Tsai (2002) investigated the knowledge acquisition advantages within a multiunit organisation that can be gained through co-opetitive networks. Tsai (2002) indicates that the co-opetitive network promotes knowledge sharing between units of an organisation, which in turn may help to enhance firms' performance. Bernal et al. (2002) suggest that freight forwarders acquire knowledge by being embedded in their co-opetitive networks. Lado et al. (1997) also suggest that co-opetition promotes vigorous knowledge exchange between organisations.

Social network embeddedness perspective places emphasis on two inter-organisational coordination mechanisms, i.e. the structural and the relational embeddedness, which facilitate knowledge exchange among entities within a network (Burt 1992; Gulati 1998; Nahapiet and Ghoshal 1998). This structural embeddedness highlights the co-ordination mechanism on how many entities are inter-connected in a network. It points out that if a focal firm is embedded in a dense network by establishing many connections with other network players, the firm can gain more informational benefits than the entity who does not engage in the same behaviour (Gulati 1998). This is possible, because the greater number of ties firms have in a network may provide more chances to access other firms' knowledge, and this facilitates knowledge exchange between players in the network. The relational mechanism relates to how closely players in a network are interconnected with each other (Gulati 1998). Previous literature contends that strong relationships with other players in a network enable a focal firm to build up trust with other entities, which it may in turn promote in-depth, two-way communication, and facilitate the exchange of solid information between organisations (Granovetter 1985; Krackhardt 1992; Uzzi 1997; Gulati 1998; Rowley et al. 2000).

The knowledge acquisition practices of maritime logistics operators within a co-opetitive network can be thoroughly explained from the social network embeddedness perspective. For example, as maritime operators expand their business scope into global markets, they can build up new partnerships with their rivals in the forms of strategic alliances, joint ventures, associations and consortium, and various types of informal relations such as personal meetings, phone conversations, or emails. If a maritime operator has a great number of co-operative relations with their rivals in a network, they can gain further exposure to greater information flows in the network, and as a result they could share more knowledge about the industry, market, or the firms' own technology (Song and Lee 2012). Song and Lee (2012) empirically investigated these practices of the maritime logistics industry, and revealed that maritime logistics operators in Korea can access the common pool of knowledge and gain a great deal of information through establishing a great number of partnerships with their competitors. These contentions ensure that high numbers of ties within a co-opetitive network may facilitate the greater volume of knowledge acquisition of maritime logistics operators (Galaskiewicz 1979).

Maritime operators can acquire knowledge through the relational embeddedness mechanism within the co-opetitive network, i.e. the closeness of ties. For instance, if a maritime logistics operator establishes close relationships with other network players by interacting frequently with them and accumulating mutual trust with each other, they are more likely to open their minds and prevent any potential opportunistic behaviours among organisations. This would lead to network players proactively exchanging useful information and knowledge with each other (Lee and Song 2010). There is empirical evidence investigated by Song and Lee (2012), which reveals that maritime logistics operators in Korea, which keep close relationship with other network entities by frequently interacting via email or telephone and holding both formal and informal meetings, have shared useful information and knowledge about the industry, customers, and strategic behaviours of their competitors. These previous contentions may ensure that maritime operators with strong ties are more likely to share more information and knowledge with one another.

The positive association between the extent of co-operative relationship in a network and knowledge acquisition has been explored in the above. However, there's a contention that although the inter-organisational co-operative relationship may encourage firms to acquire knowledge from their competitors, its effectiveness may in turn be further affected by the extent of competition in the network (Tsai 2002). Competition per se may hamper vigorous knowledge exchange between competing organisations, because normally firms may hesitate to share with their rivals the useful knowledge which would help the competitors to enhance their competitive advantages. However, the competition within a co-opetitive network can facilitate vigorous knowledge sharing between organisations in the network, because the network entities cannot act independently as a sole player. Instead, their strategic behaviours are inevitably affected by the governance mechanisms of a network relationship such as mutual gain, reciprocity and reputation effect (Coleman 1988; Powell & Brantley 1992; Jones et al. 1997).

For instance, previous studies have argued that network entities are affected by social governance mechanisms that implicitly monitor and co-ordinate inter-organisational strategic behaviours by enforcing individual organisations to act in a way to increase common interest rather than pursuing their own benefit (Jones et al. 1997). This mechanism works as firms in a network are inter-dependent entities whose action may mutually affect each other. If a firm does not observe the social mechanism by attacking their rivals indiscriminately in order to pursue their own interest, the firm would fall into disrepute with the other players in the network, and as a result the firm would be disadvantaged due to the nonobservance of these social mechanisms. Thus, network governance mechanisms may force the competing actors in a network to act in a mutually helpful way, by encouraging them to proactively exchange useful assets and resources rather than completely protecting them (Song and Lee 2012).

Applying the above argument to maritime logistics operators, if a maritime operator competes intensively with one player and competes less with another player in a network, the firm may be more enthusiastic to acquire the knowledge of the one with whom they are in a more competitive relationship, in order to quickly determine new skills or know-how of the competitors and win the competition. But, under the social network governance mechanism, they may need to open their knowledge as much as they wish to acquire their competitors' knowledge. This tendency may be more apparent when the competition is more intense. Therefore, the knowledge acquisition of maritime operators, which is obtained through high numbers of and strong ties of co-opetitive network relationships, may be promoted more when the competition is high (Lado et al. 1997; Tsai 2002; Song and Lee 2012).

16.5.2 Knowledge Acquisition and Maritime Logistics Value

As discussed in the previous section, the key objective of maritime operators would be the maximisation of maritime logistic value. Knowledge acquisition would help maritime operators to improve their maritime logistics value. The key dimensions of maritime logistics value are reducing lead time and business costs (i.e. operational

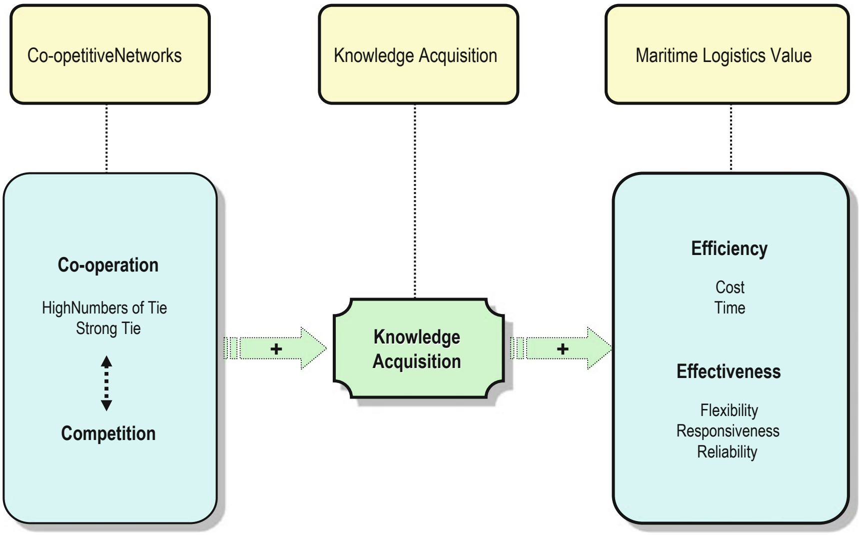


Fig. 16.5 Conceptual Framework. (Source: Song and Lee 2012, p. 20)

efficiency), and improving flexibility, responsiveness and reliability of their services (service effectiveness) (Lee and Song 2010). The knowledge that maritime operators acquire from their competitors may encompass useful information about markets, customers, and business environments of the industry, or other firms' firm-specific knowledge such as a firm's information technology, operational know-how and organisational skills. The knowledge may help maritime operators to reduce environmental uncertainty and to better understand customers' various demands. This may enable maritime operators to better reflect the voice of the customers on their operations and from there allow them to upgrade their service effectively. As a result the maritime operators can refine their service in a more responsive and flexible way, and then can build up trust from their customers.

By applying this newly acquired knowledge, maritime operators could develop new technologies and provide their customers with newer services. Maritime operators can also utilise the acquired knowledge differently according to their respective situations, and then create unique services which can make them more differentiated from their competitors. By making better use of the acquired knowledge, they can also develop new operational skills, and this may help them to eliminate wasteful activities so as to further save business costs and reduce lead time. This may in turn lead to an improvement in operational efficiency in maritime operations. These contentions are empirically supported by the existing study (Song and Lee 2012). This may therefore ensure that knowledge acquisition may have a positive influence on the improvement of maritime logistics value. The aforementioned theoretical contentions are described in Fig. 16.5.

16.6 Discussion and Conclusion

This chapter explored the current situations of competition and co-operation of maritime logistics operators and investigated the way to effectively manage inter-organisational co-ordination, which allows maritime operators to successfully acquire knowledge and maximise maritime logistics value. The principal mechanism of the inter-organisational coordination in this chapter is an extent of co-operation and competition within a co-opetitive network among maritime operators. The discussion ensures that the greater extent of co-operation in a co-opetitive network has a positive impact on knowledge acquisition. In particular, the higher numbers of and stronger relationships between entities in a network would be the central relational resource in facilitating knowledge acquisition of maritime operators. Given the positive relationship between co-operation in a network and knowledge acquisition, inter-organisational competition in the network promotes more vigorous knowledge sharing between the co-operating partners.

Furthermore, this chapter also argues that knowledge acquisition may have a positive impact on the enhancement of maritime logistics value through the improvement of operational efficiency and service effectiveness. This contention is supported by much of the previous works, indicating that knowledge acquisition contributes to the reduction of costs, price, operational time (i.e. efficiency) and the enhancement of firms' responsiveness, flexibility and reliability (i.e. effectiveness) (Nonaka 1994; Grant 1996; Li and Calantone 1998; Tsai 2001; Zhao et al. 2001). Thus, this study therefore ensures that successful co-ordination of inter-organisational relationships and inter-firm learning would be a key strategic tool towards the greater outperformance of maritime logistics operators.

In conclusion, this chapter provides a systematic review on the environmental challenges maritime operators have faced, the key strategic goal that the maritime operators should achieve, and essential strategy to be adopted for maritime logistics operators to achieve this strategic objective. This chapter may give a meaningful strategic insight into the effectiveness of knowledge-based strategy and the significance of strategic management of co-opetitive partnership in maritime logistics operations. Further, it is also believed that this study provides an interesting research agenda for academics to facilitate a further empirical discussion on such matters associated with the effectiveness of organisational learning and knowledge management strategy in the maritime logistics field.

References

Alderton, P. (1999). *Port management and operations, Lloyd's practical shipping guides*. London.

Baudin, M. (2004). *Lean logistics: The nuts and bolts of delivering materials and goods*. New York: Productivity Press.

Bernal, S. M. H., Burr, C., & Johnsen, R. E. (2002). Competitor networks: International competitiveness through collaboration. *International Journal of Entrepreneurial Behaviour and Research, 8*(5), 239–253.

Bowersox, D. J. (1978). *Logistics management: A systems integration of physical distribution management and material management*. New York: Collier Macmillan.

Bradley, P., Gooley, T., & Cooke, J. A. (1999). Forwarding industry poised to thrive. *Logistics Management and Distribution, 38*(8), 28.

Branch, A. E. (2007). *Elements of shipping* (8th ed.). New York: Routledge.

Brandenburger, A. M., & Nalebuff, B. J. (1995). The right game: Use game theory to shape strategy. *Harvard Business Review, 73*(4), 57–71.

Brooks, M. (2000). *Sea change in liner shipping: Regulation and managerial decision-making in a global industry*. New York: Pergamon.

Burt, R. S. (1992). *Structural holes: The social structure of competition*. Boston: Harvard University Press.

Busan Port Authority. (2007). *2007 Port of Busan container Statistics*. Busan: Busan Port Authority.

Christopher, M. (2005). *Logistics and supply chain management: Creating value-adding networks*. New York: Financial Times Prentice Hall.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94*(1), 95–120.

Containerisation International. (February 2013a).London: A Lioyd's List.

Containerisation International (2013b). *Top 100 ports 2013*. London: A Lioyd's List.

Coyle, J.J., Bardi, E.J., & Novack, R.A. (1999). *Transportation*. New York: South-Western College Publishing.

De Souza, G.A., Beresford, A.K.C., & Pettit S.J. (2003). Liner shipping companies and terminal operators: Internationalization or globalization? *Maritime Economics and Logistics, 5*(4), 393–412.

Frankel, E.G. (1982). *Management and operations of american shipping*. Boston: Auburn House Publishing Company.

Fremont, A. (2007). Global maritime networks, the case of Maersk. *Journal of Transport Geography, 15*(4), 431–442.

Galaskiewicz, J. (1979). *Exchange networks and community politics*. Beverly Hills: Sage Publication.

Global Logistics and Supply Chain Magazine. (2006). http://glscs.texterity.com/glscs/200606/?pg=60. Accessed 7 July 2009.

Gnyawali, D. R., & Madhavan, R. (2001). Cooperative networks and competitive dynamics: A structural embeddedness perspective. *Academy of Management Review, 26*(3), 431–445.

Grant, R.M. (1996). Prospering in dynamically-competitive environments organizational capability as knowledge integration. *Organization Science, 7*(4), 375–387.

Grantovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology, 91*(3), 481–510.

Gulati, R. (1998). Alliances and networks. *Strategic Management Journal, 19*(3), 397–420.

Haasis, H. D. (2008). Knowledge management in intermodal logistics networks. In H. D. Haasis, H. J. Kreowski, & B. Sholz-Reither (Eds.), *Dynamics in logistics*. Berlin: Springer.

Heaver, T., Meersman, H., Moglia, F., & Van de Voorde, E. (2000). Do mergers and alliances influence European shipping and port competition? *Maritime Policy and Management, 27*(4), 363–373.

Jones, C., Hesterly, W. S., & Borgatti, S. P. (1997). A general theory of network governance: exchange conditions and social mechanisms. *Academy of Management Review, 22*(4), 911–945.

Korean Fair Trade Commission. (2009). http://ftc.korea.kr/gonews/main.do. Accessed 3 July 2009.

Krackhardt, D. (1992). The strength of strong ties: The importance of philos in organizations. In N. Nohria & R. Eccles (Eds.), *Networks and organizations: structure, form, and action* (pp. 216–239). Boston: Harvard Business School Press.

Lado, A. A., Boyd, N. G., & Nalón, S. C. (1997). Competition, co-operation, and the search for economic rents: A syncretic model. *Academy of Management Review, 22*(1), 110–141.

Lai, K.-H., Ngai, E. W. T., & Cheng, T. C. E. (2002). Measures for evaluating supply chain performance in transport logistics. *Transport Research Part E, 38*(7), 439–456.

Lei, L., Fan, C., Boile, M., & Theofanis, S. (2008). Collaborative vs. non-collaborative container-vessel scheduling. *Transportation Research Part E, 44*(3), 504–520.

Lee, E-S. (2010). Knowledge acquisition and maritime logistics value: An inter-organisational relationship perspective, Unpublished PhD Thesis, Heriot-Watt University, Edinburgh.

Lee, E-S., & Song, D-W. (2010). Knowledge management for maritime logistics value: Discussing conceptual issues. *Maritime Policy and Management, 37*(6), 563–583.

Li, T., & Calantone, R.J. (1998). The impact of market knowledge competence on new product advantage: Conceptualization and empirical examination. *Journal of Marketing, 62*(4), 13–29.

Luo, Y. (2004). *Coopetition in international business*. Copenhagen: Copenhagen Business School Press.

Marlow, P. B., & Paixao, A. C. (2003). Measuring lean ports performance. *International Journal of Transport Management, 1*(3), 189–202.

Murphy, P. R., & Daley, J. M. (2001). Profiling International Freight Forwarders: An Update. *International Journal of Physical Distribution and Logistics, 31*(3), 152–168.

Nahapiet, J., & Ghoshal, S. (1998). Social capital, intellectual capital, and the organizational advantage. *Academy of Management Review, 23*(2), 242–266.

Nonaka, J. (1994), A dynamic theory of organizational knowledge creation. *Organization Science, 5*(1), 14–37.

Notteboom, T. E. (2004). Container shipping and ports: An Overview. *Review of Network Economics, 3*(2), 86–106.

Notteboom, T. E. (2006). The time factor in liner shipping services. *Maritime Economics and Logistics, 8*(1), 19–39.

Notteboom, T. E., & Winkelmans, W. (2001). Structural changes in logistics: How will port authorities face the challenge? *Maritime Policy and Management, 28*(1), 71–89.

O'Leary-Kelly, S. W., & Flores, B. E. (2002). The integration of manufacturing and marketing/sales decisions: Impact on organizational performance. *Journal of Operations Management, 20*(3), 221–240.

Panayides, P. M. (2006). Maritime logistics and global supply chains: Towards a research agenda, *Maritime Economics and Logistics, 8*(1), 3–18.

Panayides, P. M., & Gray, R. (1999). An empirical assessment of relational competitive advantage in professional ship management. *Maritime Policy and Management, 26*(2). 111–125.

Panayides, P. M., & Song, D.-W. (2008). Evaluating the integration of seaport container terminals in supply chains. *International Journal of Physical Distribution and Logistics Management, 38*(7), 562–584.

Powell, W. W., & Brantley, P. (1992). Competitive cooperation on biotechnology: learning through networks? In N. Nohria & R. Eccles (Eds.), *Networks and Organizations* (pp. 366–394). Boston: Harvard Business School Press.

Roh, H. S., Lalwani, C. S., & Naim, M. M. (2007). Modelling a port logistics process using the structured analysis and design technique. *International Journal of Logistics: Research and Applications, 10*(3), 283–302.

Rowley, T., Behrens, D., & Krackhardt, D. (2000). Redundant governance structures: An analysis of structural and relational embeddedness in the steel and semiconductor industries. *Strategic Management Journal, 21*(3), 369–386.

Rushton, A., & Walker, S. (2007). *International logistics and supply chain outsourcing: From local to global*. London: Kogan Page.

Shang, K. C., & Marlow, P. B. (2005). Logistics capability and performance in Taiwan's major manufacturing firms. *Transportation Research Part E, 41*(3), 217–234.

Slack, B., & Fremont, A. (2005). Transformation of port terminal operations: From the local to the global. *Transport Reviews, 25*(1), 117–130.

Song, D.-W. (2003). Port co-opetition in concept and practice. *Maritime Policy and Management, 30*(1), 29–44.

Song D-W., & Lee, E-S. (2012). Coopetitive networks, knowledge acquisition and maritime logistics value. *International Journal of Logistics Research and Applications, 15*(1), 15–35.

Tsai, W. (2001). Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *Academy of Management of Journal, 44*(5), 996–1004.

Tsai, W. (2002). Social structure of "Coopetition" within a multiunit organization: Coordination, competition, and intraorganizational knowledge sharing. *Organization Science, 13*(2), 179–190.

UNCTAD. (2013). *Review of maritime transport 2013*. New York: United Nations.

Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly, 42*(1), 35–67.

Waters, D. (2003). *Logistics: An introduction to supply chain management*. London: Palgrave.

World Bank. (2006). *Port reform toolkit: Effective decision support policy markets and practitioners, Module 3*. Washington: Public-Private Infrastructure Advisory Facility.

Zhao, M., Droge, C., & Stank, T. P. (2001). The effects of logistics capabilities on firm performance: Customer-focused versus information-focused capabilities. *Journal of Business Logistics, 22*(2), 91–107.

Chapter 17
Hinterland Transportation in Container Supply Chains

Yann Bouchery, Stefano Fazi and Jan C. Fransoo

Abstract The increase in traded container volumes worldwide puts pressure on the hinterland road network, leading congestion and emission problems. This leads to a requirement to develop intermodal transportation systems. In this chapter, we analyze the most important features of such container transportation systems for the hinterland supply chain. At the network design level, we review the current state of the art and we identify avenues for future research. Among others, we highlight that the coordination of container shipments across the container supply chain is a particularly relevant issue as hinterland networks involve several actors. At the operational level, we characterize the most important factors influencing the trade-off between intermodal transportation and truck-only deliveries. In addition, we provide a case study of coordination at an intermodal barge terminal in the Netherlands. We highlight that the exchange of information is the key enabler for efficient hinterland intermodal transportation and we show that a better information system can be of crucial importance.

17.1 Introduction

Over the last decades, traffic of containers has increased substantially. Growth in international trade leads to increased growth in transport, and due to extensive containerization of an ever-increasing number of commodities, container transport has grown substantially (Fransoo and Lee 2013). Apart from the growth in intercontinental maritime transport, also the container traffic in the associated hinterland has grown substantially. Transportation means such as barges, trains and trucks have been adapted to be able to transport containers to and from the deep sea ports.

The transport of containers involved many actors and activities along the supply chains. If we consider an intercontinental shipment of a container that includes an ocean leg, the process is usually initiated by a company (such as a manufacturer) that orders a container. As the empty container is received, it is loaded and then transported

---

J. C. Fransoo (✉) · Y. Bouchery · S. Fazi
School of Industrial Engineering, Technische Universiteit Eindhoven,
Eindhoven, The Netherlands
e-mail: J.C.Fransoo@tue.nl

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_17

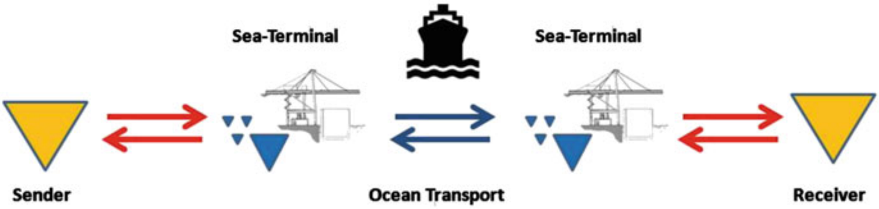


Fig. 17.1 A generic intercontinental shipment of a container. The *red* flows represent the hinterland transportation

to a terminal in a deep sea port, where it is handled and loaded onto a vessel. The container is then shipped to another seaport (potentially being transshipped along the way), discharged and delivered to a receiver (consignee), who unloads it. Finally, the container is delivered either to a depot for empties or to a deep sea terminal. Figure 17.1 contains a visual description of this cycle.

From this description we can identify 3 generic elements in this supply chain, namely:

- Ocean transport
- Terminal operations
- Hinterland transport

Each of these stages has been affected to a large extent by the increase of containerized transport.

In ocean transport, naval engineering pushed the physical limits of the ships remarkably. The latest ships provide capacity for more than 18,000 TEU (Kohr et al. 2013). Also civil engineering contributed to the ability to create economies of scale. As the size of the ships increased, also the infrastructures like canals, locks and docks had to be adapted. Artificial canals and expansions of current waterways have been made to allow more and larger ships to sail through. The case of the Panama Canal is one of the most well-known engineering processes of the artificial expansion of intercontinental waterways.

Deep sea terminals have generally been seen as the major element in the container transport chain affected by increasing container traffic, due to limited capacity in terms of storage spaces and handling equipment (Roso et al. 2009). A vast literature addressing this topic treats remedies for such limitations; from the optimization of crane scheduling to berth and yard allocation, see Vis and De Koster (2003), Kim (2005), Stahlbock and Voss (2008), and Bierwirth and Meisel (2010) for recent overviews. Besides making the best use of the available capacity, recently some deep sea ports are also expanding their premises; this requires certainly a bigger effort in terms of costs and time (Roso et al. 2009; Jarzemskis and Vasiliauskas 2010). For instance, in the Port of Rotterdam (Netherlands), a new extension of the Maasvlakte area, Maasvlakte 2, has been developed recently to allow more deep-water access. Approximately 2000 ha have been reclaimed behind a 4 km dike and approximately 1000 ha will be used by port related industries. Also with the latest larger vessels,

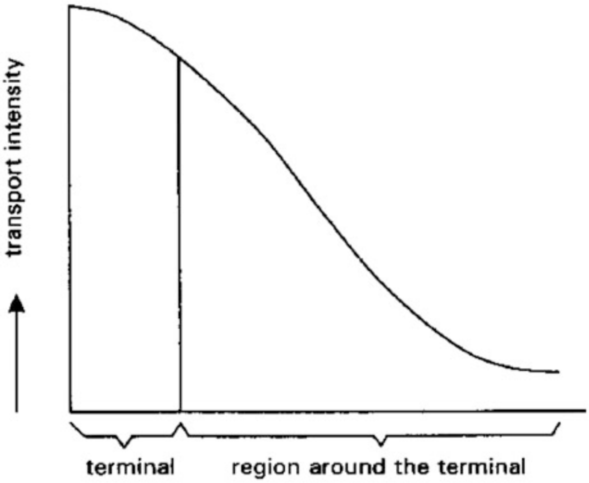


Fig. 17.2 Transport intensity. (Source: Konings 1996)

deep sea terminals have to catch up with the pace of technologies and global demands and face the challenge of handling more containers per time unit (Roso et al. 2009).

In recent years, hinterland transport has been receiving increased attention due to problems of road congestion, environmental concerns and traffic safety (Van Schijndel and Dinwoodie 2000). As the amount of containers handled by sea ports has been increasing substantially, the hinterland has been facing problems of congestion, especially in the area close to the sea port where the flows do not scatter yet (see Fig. 17.2).

The problem is very noticeable when the majority of containers are transported by truck. In Europe—especially in the North-Western area between Le Havre and Hamburg, where the flow of containers is the highest—the problem has become very relevant. Traffic congestion at the sea port and in the areas nearby is becoming unsustainable, and has drawn the attention of policy makers, shippers, and freight forwarders. Using different modalities than trucks is one of these (Van Schijndel and Dinwoodie 2000). Public authorities are pressing for the use of high capacity means of transport, in order to push large bundles far into the hinterland. Therefore, the use of trains and barges is favored by transport providers and policy makers alike, but criticized for the lack of flexibility (as compared to the high level of flexibility offered by trucks) and unfavorable cost structures. The so-called regionalization (Notteboom and Rodrigue 2005) of the sea ports is crucial in the success of multimodality. The idea is that connections between the deep sea port and the hinterland are strengthened, by means of inland terminals, strategically located in the region in order to ease the change of modality and the access of trains and barges (Parola and Sciomachen 2005). As the use of these means of transport can relieve congestion, regionalization is seen as alternative to sea port expansion as well. Moreover, the terminals should provide additional logistic services to make those alternative modes more profitable and competitive (Jarzemskis and Vasiliauskas 2010; Roso 2007; Notteboom 2007).

The design and coordination of these systems that connect deep sea ports and inland terminals is the topic of this chapter. In our description and analysis, we will focus on import hinterland networks, mainly motivated by the hinterland network of the Port of Rotterdam. Similar (and partially overlapping) hinterland networks exist for other ports in Northwestern Europe, and ideas and concepts could be transferred (albeit adaptation may be necessary) to hinterland networks that are primarily export oriented (such as in China, or other import hinterland networks such as in North America and Africa).

17.2 The Hinterland System and Its Evolution

The increase of container trade as well as the logistic integration and network orientation in the port and maritime industry have redefined the role of sea ports and the approach to the hinterland (Notteboom and Rodrigue 2005). The total cost of global supply chain transportation is to a large extent affected by the efficiency of the hinterland transportation system. According to Notteboom (2004), 40–80 % of the total transportation costs are the cost of hinterland transportation. The improvement in logistics and transportation in this leg has therefore a large impact on the final cost for the customer.

As discussed above, a regionalization phase is currently evolving with the strengthening of the connections between deep sea ports and their hinterland. The role of intermediary transshipment centers as inland terminals or hub-and-spoke terminals is part of the so-called regionalization that brings the perspective of port development beyond the port perimeter (Notteboom and Rodrigue 2005).The targets of the regionalization are many, including the deployment of other concepts such as "dry port". Dry ports facilitate pushing large quantities of cargo, as soon as they land, far into the hinterland, where multiservice hubs would replicate the services of sea ports, such as customs, handling, and storage (Jarzemskis and Vasiliauskas 2010; Roso et al. 2009). Further, sea ports—by connecting with inland hubs—try to reduce congestion and pollution (Roso 2007) and dwelling times of containers.

As hinterland transportation is evolving, also the role of inland terminals is becoming more and more important. Over time, inland terminals have been increasing their role in the deployment of multimodal transport. The locations and the services offered by the inland terminals can be the added value that makes high capacity means of transport, such as barges and trains, more cost-effective and also more convenient than trucks (Notteboom and Rodrigue 2009a).

Konings (1996) predicted early on the current criticalities in hinterland transport due to extensive trucking and claimed that the key of success is in integrated centers for transshipment, anticipating the concept of dry ports: "*The demand in container transport increases and clients want their goods delivered faster, cheaper and just-in-time. Road transport and its network would increasingly become unable to meet the demand and the quality criteria of the clients. The consequence would be that road transport becomes more expensive, less sustainable, more time consuming and less attractive*."

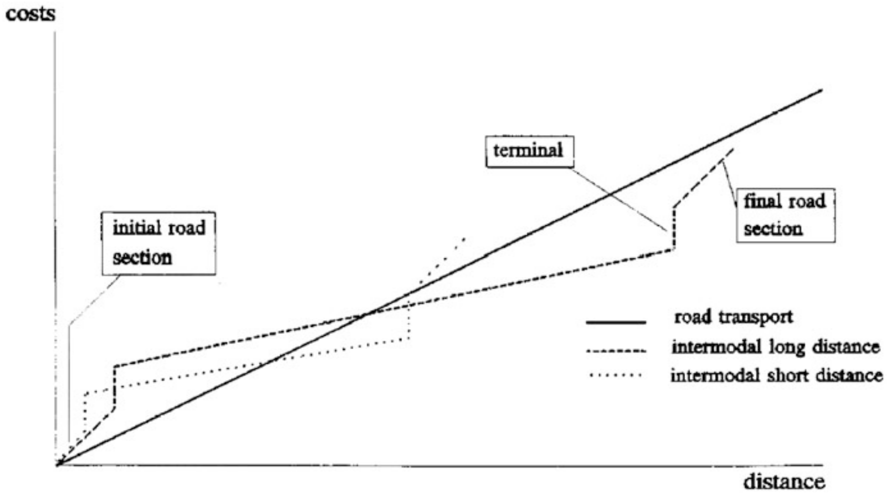


Fig. 17.3 Cost structure of unimodal road haulage versus multimodal transport. (Source: Konings 1996)

The shift to barge and rail transport can be the solution. Nevertheless, their advantages are often limited by the added costs for handling and initial and final road transport legs. These additional costs are relatively high it is difficult to absorb these costs over shorter distances and smaller volumes (see Fig. 17.3).

As road transport prices increase, so will these additional costs and combined transport will be affected proportionally, unless initial and final road transport and its cost can be somehow limited or cancelled out. In fact, it is clear that the growth of intermodal transport can entail an increase of trucking immediately around the sea-terminal, threatening the accessibility and the sustainability of the terminal itself. In a sense, as stated by Konings, multimodal transport can fall victim of its own success without smart logistic structures behind as: collection centers, new terminals, improved internal transportation. Such logistic centers have to be located both where trucking cannot affect in large part the success of the intermodality (hub-and-spoke centers in the sea port (Konings et al. 2013)) and where trucking becomes less competitive than barge and rail (faraway terminals with dry port concept (Roso et al. 2009)). We will now briefly discuss the hub-and-spoke and dry port concepts.

17.2.1 The Hub-and-Spoke System Within the Sea Terminals

A Hub-and-Spoke philosophy entails the bundling of containers in a hot-spot (Fig. 17.4). As claimed by Konings et al. (2013), a hub-and-spoke network would transform the situation in the sea port—currently characterized by separate collection and distribution centers (Caris et al. 2011)—into a system where bundles of containers for a pre-determined set of hinterland destination are gathered in one terminal. This can entail positive aspects, both for rail and barge.

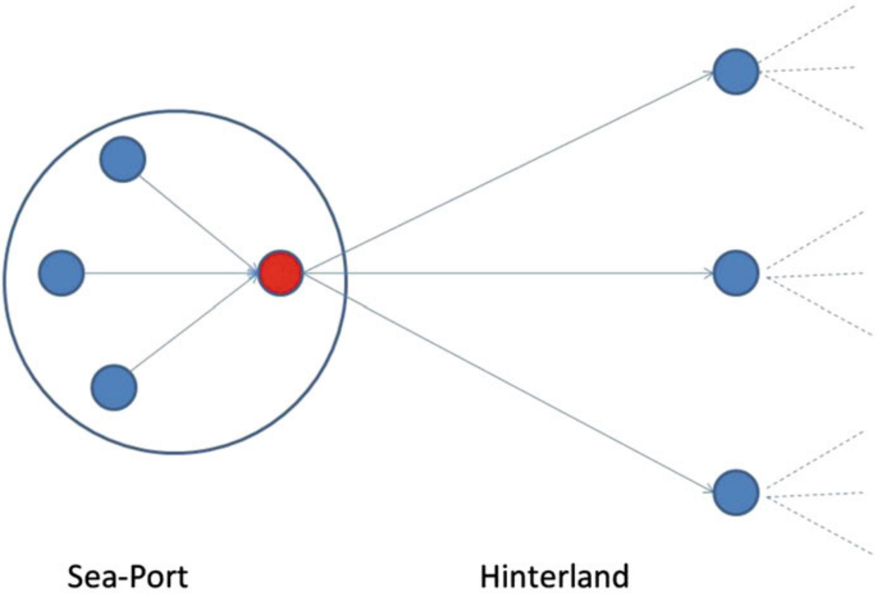


Fig. 17.4 The conceptual hub-and-spoke model, consolidating containers in a hub that is located in the deep sea port area

For hinterland transport by rail, the internal handling and transportation system of the sea port, where usually trucks but in some cases also AGVs are used, would gather containers at a central hub dedicated to rail, not far from the sea terminals.

For hinterland transport by barge, currently barges waste valuable time by making many calls at the different terminals of the deep sea ports; this is due mainly because containers meant for the same destination can arrive with different ocean vessels docking at different terminals and quays (Caris et al. 2011). Intrinsically, a route of a barge with many calls already entails a bigger effort in terms of itinerary and setup needed for the terminal operators and the cranes. The latter, when operating small size calls, do not work continuously, making the system not efficient. Moreover, long waiting times and congestion in a system with many calls is inevitable. One of the main reasons is that ocean vessels have priority over barges. Deep sea careers have usually contracts with terminal operators while barge operators do not (Fransoo and Lee 2013). That is why barge operators insert large margins in their schedules when planning terminal visits, in order to not compromise the reliability of the service. However, the competitiveness of the service is somehow undermined by such long waiting times and loose schedules. Also here, a hub-and-spoke system can be the solution to reduce the number of calls by barges and the waiting times and as a direct consequence the reliability of the service (Konings 2013; Fransoo et al. 2013).

There is an increasing interest in this topic. Designing and operating a hub-and-spoke system implies decisions on its location (Rutten 1998; Bouchery and Fransoo 2014), bundling strategies (Jourquin et al. 1999; Caris et al. 2011) and vertical integration with inland terminals (Veenstra et al. 2012). With concern to this last point, inland terminals located in the same region can co-operate along a hinterland corridor to bundle their freight. An example is a co-operation recently established in the Brabant region, in the Netherlands. *Brabant Intermodal* includes four terminals that coordinate and bundle their shipments in order to make larger call sizes in the sea port. Then the bundles consolidate in the port area and split on the way to the different inland terminals. We will further discuss the hub-and-spoke networks in Sect. 3 of this Chapter. An extensive description of the cost structure of the hub-and-spoke system in barge transportation has been described and analyzed by Fransoo et al. (2013).

17.2.2 The Dry Port Concept and Its Evolution: The Extended Gate

The dry port concept has been successfully implemented in various geographical contexts. Dry ports relieve congestion at sea terminals and the port city, by bundling large quantities of containers and pushing them to the hinterland toward equipped intermodal terminals. The use of high capacity modes is then further encouraged (Roso et al. 2009). Roso et al. (2009) investigate the pros and cons of dry ports when these are positioned on the short, medium and long distance. Practically, the dry ports have to be considered as a natural extension of the sea ports, whose common activities are replicated: stacking, custom clearance, and handling. Some port authorities, such as the Port of Barcelona, have been very active in developing the hinterland hub-and-spoke network (Van den Berg et al. (2012).

The Extended Gate is an evolution of the dry port concept, as it includes some of the natural consequences, such as integral network design and direct operational control in the transport network between the sea terminal and the dry port (Veenstra et al. 2012). The main idea of an Extended Gate is to extend the delivery point from the perspective of the shipper/receiver from the deep sea terminal along a corridor to an inland multimodal terminal and, possibly, the final destination such as a distribution center of a logistics service provider or shipping (Veenstra et al. 2012). The gate of the sea terminal is basically moved to the inland terminal. Receivers agree to pick up their containers at the inland terminal; the final leg of the journey can be arranged by the terminal or another operator. Therefore, the receiver can deal with a wide variety of inland terminals close to multiple delivery locations rather than with the sea terminal. This delivery at the inland terminal is typically offered as an additional premium service to the customer, such as the Extended Gate service provided by the Rotterdam-based Europe Container Terminals. Inland terminals strategically located in economic centers are the most suitable for this purpose, because they could facilitate the flow of import and export containers and the flow of intercontinental cargo as well.

As stated by Veenstra et al. (2012), one of the crucial factors for the success of such efficient hinterland networks in Europe is the availability of the right information on containers that arrive from overseas. This includes information on the nature of the transported goods, their quality, safety and other handling instructions, destination, shipper, receiver, intended mode of hinterland transport, and required arrival date and time. Currently, such information is not regularly available to container terminal operators or hinterland transport operators, until the very last moment. Information is usually in the hands of freight forwarders, and of the owners of the goods. Limited research has so far been conducted on the value of information in container transport networks. Early work by Zuidwijk and Veenstra demonstrates that information can be valuable under certain conditions (Zuidwijk and Veenstra 2014).

A possible solution for a smooth and real time exchange and update of information can be the introduction of a collection agent that collects and organizes the information and guarantees their availability online. Currently, such systems are being implemented, especially in those ports bringing forth the Extended Gate concept. For instance, ECT, a container terminal operator in the port of Rotterdam, is driving the deployment of the Extended Gate, by developing strong connections with a network of inland terminals. The exchange of information is managed by a new business entity created by ECT. The system is a web-based application and available to the inland terminals involved in the transportation.

17.2.3 The Transportation

Hinterland transport has recently attracted the attention of authorities and stakeholders. In fact, origins and destinations of virtually all intercontinental flows are always situated in the hinterland. However, they merge in the areas close to the ports, contributing to traffic jams, pollution and congestion (Van Schijndel and Dinwoodie 2000). These problems are more and more considered by Port Authorities, especially in the areas where the majority of the cargo is moved by trucks. The option of multimodal transportation has been acknowledged as priority measure to solve EU transport system problems (Jarzemskis and Vasiliauskas 2010) and to reduce CO_2 emissions (Liao et al. 2009; Roso 2007). As an example, in the Port of Rotterdam area new regulations are very strict and they constrain terminal operators as part of the concession to transport a high percentage of the containers by barges and trains. The target for the 2035 so-called "modal split" is to transport at least 45 % of the volumes by barge, at least 20 % by train and at most 35 % by truck (Konings et al. 2013). Interestingly, this target is part of the lease-contract with the terminal operators. Earlier, Fransoo and Lee (2013) noticed that the terminal operators do not hold contractual relationships with the hinterland transport operators. Since most of the inbound transportation in Europe is under a merchant-haulage contract, it will be interesting to see how the terminal operators can influence their modal split.

17.3 Design of Hinterland Networks

Transportation of containerized cargo has become increasingly popular in recent years. Container transportation indeed enables quick and undamaged arrival of the entire shipment to destination. As a result, world container traffic has been growing at almost three times world gross domestic product growth since the early 1990s (Fransoo and Lee 2013). This trend towards containerization is shaping the design of hinterland freight transportation networks. When focusing on hinterland supply chains, containerization provides protection against damage and theft and standardization. But the most important feature related to the transportation of containerized cargo is that the entire load may be handled in one time. Two main trends in hinterland networks may be associated with this ease of handling. First, containerization favors consolidation systems such as hub-and-spoke networks. Second, containerization facilitates intermodal transportation. These two main concepts are reviewed in Sect. 3.1 and 3.2. Then, Sect. 3.3 focuses on future research opportunities.

17.3.1 Strategic Models for Flow Consolidation in Hinterland Networks

Several hinterland transportation network topologies may enable flow consolidation compared to direct shipment from origin to destination (e.g. corridor, connected hubs, and hub and spoke networks). The most commonly used network topology is by far the hub-and-spoke system.

The first papers focusing on hub-and-spoke networks design problems may be traced back to 1986 (O'Kelly 1986a, b). From these first papers, the literature on hub location has expanded very quickly. We refer to Alumur and Kara (2008), Campbell and O'Kelly (2012) and Farahani et al. (2013) for recent reviews. Several classical formulations of the hub location problem may be found in the literature (i.e. hub median, hub center, hub covering, and hub location) but the most commonly used model for hinterland transportation network design is the p-hub median problem.

The p-hub median problem consists in finding the most appropriate way to transport demand flows from origins to destinations. To do so, the problem consists in locating $p \in N*$ hubs and to decide how to allocate these hubs to the set of origin/destination nodes in order to minimize the total transportation costs of the system. The demand from origin to destination may correspond to an amount of freight (freight transportation network), passengers (passenger air transportation network) or data (telecommunication network). In the basic settings of this static deterministic problem, the demand flow has to be routed from origin to destination through at least one hub. Routing demand flow via a hub indeed enables flow consolidation. This consolidation may occur on spokes (the arcs connecting origin/destination nodes to hub nodes) as the flow from one origin to several destinations may be combined if routed via the same hub. However, this consolidation is often much stronger on inter-hub arcs. The underlying benefit of this consolidation is that the cost of transporting one unit of flow per unit distance is discounted on inter-hub arcs. This creates an incentive to route origin/destination flows through more than one hub as this will increase the distance travelled but may lead to cost reduction. The hubs are generally assumed to be fully interconnected; thus, there is no reason for routing any origin/destination flow through more than two hubs.

The most important feature of the p-hub median problem is that the location of the hubs and the allocation of these hubs to origin/destination are two inter-related questions. As the allocation part of the problem (where the locations of hubs are fixed) is known to be NP-hard, the p-hub median problem is also NP-hard. However, a lot of research has been conducted to find efficient ways of solving the problem (either optimally or by using heuristic approaches) and large scale problems are nowadays solved very efficiently. The p-hub median problem may be considered as an extension of the p-median problem proposed for facility location (Hakimi 1964, 1965) that takes interdependency between facilities into account. As for the facility location research, a tremendous number of extensions from the basic models have been investigated. Reviewing the entire literature on hub location is outside of the scope of this chapter. Note that some extensions are reviewed in Sect. 3.3 as these ones are of great interest for future research on hinterland network design problems.

Even if the advantages of consolidation are undeniable in terms of cost per unit of flow per unit of distance, using a hub-and-spoke network implies increasing the traveled distance. The comparison of direct versus terminal (equivalent to hub) freight routing has been extensively studied (see e.g. Blumenfeld et al. (1985); Campbell (1990); Daganzo (1987); Hall 1987a, b). These papers focus on continuous approximation models in order to obtain analytical formulations. Moreover, only the allocation decisions are made. Even though approximating the demand for transportation as continuous over a region (by using density) may be viewed as unrealistic, these models are mainly used to provide insights and guidelines. Moreover, these models are proven to be quite robust when used to approximate the optimal transportation cost for discrete demand hub location problems (Campbell 1993). Numerical optimization and continuous approximation methods could thus be viewed as complementary and should be used together (Smilowitz and Daganzo 2007). We refer to Langevin et al. (1996) for a review of models applying the continuous approximation.

17.3.2 Design of Intermodal Hinterland Networks

Intermodal freight transportation implies transporting the load from origin to destination in the same transportation unit without handling of the goods themselves when changing modes (Crainic and Kim 2007). The main characteristic of intermodal transportation is the use of more than one mode of transportation. This feature also corresponds to other terminologies proposed in the literature such as multimodal or co-modal transportation. Even if the definitions are slightly different from one terminology to another, these terms are often used interchangeably

(SteadieSeifi et al. 2014). Intermodalism is clearly related to containerization as the containers are the most common transportation unit used for intermodal transportation. When focusing on hinterland networks, intermodal transportation and hub-and-spoke networks are two interrelated concepts as they share the issue of flow consolidation. We refer to Bontekoning et al. (2004) for review on the early development of the research on intermodal transportation for hinterland supply chains. The remaining of this section focuses on reviewing the most recent literature related to this issue.

Arnold et al. (2004) model an intermodal rail-road system as a hub-and-spoke network. Intermodal terminals are considered as hubs that have to be located. They apply their model to the rail-road transportation system in the Iberian Peninsula. As the size of the real world application is a concern, they propose a heuristic procedure to find an approximate solution. Racunica and Wynter (2005) extend the p-hub median problem to take into account classical features of intermodal rail-road transportation. Among others, the model accounts for non-linear concave cost functions in order to represent flow dependent economies of scale (see Sect. 3.3 for a detailed discussion on this topic). The authors propose a linearization procedure as well as two heuristics that enable solving large instances of the problem. A case study based on data from the Alpine region is also presented. Groothedde et al. (2005) focus on a road-barge transportation network in the Netherlands and explain how collaborative intermodal hub networks may be developed. Similar to Racunica and Wynter (2005), Jeong et al. (2007) extend the p-hub median problem to take some specificities of an intermodal rail-road network into account. In the same context, Limbourg and Jourquin (2009) specifically focus on the effect of considering flow dependent economies of scale in transshipment cost. Their study mainly focuses on a case study based on the European rail-road network. Ishfaq and Sox (2010) use a tabu search based meta-heuristic to solve another extension of the p-hub median problem in the context of intermodal transportation. They conduct an empirical study based on US freight data. Meng and Wang (2011) specifically include multi-type containers as well as user equilibrium constraints on a hub-and-spoke model for intermodal transportation. These user equilibrium constraints are intended to model the behavior of the users of the network who are willing to optimize their individual cost in making their route choice. This feature is discussed in more detail in Sect. 3.3. Alumur et al. (2012a) focus on including customer de-pendent service time on the intermodal p-hub median problem and consider travel times in addition to travel costs. The authors study the structural properties of the problem and propose a set of valid inequalities and a heuristic that enable to efficiently solve large in-stance of the problem. Finally, Alumur et al. (2012b) propose to apply an extension of the hub median problem called the hierarchical hub median problem Yaman (2009) to represent intermodal logistic networks. This literature mainly aims at bridging the gap between the theoretical p-hub median problem (and its extensions) with the current practices in real life intermodal transportation networks. By following this line of research, several future research directions based on current developments of hinterland networks are highlighted in the next section.

17.3.3 Future Research Opportunities

Several research directions may be emphasized based on the existing literature as well as the key trends in current hinterland transportation networks. Among them, the most straightforward feature of intermodal transportation is that economies of scale are necessary to make rail and barge transportation viable. Economies of scale is the "raison d'ê tre" of hub-and-spoke networks. However, the basic p-hub median model (as well as much of its extensions) assumes that economies of scale are somehow exogenous to the decisions taken on hub location and on origin/destination allocation. A fixed discount factor is generally used to account for economies of scale on inter-hub arcs. This hypothesis of considering economies of scale as independent of the decisions taken in the p-hub median problem is rather questionable as the volume of cargo transported on each arc strongly depends on the decisions taken. This limitation has been firstly addressed by O'Kelly and Bryan (1998) who account for flow dependent economies of scale on inter-hub arcs by considering strictly increasing concave transportation cost functions. They prove that the optimal hub locations may greatly differ from the results obtained without taking flow dependent economies of scale into account. Moreover, they propose a linearization technique to approximate the general concave strictly in-creasing function as piecewise linear. Note that several other papers have built on this idea. However, only few papers consider flow dependent economies of scale when focusing on intermodal hub location models. Racunica and Wynter (2005) take flow dependent economies of scale on each arcs (both inter-hub arcs and spokes) into account. Limbourg and Jourquin (2009) also account for flow dependent economies of scale but they take this feature into account only for transloading operations. Finally, Meng and Wang (2011) account for flow dependent economies and diseconomies of scale on inter-hub arcs in order to additionally take congestion into account. We are not aware of any paper considering flow dependent economies (and diseconomies) of scale for both transportation and terminal activities in the context of intermodal hub location problem. More research is required on this topic.

The second key feature in current hinterland networks is that several actors are generally involved. Indeed, various organizations generally control a part of the hinterland transportation chain, with no single-actor fulfilling the role of chain leader (Bontekoning et al. 2004). This idea is clearly in contradiction with the classical settings of the hub location literature as the objective of most of the models is to minimize the total transportation costs incurred in the system. Several other objectives may be chosen depending on the position of the actor in the supply chain. For instance, a terminal operator may follow the objective of maximizing the hubs utilization. A barge or rail service provider may tend to favor solutions with a high utilization of the intermodal service while a policy maker would certainly favor solutions that maximize the modal shift (Arnold et al. 2004). However, there are no theoretical reasons that allow considering that these objectives would lead to the same solution in a general setting. Taking several objectives into account is not classical in the hub location literature. To our knowledge, da Graça Costa et al. (2008) is one of the only papers proposing a multiobjective formulation of a hub location problem (the model account for both cost and time and both objective are taken into account separately). Another way taking multiple actors into account would consist in adopting a game theoretic perspective. To our knowledge, game theory has mainly been employed to account for competition among several independent hub-and-spoke networks (Lin and Lee 2010; Lüer-Villagra and Marianov 2013). We can conclude that the multiple actors setting of typical hinterland networks is not appropriately taken into account in the existing literature and that multiobjective optimization may be seen as a promising path for future research.

While combining flow dependent economies of scale with a multiple actors setting, interesting new allocation sub-problems may arise. As already pointed out by O'Kelly and Bryan (1998), "some origin-destination pairs may be routed via a path that is not their least-cost path because doing so will minimize total network travel cost". Two comments may be derived. First, the question of how to allocate costs if the objective is to minimize the total costs incurred in the network is of interest. We refer to Skorin-Kapov and Skorin-Kapov (2005) and to Skorin-Kapov (1998) for relating issues, even if the way of taking flow dependent economies of scale is somehow different in these papers. Second, O'Kelly and Bryan's statement may not hold if several actors act independently. The situation is similar to a classical problem in the traffic assignment literature. Due to congestion, the solution which minimizes the total traveling time in the system is not equivalent to the solution minimizing the travel times of each individual users. This leads to two extreme behaviors described by Wardrop (1952) as user equilibrium (where each user minimizes its own travel time) versus system optimum (where the total travel time of the system is minimized). Most of the existing literature on hub location assumes that the system optimum principle holds. Only one recent paper applies user equilibrium principle by constraining each origin-destination pair to be allocated to its lowest cost path (Meng and Wang, 2011). From our discussions with managers in industry, we would argue that classical intermodal networks do not follow any of these two principles. Indeed, several logistics service providers often use the same intermodal hub-and-spoke network they aim at optimizing their own transportation costs. Thus, real systems are generally sub-system optimal, meaning that each sub-system is independently optimized. More research is needed to account for this special feature of intermodal transportation networks.

As pointed out in the introduction of this section, containerization has some major advantages for hinterland transportation as this favors hub-and-spoke networks and intermodalism. However, we need to keep in mind that containerization has drastically modified the management of hinterlands transportation networks by raising new issues. Indeed, container transportation requires sending the container back to the shipper when the cargo has been delivered. This could be efficiently done by finding an export match. However, this is a challenging task for two main reasons. First, the import and export flows are often unbalanced, implying that some containers need to be sent back empty. Second, the containers belong to a particular shipping line who aims at reusing the container as quickly as possible (the shipping lines are charging detention fees if the containers are not sent back after a definite time limit).

Empty container management is thus a crucial issue that has deserved a lot of research (see e.g. Crainic et al. (1993)). Moreover, due to the preeminent role of ocean transportation in global supply chains (Fransoo and Lee 2013), the most widely used containers are specifically designed for ocean shipping. Thus, they are not optimized for hinterland transportation. For instance in Europe, conventional 40ft containers may contain up to 26 euro pallets instead of 33 euro pallets for conventional trailers. Even if these drawbacks seem to be counterbalanced by the advantages resulting from containerization, further improvements in hinterland container transportation efficiency may be obtained by implementing innovative solutions. Among these solutions, several projects in the Netherlands focus on assessing if cross-docking in the port area (Mangan et al. 2008) or at an inland terminal (Notteboom and Rodrigue 2005) may be valuable for optimizing hinterland networks. In general, the cross-docking activity takes place at the retailer's distribution center further down-stream in the supply chain. The idea of cross-docking upstream in the hinterland supply chain is to empty the maritime containers and to use special types of containers designed for hinterland transportation for delivering the cargo to the final customer. By doing so, multiple items may be loaded within the same container and this could improve the efficiency of hinterland transportation systems. Such innovative ideas may deserve future research as they are promising from an industrial perspective as well as challenging from an academic perspective.

Finally, it is very striking to note that eventhough intermodal transportation is generally claimed to be environmentally friendly in the introduction of the papers, no further investigations are conducted on assessing the environmental impacts of intermodal transportation. To our knowledge, Craig et al. (2013) is the only published paper focusing on this issue. Further research is definitely needed.

As a conclusion of this section, we would like to emphasize that several methodologies would be required to appropriately address the highlighted new issues. Indeed, even if mixed-integer programming techniques would continue to be very useful to solve real life problems, single hub formulations as well as continuous approximation models should supplement mathematical programming techniques and may help gaining better insights on these challenging issues.

17.4 The Trade-Offs in Multimodal Transport Operations

In North-West Europe, excellent waterway networks favor the use of barges. For instance, in the Ports of Rotterdam and Antwerp, in 2010, respectively 33 and 34 % of the total volume were handled by barge. These volumes compared to the share of container barging in other ports (Hamburg 1 %, Le Havre 7 %) are remarkable. In the Netherlands, high-quality waterways guarantee the success of the deployment of this modality by offering access to major industrialareas in Germany (Konings et al. 2013).

Although other regions are less favored by the geographical conditions, container barging can still be considered as a valid option. Fremont et al. (2010), describe the particular condition of the Port of Le Havre in France and its connection to the Paris region. The Port of Le Havre does not handle volumes that are as high as those in Antwerp and Rotterdam, and therefore high-frequency multimodal services may encounter difficulties to obtain cost effectiveness. However, Fremont et al. (2010) state that even under these unfavorable conditions, multimodality can compete with trucks, especially when these deliver an empty container in one transportation leg. However, the sole cost structure cannot guarantee the success of multimodality as the distances are short. They emphasize the fact that additional changes are of importance to make rail and barge transport more attractive. A first required additional change is to offer more flexibility with regard to dwelling times. In fact, the use of high capacity modalities can trigger overstays at the sea terminal premises. For instance, shipping lines could extend the periods of free demurrage and detention when the container is delivered by barge or train. A second additional change is to provide customs facilities to shippers. In France, for import flows, French customs and some multimodal operators made deals to ease the customs procedures for such flows. Fremont et al. (2010 claim that shippers can almost wait until their products are sold at the outlet before paying the customs. Therefore, such additional services, which are not provided by road transport, can make the difference. Strangely enough, the time factor can be on the side of multimodal transport.

17.4.1 Operational Decision Processes

For transport providers, the operational decision of choosing certain modes of transport is not only a matter of costs. When the positions of hubs, inland terminals and dry-ports are defined, the trade-off of choosing high capacity modes rather than trucks can be difficult to resolve.

In general a transport planner has to consider several features concerning both the fleet and the containers; the decision is made according to the available information. As mentioned earlier in this Chapter, information is generally not shared or becomes only available gradually over time. As the system is highly dynamic and the exchange of information between sea terminals and inland terminal planning systems is usually not in real time, the planner has to face also critical decisions in a rolling horizon manner. Planners can make their decision at any point in time; they can wait until more data becomes available in the system. For instance, new containers can become available at the deep sea terminal and therefore be ready to be picked up. As a consequence, the schedules may change until a certain moment, when a planner has to confirm a final schedule. The final decision is usually required when scheduled containers cannot wait any more time to be processed. After the decision is made, there is a time range where it is still possible to slightly modify the decision. As an example, assume that a barge leaves the inland terminal at time 0 and that it takes 12 h to reach the sea terminal. The planner can make a call (appointment) to pick up containers at the sea terminal with some margin. In the port of Rotterdam, a call can be made at most 2–3 h before the barge arrives to pick up containers. Then, it is clear

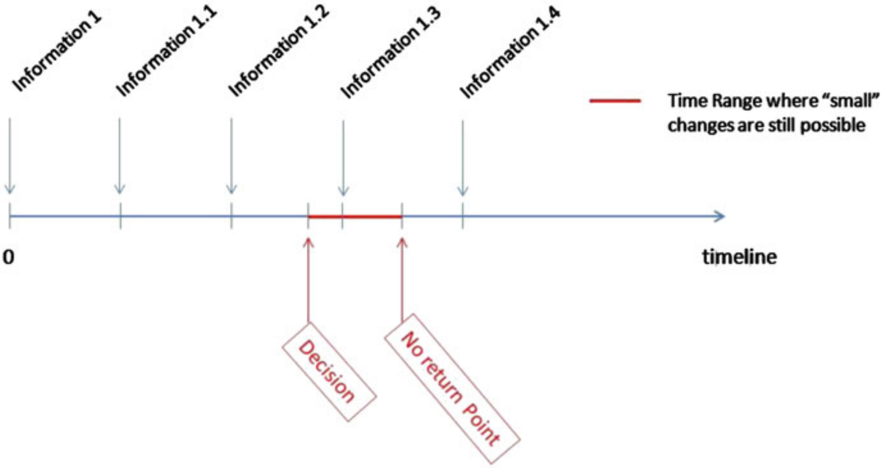


Fig. 17.5 Dynamic updating: after a schedule is made, small changes are still possible

that planners can have a margin of 9–10 h, where they can adapt the schedules to the newly available information and in that case pick up more containers. See Fig. 17.5 for a graphical explanation.

A similar problematic was tackled by Fazi (2014). In their paper, they add time features to the classical variable-size bin packing problem. Bins of larger size (barges) are able to process items slower than small bins (trucks), but they can generate economies of scale. The authors generated instances with data affected by variability. Numerical experiments showed that in case of uncertainty in the availability of the items (containers), planners are pushed to use more small bins (trucks) to face variability. In parallel, in the real system, planners have to deal with uncertain information and trucks are the only way to face the variability in the system.

With regard to information about containers, planners usually consider several features to come up with a final schedule. These are:

- Size
- Availability
- Initial location and destination
- Deadline
- Demurrage and detention deadlines
- Closing dates
- Status of the container: empty/full
- Particular requests of the customer

The status of the container can affect the decision of the planner. If the container is empty, the trade-off needs to be made whether to send back the container, incurring transportation costs but avoiding detention costs, or whether to keep the container until an export load becomes available, avoiding the transportation costs of (possibly twice) sending an empty container but incurring detention costs.

The fleet is usually limited in terms of capacity and availability. Planners have to consider in their schedules the following features of the fleet:

- Current and future availability
- Cost
- Capacity
- Transportation time
- Utilization
- Contract with the carrier

Utilization of the means of transport is highly related to whether the vehicle is owned or is under a specific contract. When the vehicle is owned, ideally the planner follows the strategy of the company: high frequency service vs. efficient use of the vehicle. When the vehicle is not owned, then the specific contracts play a role in the usage. When a vehicle is paid per ride, the planner tries to fill the capacity. When a vehicle is paid at a flat rate, then the planner may want to increase the frequency of the service by letting vehicle to be not fully utilized.

A case study that can explain the trade-off between barges and trucks in this planning process is presented in the next section.

17.4.2 A Case Study

In the Netherlands many inland terminals have become transport providers and offer barge and truck service for import and export container flows. In the Brabant region, canals have still relatively small size and allow the sailing of barges of 28 TEU capacity. Despite this small size and the short 120 km distance from the Port of Rotterdam, container barging can provide economies of scale, on condition that barges are fully utilized for both legs. Considering solely the rough transportation costs, for each leg an amount of six 2 TEU containers on a barge can compete with the cost of six trucks sent either from or to the port, based on our calculations for the Veghel Inland Terminal.

The system includes import and export containers. For import containers, the inland terminal typically has information on their arrival time at the sea terminal, the quay they are located, the time windows when is possible to pick them up and the deadline at the customer site. Analogously, export containers are available at the inland terminal and need to be delivered to the sea terminal. When these are packed with goods, usually the delivery has to take place before a closing date, which is the departure time of the ocean vessel. Otherwise, they just need to be repositioned to a depot for empties before the end of the detention period. According to these available information, the operational decision making process occurs.

Planners decide the allocation of containers to the fleet, the schedules and the routes of the barges. The terminal is dealing mainly with the Port of Rotterdam area and two main sea terminals areas: the Maasvlakte and the Rotterdam city terminal, see Fig. 17.6.

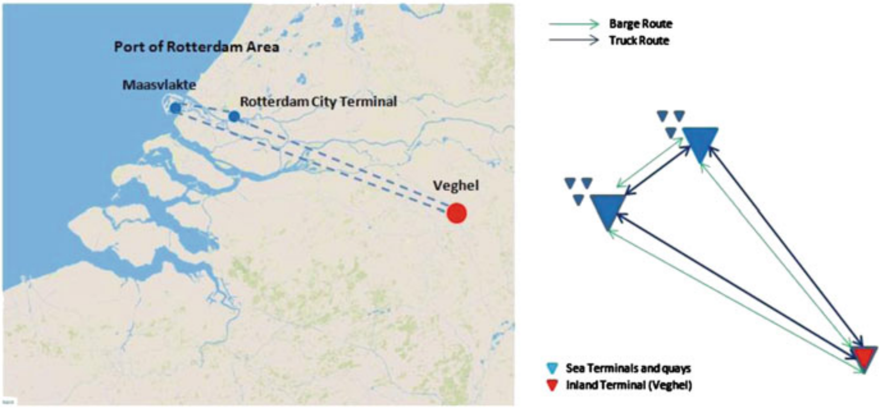


Fig. 17.6 Geographical location of the players and main flows

Both deep sea terminal areas have several quays to pick up the available containers. So the planner has to consider also a routing problem. In the work of Fazi (2014), the same problem was tackled under a pure operational point of view. Basically, the problem was modeled as a Vehicle Routing Problem with Pick-up and Deliveries (VRPPD) and heterogeneous fleet, see (Berbeglia et al. 2007) and (Toth and Vigo 2001) for a review. The particular time windows of the system, where the deadlines are defined at the destination of the single containers, have also been added to the formulation. The aim is to replicate exactly the decision making process that the planners face every day. The authors proposed some data set from a case study and solved them heuristically.

With regard to barge transport, barges have to visit the quays, meeting the capacity constraints. At each quay, the containers are first dropped and then picked up. Figure 17.7 shows a typical barge route. In the literature the transportation of containers between a depot and one or multiple destination has seen an exponential growth in the last two decades. The general terminology that includes these problems is ship routing problems. In his review paper, Ronen (1993) defined the ship routing problem as the assignment of shipments to ships and which sequence of ports to do. In his previous work, a heterogeneous fleet of vessels is scheduled on a set of predefined routes, from a single origin and multiple destination. Many papers have network design perspectives. Routes are generated a priori and then selected at later stage using Linear Programming formulations. Further, many problems are solved by set partitioning. When the total number of schedules are too large to enumerate, the most promising schedules are generated heuristically; see Christiansen et al. (2004) for a review of these papers. To the best of our knowledge, few papers address the problem of finding a route for a fleet of ships in a direct manner. Karlaftis et al. (2009) investigate route scheduling for a homogeneous fleet of containerships, performing pick-ups and deliveries between some ports and considering deadline constraints. They develop a VRPPD formulation. Every port has quantities to be picked up and

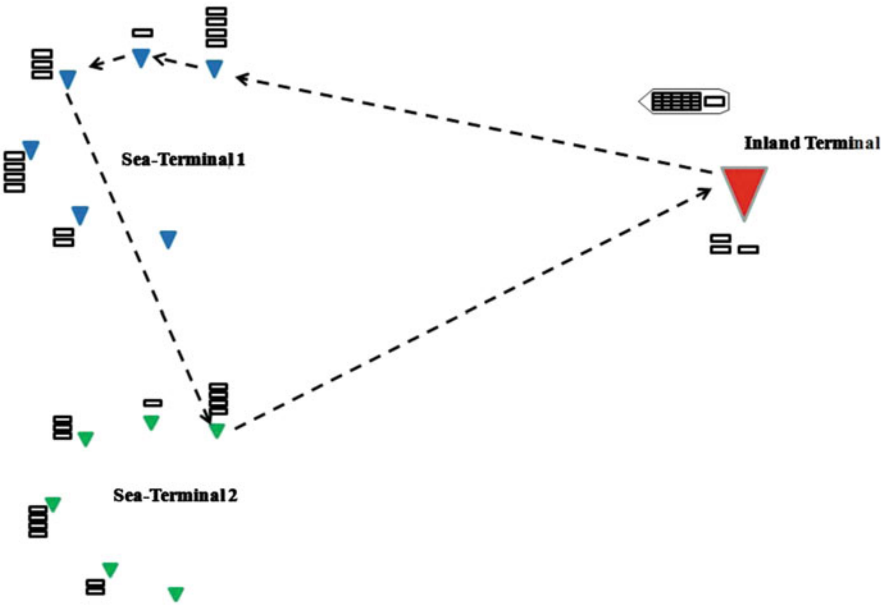


Fig. 17.7 A typical barge route. From the inland terminal through the sea terminals

delivered and it has to be processed before a certain deadline. They solve the problem by means of a genetic algorithm and they process real-world instances from the Aegean islands. Karlaftis et al. (2009) directly generate best containership routes and schedules, different from the existing literature.

With regard to truck transport, Imai et al. (2007) consider one depot and pick-up and delivery locations. In its round trip, a truck can go empty in one leg and full in the other (individual trips), or it can go full in both legs (merged trip), see Fig. 17.8. A VRPPD has been applied. They propose a Lagrangian relaxation-based heuristic to solve the problem. The problem is a special case of the models known in literature as VRP with backhauling or truck backhauling. In these models pickup customers can only be served after all delivery operations have taken place.

Fazi (2014) is the first of its kind, where different modalities, with such a strong trade-off, are considered together. In the literature, VRP with heterogeneous fleet has been tackled in many papers (Baldacci et al. 2008), but the considered test instances are basically for means of transport with same speed and with costs not reflecting an economy of scale.

Nowadays, many planners are not supported by computer tools for their decisions. Such tools, jointly with the new frontiers and restructuring of the sea-land system are essential to make the system smoother and limit the drawbacks of transportation in the hinterland. As underlined in the previous section, the role of information is also important. A better information system can be crucial in the quality of the schedules. The dry port with its evolution, the extended gate, entails that information is shared and real time. We suggest that this is the direction that sea-land systems have to follow to limit drawbacks in hinterland transportation.

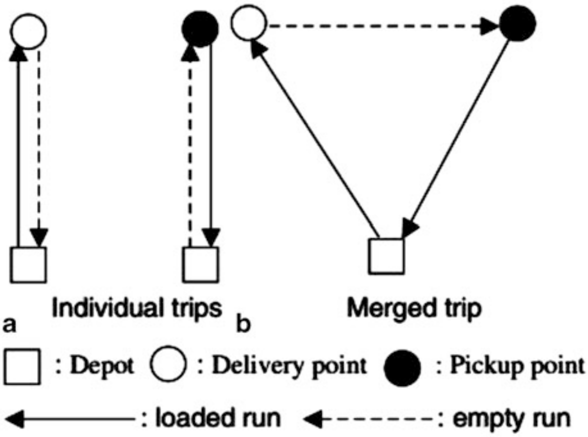


Fig. 17.8 Truck routes. (Source: Imai 2007)

17.5 Conclusion

The trend toward containerization and the huge increase in traded volumes worldwide put a lot of pressure on the hinterland transportation network. Volumes to be transportation are large and in many countries the road network cannot handle this sufficiently, leading to congestion problems. Congestion problems occur both in the port and port-city area, but also in the more extended hinterland, such as on main transportation axes between the Ports of Rotterdam and Antwerp and the German hinterland. On the other hand, the pressure on the road transportation system provides an extraordinary opportunity for the development of efficient hinterland transportation systems based on intermodal transportation and hub-and-spoke networks. In this chapter, we have analyzed the most important features of such container transportation systems for the hinterland supply chain.

At the network design level, we review the current state of the art and we identify avenues for future research. Among others, we highlight that the coordination of container shipments across the container supply chain is a particularly relevant issue as hinterland networks involve several actors. At the operational level, we characterize the most important factors influencing the trade-off between intermodal transportation and truck-only deliveries. In addition, we provide a case study. We highlight that the exchange of information is the key enabler for efficient hinterland intermodal transportation and we show that a better information system can be of crucial importance.

The management of the hinterland container supply chain has attracted substantial attention from the maritime economics community. The results obtained document very well the current trends and enable identifying and understanding the most important challenges faced by the hinterland supply chain. However, studies in the field of maritime economics either address strategic questions conceptually or focus on descriptive empirical research. We show in this chapter that model-based research addressing the issues raised by the management of the container supply chain is still scarce. Consequently, we incite the operations management community to explore the challenging context provided by the hinterland container supply chain. The expected results would be supplementary to the results obtained by the maritime economics community and they may help achieving substantial improvements.

Acknowledgements The research was partly financed by Dinalog, the Dutch Institute for Advanced Logistics.

References

Alumur, S., & Kara, B. Y. (2008). Network hub location problems: The state of the art. *European Journal of Operational Research, 190,* 1–21.

Alumur, S. A., Kara, B. Y., & Karasan, O. E. (2012a). Multimodal hub location and hub network design. *Omega, 40,* 927–939.

Alumur, S. A., Yaman, H., & Kara, B. Y. (2012b). Hierarchical multimodal hub location problem with time-definite deliveries. *Transportation Research Part E: Logistics and Transportation Review, 48,* 1107–1120.

Arnold, P., Peeters, D., & Thomas, I. (2004). Modelling a rail/road intermodal transportation system. *Transportation Research Part E: Logistics and Transportation Review, 40,* 255–270.

Baldacci, R., Battarra, M., & Vigo, D. (2008). Routing a heterogeneous feet of vehicles. In B. Golden, S. Raghavan, & E. Wasil (Eds.)*, The vehicle routing problem: Latest advances and new challenges* (Chap. 1, pp. 3–27). Springer, New York.

Berbeglia, G., Cordeau, J.-F., Gribkovskaia, I., & Laporte, G. (2007). Static pickup and delivery problems: A classification scheme and survey. *Top, 15,* 1–31.

Bierwirth, C., & Meisel, F. (2010). A survey of berth allocation and quay crane scheduling problems in container terminals. *European Journal of Operational Research, 202*(3), 615–627.

Blumenfeld, D. E., Burns, L. D., Diltz, J. D., & Daganzo, C. F. (1985). Analyzing trade-offs between transportation, inventory and production costs on freight networks. *Transportation Research Part B: Methodological, 19,* 361–380.

Bontekoning, Y., Macharis, C., & Trip, J. (2004). Is a new applied transportation research field emerging?—-A review of intermodal rail–truck freight transport literature. *Transportation Research Part A: Policy and Practice, 38,* 1–34.

Bouchery, Y., & Fransoo, J. C. (2014). Intermodal hinterland network design with multiple actors. BETA Working paper 449, Technische Universiteit Eindhoven.

Campbell, J. F. (1990). Freight consolidation and routing with transportation economies of scale. *Transportation Research Part B: Methodological, 24,* 345–361.

Campbell, J. F. (1993). Continuous and discrete demand hub location problems. *Transportation Research Part B: Methodological, 27,* 473–482.

Campbell, J. F., & O'Kelly, M. E. (2012). Twenty-five years of hub location research. *Transportation Science, 46,* 153–169.

Caris, A., Macharis, C., & Janssens, G. K. (2011). Network analysis of container barge transport in the port of Antwerp by means of simulation. *Journal of Transport Geography, 19,* 125–133.

Christiansen, M., Fagerholt, K., & Ronen, D. (2004). Ship routing and scheduling: Status and perspectives. *Transportation Science, 38,* 1–18.

Craig, A. J., Blanco, E. E., & Sheffi, Y. (2013). Estimating the CO2 intensity of intermodal freight transportation. *Transportation Research Part D: Transport and Environment, 22,* 49–53.

Crainic, T. G., & Kim, K. H. (2007). Intermodal transportation. In: C. Barnhart and G. Laporte, *Transportation. Volume 14 of handbooks in operations research and management science* (pp. 467–537), Amsterdam: North Holland.

Crainic, T. G., Gendreau, M., & Dejax, P. (1993). Dynamic and stochastic models for the allocation of empty containers. *Operations Research, 41,* 102–126.

Da Graça Costa, M., Captivo, M. E., & Clímaco, J. (2008). Capacitated single allocation hub location problem—A bi-criteria approach. *Computers & Operations Research, 35,* 3671–3695.

Daganzo, C. F. (1987). The break-bulk role of terminals in many-to-many logistic networks. *Operations Research, 35,* 543–555.

Farahani, R. Z., Hekmatfar, M., Arabani, A. B., & Nikbakhsh, E. (2013). Hub location problems: A review of models, classification, solution techniques, and applications. *Computers & Industrial Engineering, 64,* 1096–1109.

Fazi, S. (2014). Mode selection, routing and scheduling for inland container transport, PhD Thesis, *Eindhoven University of Technology*, Netherlands.

Fransoo, J. C., & Lee, C.-Y. (2013). The critical role of ocean container transport in global supply chain performance. *Production and Operations Management, 22,* 253–268.

Fransoo, J. C., De Langen, P. W., & Van Rooy, B. (2013). Business models and network design in hinterland transport. In J. H. Bookbinder (Ed.), *Handbook of global logistics, international series in operations research & Management Science*, (Vol. 181, Part 5, Chap. 15, pp. 367–389), Springer, New York.

Fremont, A., & Franc, P. (2010). Hinterland transportation in Europe: Combined transport versus road transport. *Journal of Transport Geography, 18,* 548–556.

Groothedde, B., Ruijgrok, C., & Tavasszy, L. (2005). Towards collaborative, intermodal hub networks: A case study in the fast moving consumer goods market. *Transportation Research Part E: Logistics and Transportation Review, 41,* 567–583.

Hakimi, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research, 12,* 450–459.

Hakimi, S. L. (1965). Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations Research, 13,* 462–475.

Hall, R. W. (1987a). Direct versus terminal freight routing on a network with concave costs. *Transportation Research Part B: Methodological, 21,* 287–298.

Hall, R. W. (1987b). Comparison of strategies for routing shipments through transportation terminals. *Transportation Research Part A: General, 21,* 421–429.

Imai, A., Nishimura, E., & Current, J. (2007). A Lagrangian relaxation-based heuristic for the vehicle routing with full container load. *European Journal of Operational Research, 176*(1), 87–105.

Ishfaq, R., & Sox, C. R. (2010). Intermodal logistics: The interplay of financial, operational and service issues. *Transportation Research* Part E: *Logistics and Transportation Review, 46,* 926–949.

Jarzemskis, A., & Vasiliauskas, A. V. (2010). Research on dry port concept as intermodal node. *Transport, 22,* 207–213.

Jeong, S.-J., Lee, C.-G., & Bookbinder, J. H. (2007). The European freight railway system as a hub-and-spoke network. *Transportation Research Part A: Policy and Practice, 41,* 523–536.

Jourquin, B., Beuthe, M., & Demilie, L. D. (1999). Freight bundling network models: Methodology and application. *Transportation Planning and Technology, 23,* 157–177.

Karlaftis, M. G., Kepaptsoglou, K., & Sambracos, E. (2009). Containership routing with time deadlines and simultaneous deliveries and pick-ups. *Transportation Research Part E: Logistics and Transportation Review, 45*(1), 210–221.

Khor, Y. S., Dohlie, K. A., Konovessis, D., & Xiao, Q. (2013). Optimum speed analysis for large containerships. *Journal of Ship Production and Design, 29*(3), 93–104.

Kim, K. H. (2005). Models and methods for operations in port container terminals. In A. Langevin & D. Riopel (Eds.), *Logistics systems: Design and optimization* (pp. 213–243). Berlin: Springer.

Konings, J. W. (1996). Integrated centres for the transshipment, storage, collection and distribution of goods. A survey of the possibilities for a high-quality intermodal transport concept. *Transport Policy, 3,* 3–11.

Konings, R., Kreutzberger, E., & Maras, V. (2013). Major considerations in developing a hub-and-spoke network to improve the cost performance of container barge transport in the hinterland: The case of the port of Rotterdam. *Journal of Transport Geography, 29,* 63–73.

Langevin, A., Mbaraga, P., & Campbell, J. F. (1996). Continuous approximation models in freight distribution: An overview. *Transportation Research Part B: Methodological, 30,* 163–188.

Liao, C.-H., Tseng, P.-H., & Lu, C.-S. (2009). Comparing carbon dioxide emissions of trucking and intermodal container transport in Taiwan. *Transportation Research Part D, 14,* 493–496.

Limbourg, S., & Jourquin, B. (2009). Optimal rail-road container terminal locations on the European network. *Transportation Research Part E: Logistics and Transportation Review, 45,* 551–563.

Lin, C.-C., & Lee, S.-C. (2010). The competition game on hub network design. *Transportation Research Part B: Methodological, 44,* 618–629.

Lüer-Villagra, A., & Marianov, V. (2013). A competitive hub location and pricing problem. *European Journal of Operational Research, 231,* 734–744.

Mangan, J., Lalwani, C., & Fynes, B. (2008). Port-centric logistics. *International Journal of Logistics Management, 19,* 29–41.

Meng, Q., & Wang, X. (2011). Intermodal hub-and-spoke network design: Incorporating multiple stakeholders and multi-type containers. *Transportation Research Part B: Methodological, 45,* 724–742.

Notteboom, T. (2004). Container shipping and ports: An overview. *Review of Network Economics, 3,* 86–106.

Notteboom, T. (2007). Inland waterway transport of containerized cargo: From infancy to a fully-fledged transport mode. *Journal of Maritime Research, 4,* 63–80.

Notteboom, T., & Rodrigue, J.-P. (2005). Port regionalization: Towards a new phase in port development. *Maritime Policy & Management, 32,* 297–313.

Notteboom, T., & Rodrigue, J.-P. (2009a). The future of containerization: Perspectives from maritime and inland freight distribution. *Geo Journal, 74,* 7–22.

O'Kelly, M. E. (1986a). The location of interacting hub facilities. *Transportation Science, 20,* 92–106.

O'Kelly, M. E. (1986b). Activity levels at hub facilities in interacting networks. *Geographical Analysis, 18,* 343–356.

O'Kelly, M., & Bryan, D. (1998). Hub location with flow economies of scale. *Transportation Research Part B: Methodological, 32,* 605–616.

Parola, F., & Sciomachen, A. (2005). Intermodal container flows in a port system network: Analysis of possible growths via simulation models. *International Journal of Production Economics, 97,* 75–88.

Racunica, I., & Wynter, L. (2005). Optimal location of intermodal freight hubs. *Transportation Research Part B: Methodological, 39,* 453–477.

Ronen, D. (1993). Ship scheduling: The last decade. *European Journal of Operational Research, 71,* 325–333.

Roso, V. (2007). Evaluation of the dry port concept from an environmental perspective: A note. *Transportation Research Part D, 17,* 523–527.

Roso, V., Woxenius, J., & Lumsden, K. (2009). The dry port concept: Connnecting container seaports with the hinterland. *Journal of Transport Geography, 17,* 338–345.

Rutten, B. C. M. (1998). The design of a terminal network for intermodal transport. *Transport Logistics, 1,* 279–298.

Skorin-Kapov, D. (1998). Hub network games. *Networks, 31,* 293–302.

Skorin-Kapov, D., & Skorin-Kapov, J. (2005). Threshold based discounting networks: The cost allocation provided by the nucleolus. *European Journal of Operational Research, 166,* 154–159.

Smilowitz, K. R., & Daganzo, C. F. (2007). Continuum approximation techniques for the design of integrated package distribution systems. *Networks, 50,* 183–196.

Stahlbock, R., & Voss, S. (2008). Operations research at container terminals: A literature update. *OR Spectrum, 30,* 1–52.

SteadieSeifi, M., Dellaert, N. P., Nuijten, W., Van Woensel, T., Raoufi, R. (2014). Multimodal freight transportation planning: A literature review. *European Journal of Operational Research, 233*(1), 1–15.

Toth, P., & Vigo, D. (2001). The Vehicle Routing Problem. Philadelphia: Siam.

Van den Berg, R., De Langen, P. W., & Rúa Costa, C. (2012). The role of port authorities in new intermodal service development; The case of Barcelona Port Authority. *Research in Transportation Business & Management, 5,* 78–84.

Van Schijndel, W.-J., & Dinwoodie, J. (2000). Congestion and multimodal transport: A survey of cargo transport operators in the Netherlands. *Transport Policy, 7,* 231–241.

Veenstra, A., Zuidwijk, R., & van Asperen, E. (2012). The extended gate concept for container terminals: Expanding the notion of dry ports. *Maritime Economics & Logistics, 14,* 14–32.

Vis, I. F. A., & De Koster, R. (2003). Transhipment of containers at a container terminal: An overview. *European Journal of Operational Research, 147,* 1–16.

Wardrop, J. G. (1952). Road paper: Some theoretical aspects of road traffic research. *ICE Proceedings: Engineering Divisions, 1,* 325–362.

Yaman, H. (2009). The hierarchical hub median problem with single assignment. *Transportation Research Part B: Methodological, 43,* 643–658.

Zuidwijk, R. A., & Veenstra, A. W. (2014). The value of information in container transport. *Transportation Science* (in press).

Chapter 18
Green Corridors and Their Possible Impact on the European Supply Chain

George Panagakos, Harilaos N. Psaraftis and Even Ambros Holte

Abstract The purpose of this chapter is to present the concept of green corridors and analyse their possible impact on the supply chain. The basis of this material is work conducted in the context of the EU SuperGreen project and therefore the geographical setting of the chapter is Europe. The general objective of the SuperGreen project has been to support the development of sustainable transport networks by fulfilling requirements covering environmental, technical, economic, social and spatial planning aspects. The chapter deals only with surface freight transport, including maritime transport, noting however that the quality of transport and logistics services is also affected by passenger transport competing for route capacity. Aviation is outside the scope of our analysis, as is the use of pipelines for liquid cargoes.

In addition, the chapter provides examples of the corridor development approaches employed in Europe, and describes the performance monitoring methodology developed by SuperGreen. The deep sea service linking China to Europe is compared to the trans-Siberian rail link between Beijing and Duisburg as an example. Finally, the new transport infrastructure policy of the European Union is reviewed to investigate the relationship between green corridors and the recently introduced concept of TEN-T core network corridors in order to derive implications for corridor governance.

Abbreviations

CEF	Connecting Europe Facility
CO_2	Carbon dioxide
CO_2-eq	Carbon dioxide equivalent unit
DEFRA	Department for Environment, Food and Rural Affairs (U.K.)
DSS	Deep Sea Shipping

H. N. Psaraftis (✉) · G. Panagakos
Department of Transport, Technical University of Denmark,
Bygningstorvet 1, 2800 Kgs. Lyngby, Denmark
e-mail: hnpsar@transport.dtu.dk

E. A. Holte
Marintek, Trondheim, Norway

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8_18

EDU	Equivalent Delivery Unit
ERTMS	European Rail Traffic Management System
EWTC	East-West Transport Corridor
GHG	Greenhouse gas
HC	High Cube (for containers)
ICT	Information and Communication Technologies
IMO	International Maritime Organisation
IQ-C	International Group for Improving the Quality of Rail Transport in the North-South-Corridor (Rotterdam-Genoa)
ITS	Intelligent Transport Systems
KPI	Key Performance Indicator
LCA	Life Cycle Assessment
MoU	Memorandum of Understanding
NOx	Nitrogen oxides (NO and NO_2)
R&D	Research and Development
RIS	River Information Services
RNE	RailNetEurope
SESAR	Single European Sky Air traffic management Research
SOx	Sulphur oxides (SO_2 and SO_3)
SSS	Short Sea Shipping
TEN-T	Trans-European Transport Network
TEU	Twenty-foot Equivalent Unit (for containers)
VAT	Value Added Tax
VTMIS	Vessel Traffic Management Information System

18.1 Introduction

The purpose of this chapter is to present the concept of green corridors and analyse their possible impact on the supply chain. The basis of this material is work conducted in the context of the EU SuperGreen project and therefore the geographical setting of the chapter is Europe. Much of the material of the chapter is a condensed version of the so-called “Green Corridors Handbook,” Vol. II (Panagakos et al. 2012).

The general objective of the SuperGreen project has been to support the development of sustainable transport networks by fulfilling requirements covering environmental, technical, economic, social and spatial planning aspects. More specifically the project aimed at:

- giving overall ***support and recommendations*** on green corridors to the EU’s Freight Transport Logistics Action Plan;
- ***encouraging co–modality*** for sustainable solutions;
- ***benchmarking*** green corridors based on selected Key Performance Indicators (KPIs) covering all aspects of transport operations and infrastructure (emissions, internal and external costs);

- conducting a programme of ***networking activities between stakeholders*** (public and private);
- ***delivering policy recommendations*** at a European level for the further development of green corridors; and
- providing ***recommendations concerning new calls for R&D*** proposals to support the development of green corridors.

It should be clarified that this chapter does not seek to present all the work performed under SuperGreen, not even a summary of it. The reader should be referred to the project's web site (www.supergreenproject.eu) and to other related publications. Instead, this chapter attempts to clarify the concept of 'green transport corridors' (or simply 'green corridors') as much as possible, encourage a standardised approach for developing and implementing a green corridor, and assist the customers of freight transport operators who may wish to understand the repercussions for their supply chain.

We also clarify that the chapter deals only with surface freight transport, noting however that the quality of transport and logistics services is also affected by passenger transport competing for route capacity. Aviation is outside the scope of our analysis, as is the use of pipelines for liquid cargoes.

The rest of this chapter is structured as follows. Sections 1 and 2 respectively define the concept of a transport corridor and of a green corridor. Sections 3 and 4 respectively examine the questions why do we need green corridors and how one can develop a green corridor. Section 5 discusses how one can monitor its performance. Section 6 investigates the relation between green corridors and the Trans-European Transport Network (TEN-T). Section 7 examines a maritime green corridor and Sect. 8 presents the chapter's conclusions.

18.2 What is a Transport Corridor?

Despite being used for years as a concept, there is no precise definition for a "transport corridor". The description that suits best the way the term is used in the present chapter is this of Arnold (2005):

> [A transport corridor] has both a physical and functional dimension. In terms of physical components, a corridor includes one or more routes that connect centres of economic activity. These routes will have different alignments but with common transfer points and will be connected to the same end points. These routes are composed of the links over which the transport services travel and the nodes that interconnect the transport services. The end points are gateways that allow traffic with sources or destinations outside the corridor (and its immediate hinterland) to enter or exit the corridor.
>
> An international transport corridor connects one or more neighbouring countries. It may also connect countries that are separated by one or more transit countries or provide a landlocked country with access to the sea. Some corridors have a single mode or a single route, but most have multiple routes and modes. Some are relatively short and defined by a principal gateway like a port. Others are defined by the region they serve. Still others are defined as part of a network serving a larger region.

Fig. 18.1 Rail Corridor A serving the "Blue Banana" region

> While it is important to separate the concepts of economic corridors and transport corridors, the fact is that most transport corridors are developed to support regional economic growth. They provide transport and other logistics services that promote trade among the cities and countries along the corridor.

Rail Corridor A, the corridor from Rotterdam to Genoa is a good example. It stretches from the sea ports of Rotterdam, Zeebrugge and Antwerp to the port of Genoa, right through the heart of the EU along the so-called "Blue Banana". This is the most heavily industrialised North-South route in Central Europe and connects Europe's prime economic regions (Fig. 18.1).

The "Blue Banana" includes economically strong urban centres such as Rotterdam, Amsterdam, Duisburg, Cologne, Frankfurt, Mannheim, Basle, Zurich, Milan and Genoa. All these centres are served and connected by the corridor, also indirectly including London and Brussels. The countries directly involved are The Netherlands, Belgium, Germany, Switzerland and Italy.

This outstanding position together with the resulting fact that this corridor carries by far the greatest transport volume in Europe, makes the Rotterdam-Genoa route with its branch to Zeebrugge and Antwerp the pioneer for international rail freight transport in Europe.

18.3 What is a 'green' Transport Corridor?

In a strict sense, a precise answer to this question is not available, and in fact one of the most important contributions of ongoing research on the topic would be to develop an explicit and workable definition of the 'green corridor' term.

The concept was introduced in 2007 by the *Freight Transport Logistics Action Plan* of the European Commission[1]. According to this document:

> . . . transport corridors are marked by a concentration of freight traffic between major hubs and by relatively long distances . . .
> . . . Industry will be encouraged along these corridors to rely on co-modality and on advanced technology in order to accommodate rising traffic volumes, while promoting environmental sustainability and energy efficiency . . .
> . . . Green transport corridors will . . . be equipped with adequate transhipment facilities at strategic locations . . . and with supply points initially for bio-fuels and, later, for other forms of green propulsion . . .
> . . . Green corridors could be used to experiment with environmentally-friendly, innovative transport units, and with advanced Intelligent Transport Systems (ITS) applications . . .
> . . . Fair and non-discriminatory access to corridors and transhipment facilities should be ensured in accordance with the rules of the Treaty.

Some years later, the Swedish Logistics Forum (Kyster-Hansen et al. 2011) worked out a more structured definition. According to them:

> Green Corridors aim at reducing environmental and climate impact while increasing safety and efficiency. Characteristics of a green corridor include:
>
> - sustainable logistics solutions with documented reductions of environmental and climate impact, high safety, high quality and strong efficiency,
> - integrated logistics concepts with optimal utilisation of all transport modes, so called co-modality,
> - harmonised regulations with openness for all actors,
> - a concentration of national and international freight traffic on relatively long transport routes,
> - efficient and strategically placed transhipment points, as well as an adapted, supportive infrastructure, and
> - a platform for development and demonstration of innovative logistics solutions, including information systems, collaborative models and technology.

If this is so, what makes a freight corridor green?

A careful examination of the aforementioned definitions lead to the conclusion that, with the exception of characteristics that relate to the efficiency of a corridor regardless of its colour, they can be decomposed into the following prerequisites that distinguish a green corridor from its non-green counterpart:

a. Reliance on co-modality, which in turn requires:
   - adequate transhipment facilities at strategic locations; and
   - integrated logistics concepts.
b. Reliance on advanced technology, allowing:
   - energy efficiency; and
   - use of alternative clean fuels.
c. Development and demonstration capabilities of environmentally-friendly and innovative transport solutions, including advanced telematic applications.
d. Collaborative business models.

---

[1] The *Freight Transport Logistics Action Plan* of the European Commission can be found at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2007:0607:FIN:EN:PDF.

It is important to note that significant characteristics, like the harmonised regulations prescribed by the Swedish definition, are excluded from the list of green prerequisites as they relate more to the efficiency rather than the environmental sustainability of a corridor. The collaborative business models also fall in a rather grey area, as they are usually needed in all types of transport corridors. However, they are much more important in formulating the integrated logistics concepts of the green corridors, and as such they have been included in the list.

The above discussion leads to the conclusion that a green corridor is efficient, but an efficient corridor is not necessarily green.

18.4 Why do We Need Green Corridors?

Those who follow the evolution of the EU transport policy cannot escape noticing that the corridor approach gains more and more importance as a response to the new and old challenges that the common transport policy faces in Europe.

- In March 2005, the European Commission and the railway sector agreed on a MoU referring to the implementation of ERTMS on six corridors to define a European migration strategy for the deployment of ERTMS.
- In October 2007, The European Commission published its "Freight Transport Logistics Action Plan", which introduced the concept of 'green corridors' as a means to improve the efficiency and sustainability of freight transport in Europe.
- In November 2010, the European Parliament and the Council adopted the EU Regulation No 913/2010 concerning a European rail network for competitive freight. This Regulation defines nine initial corridors along which, sufficient priority is given to freight trains crossing at least one border.
- In March 2011, the latest White Paper on transport that describes the European Commission's vision of future transport and the corresponding strategy for the next decade, introduced 'multimodal freight corridors' as a means to improve governance and to support pilot projects for innovative and clean transport services.
- In December 2013, the European Parliament and the Council adopted a Regulation on the new TEN-T guidelines, which introduced the concept of 'core network corridors' as an instrument to facilitate the coordinated implementation of the parts of the TEN-T with the highest strategic importance (core network).

At a lower level, the initiatives listed below comprise only a selection among a wide range of corridor applications in Europe:

- In December 2002, Germany, Austria and Italy adopted the Brenner Action Plan aiming at a significant and sustainable increase in intermodal volume along the Brenner corridor, one of the most trafficked international transit corridors, where—on a length of only 448 km between Munich and Verona—three countries and thus railway infrastructures and the Alps are being bridged (Mertel et al. 2007).

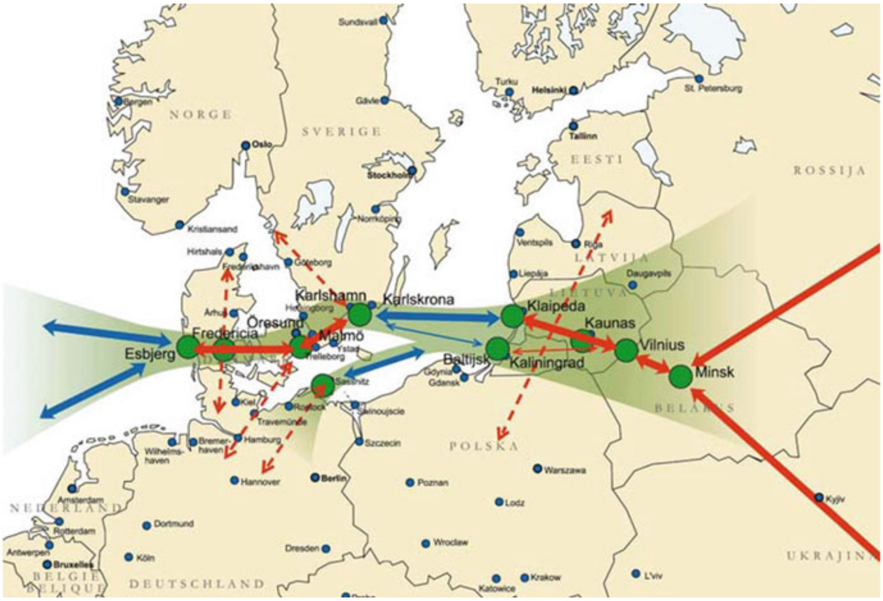


Fig. 18.2 The east-west transport corridor

- In January 2003, the Ministries of Transport of The Netherlands, Germany, Switzerland and Italy agreed on a MoU establishing an international working group to develop a comprehensive action plan aiming at bringing about numerous quantitative and qualitative improvements on the rail corridor from Rotterdam to Genoa (Corridor A/ IQ-C 2011). The so-called Corridor A was born (refer also to Sect. 1).
- In 2006, 42 partners (local, regional and national authorities, universities, harbours and private stakeholders) from Denmark, Lithuania, Russia and Sweden joined forces to strengthen transport development along the so-called “East-West Transport Corridor—EWTC” through infrastructure improvements, new solutions for business, logistics and cooperation between researchers. The success of EWTC led to the follow up project EWTC II, which aims at transforming the EWTC into a green corridor in line with the EU’s policy (Fig. 18.2).
- In 2008, the Swedish “Green Corridors” initiative was introduced focusing on transport routes and collaboration among shippers, forwarders, industry and haulers in order to optimise the use of available transport capacity. Today the project collaborates with the governments of Denmark, Finland and Norway.
- In 2009, the Scandria project was introduced, covering the corridor from the Region of Halland, via Zealand to Mecklenburg-Vorpommern and Berlin. The project cooperates with SoNorA, which extends coverage from Berlin to the Adriatic Sea.

- In 2009, the TransBaltic project was also introduced covering corridors across the Baltic Sea. Its overall objective was to provide incentives for the creation of a comprehensive multimodal transport system in the Baltic Sea Region.

There are a number of good reasons for making green corridors so popular:

- The consolidation of large volumes of freight for transport over long distances improves the competitiveness, and thus the possibilities of engagement, of modes like rail and waterborne transport, which are environmentally friendlier than trucks.
- The shift of cargoes away from European roads will alleviate the serious congestion problem that this transport mode faces, producing positive externalities to the other users of the road network through improvements in reliability and reduction of transport time.
- Additional environmental and financial (through lower operating costs) gains can also result from optimisation in terms of energy use and emissions, further enabled by the scale and length of such freight corridors.
- The international character of the corridors (involve at least three Member States) addresses the fragmented nature of transport networks, especially rail, dealing with the haunting interoperability issues in geographical terms. At the same time, focusing on a subset of the network improves the chances of identifying workable solutions by limiting the overwhelming scale of the problem.
- The realisation of international multimodal corridors cannot be implemented without appropriate corridor structures. These structures will bring together the Commission, Member States, the regions, the local authorities, but also the infrastructure owners and managers, transport operators, shippers, financiers and, when appropriate, neighbouring countries. The involvement of such structures is absolutely necessary in promoting multimodal logistics, where lack of coordination comprises probably the most persisting problem.
- The establishment of corridors that enhance the efficiency of transport modes (alone and in combination) through better utilisation of resources will limit the considerable investments needed for expanding the capacity of the transport networks in an environment of budgetary consolidation and increasing public opposition to major transport infrastructure projects especially in the vicinity of urban areas.

18.5 How do We Develop a Green Corridor?

Corridors are rarely developed as ‘greenfield’ projects. Most have been developed from existing routes, many of which date back to ancient trading routes, e.g. the Silk route. Nearly all evolved from existing land-based multimodal transport networks. Coastal and shortsea routes are less common but important for archipelagic countries. Inland water routes, too, are less common although important in riverine countries. Ocean routes are not usually included in the definition of the corridor because there is little need to develop the links on these routes. However, seaports are included since they serve as the international gateways.

The development of a corridor is closely related to the functions it serves. Having examined a number of international transport corridors, Arnold (2005) concludes that there are three general functions requiring management oversight:

- **Infrastructure and facilities**, including links and nodes along the routes, are developed and funded primarily by the public sector but increasingly constructed and maintained by the private sector. Management's role is to guide the planning and procurement of these assets. Its goal is to insure that these assets are:
  - designed to provide efficient movement of cargo along the infrastructure and through the facilities;
  - constructed and maintained so as meet required standards;
  - of sufficient capacity to meet projected demand;
  - used efficiently; and
  - fully utilised.
- **Transport and logistics** services. Increasingly these activities are undertaken by the private sector in a competitive market with costs recovered through user charges. The objective of the managers of individual services is to capture significant market share by offering a competitive combination of cost, time and reliability. To the extent that corridor management is responsible for overseeing these services, its objective should be to promote more efficient services, usually by encouraging competition but often by allowing vertical and horizontal integration.
- **Regulatory procedures** that affect the movement of goods in the corridor and the transport and logistics providers operating in the corridor. Rarely is corridor management involved in the enforcement of the regulations or even in the enactment of these regulations. Instead it performs an advocacy role discouraging excessive regulation and reforming regulation that leads to inefficiencies. The management can encourage reform by supporting efforts to harmonise procedures across borders, to simplify documentation and procedures, and to enhance transparency.

These corridor functions require different management approaches. The first one involves the public sector, the second the private sector, and the third both. One involves provision of assets in a market with limited competition and partial cost recovery, another provision of services in a competitive market with full cost recovery, while the third deals with enforcement of laws/regulations and tax collection. It is difficult to imagine a management structure that encompasses all three.

More recently, Engström (2011) reports that the Swedish Transport Administration views green corridors projects/initiatives as being divided into three main categories that interact and complement each other. These categories promote the view of logistics/transports as a system of integrated services and properties aiming at increased efficiency and a reduced negative ecologic impact. The three parts are:

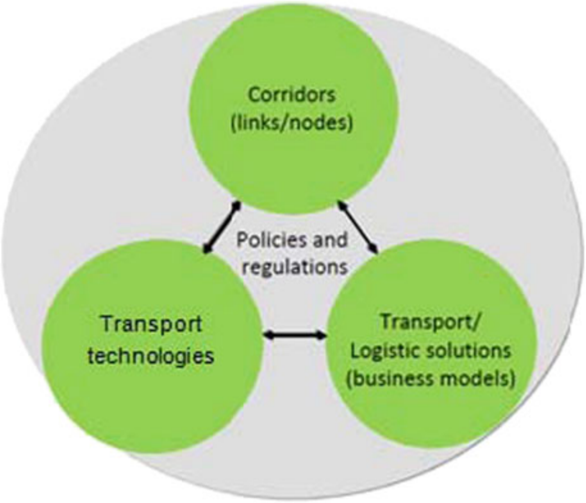


Fig. 18.3 The three pillars of green corridors. (Source: Engström 2011)

- **Corridors (links and nodes)**: A corridor project is a geographic subset of a designated main European Green Corridor. It is based on the needs of an efficient transport infrastructure in a physical and/or communicative aspect. A corridor project promotes optimal use of transport modes including transhipment nodes (hubs, cross docks etc). It can be of either a national or international character.
- **Transport technologies**: Projects related to transport technologies encompass features and properties of various types of equipment used in transport operation. The main focus is on the different transport modes, transport/load units and transfer/reloading of goods between different modes. Examples are technologies related to trucks, trailers, railway engines, rail wagons, ships, port handling, containers, packaging, cranes, stackers etc (Fig. 18.3).
- **Transport/logistics solutions**: Refers to complete solutions which integrate different partners and stakeholders mutually forming a business case that promotes efficiency and lowers environmental impact. In general terms, it is a complete freight logistic/ transport setup that meets a shipper's demand often linked to a new business model.

Although not seen as a 'pillar' in the Swedish schematic, the underlying policies and regulations are also recognised as a prerequisite for the implementation of green corridors.

Based on these functions, Arnold (2005) distinguishes between three general models that have been applied in corridor development:

Disjointed Incrementalism Viewed as part of a general development model, this approach is characterised by a project focus. Governments undertake improvements in the corridor infrastructure based on local requirements and problems. This model has been most effective in providing improvements in infrastructure. However, it lacks a formal corridor organisation or other mechanism to identify and prioritise initiatives.

Legislative Development This is characterised by the use of legislation to provide formal recognition of the importance of corridors, designation of specific routes, harmonisation of standards, simplification of cross-border movements and funding for corridor infrastructure. Implementation is left to individual jurisdictions and government agencies. Coordination is undertaken at the regional or ministerial level and is characterised by formal meetings to review progress made by others. Development of services on the corridor is left to private sector competition. Improvements in infrastructure are undertaken by government agencies responsible for transport. This approach is effective in targeting funding infrastructure and reducing formal impediments to movement of goods on these corridors.

Consensus-Building This approach uses a regional institution to mobilise stakeholder support for improvements in the corridor and to push for trade facilitation reforms including improving border-crossing procedures. Its primary function is to provide information to stakeholders, including government agencies, concerning current performance, needs for improvement, and success of previous initiatives. The success of this model depends on the active participation of public and private sector stakeholders in a partnership to address issues related to regulation, investment and quality of service.

Bringing this taxonomy into the current European environment, one could distinguish between two models:

Top-Down It corresponds to Arnold's legislative development model. It has been followed in all corridor development initiatives of the European Commission, such as the RNE corridors, the ERTMS corridors, the rail freight corridors of Regulation No 913/2010 and, more recently, the proposed TEN-T core network corridors. In a smaller scale, the Brenner corridor is a good example of a top-down model application.

Bottom-Up It corresponds to Arnold's consensus-building model. All Scandinavian projects such as the EWTC II, Scandria, TransBaltic, and Bothnian corridors comprise applications of this type of model.

No European equivalent to Arnold's disjointed incrementalism model is necessary, as activities such as priority setting and project identification under this model are more or less left uncoordinated, which is not the normal case of infrastructure development in Europe.

The comparison between these two models, bottom-up and top-down, is in essence meaningless. Their distinction basically relates to the origin of the initiative. In the top-down model the initiative comes from regional organisations, national governments or even local authorities. On the contrary, it is the transport and logistics companies themselves who take the initiative in the bottom-up model.

Nevertheless, as the corridor structures mature, their success will depend on whether they exhibit features like:

- the cooperation between public and private sectors; and
- the active participation of stakeholders.

In this respect, in the long run the two models will have to converge.

If the idea of a green corridor is more popular among private businesses, the bottom-up approach should be followed. The idea is cultivated among all types of stakeholders and once sufficient support is secured, the public sector is engaged. In any event, its involvement is necessary for signing the necessary bilateral or multilateral agreements.

If, on the other hand, the idea is originated in the ministerial offices or among infrastructure managers closely related to national governments, the top-down model seems to be more appropriate. Intensive information campaigns are needed to engage the private sector in the process as early as possible.

18.6 How do We Monitor Performance?

18.6.1 The Key Performance Indicators

It is important to understand that the indicators used for monitoring the performance of a green corridor are selected by the corridor management based on the objectives being pursued. Following a cumbersome methodology that heavily involved stakeholders, SuperGreen has concluded in the following KPIs (Psaraftis et al. 2012; Pålsson et al. 2010):

- **Out-of-pocket costs** (excluding VAT), measured in € /tonne-km;
- **Transport time**, measured in hours (or average speed, measured in km/h, depending on the application);
- **Reliability** of service (in terms of timely deliveries), measured in percentage of consignments delivered within a pre-defined acceptable time window;
- **Frequency** of service, measured in number of services per year;
- **CO_2 emissions**, measured in g/tonne-km; and
- **SOx emissions**, measured in g/tonne-km.

Others suggest different indicators. Arnold (2005) proposes the use of cost, time, reliability and flexibility (C/T/R/F). The management of Corridor A (Rotterdam-Genoa) has selected indicators concerning traffic volume, modal split, punctuality and commercial speed. The defined quality objectives of the BRAVO project (Brenner corridor) were punctuality, reliability, flexibility, customer information, employment rate of agreed rolling stock, and reliability of transport documents.

Once the indicators have been selected, the corridor performance can be monitored periodically as follows:

- **Step 1**: Disintegrate the corridor into transport chains.
- **Step 2**: Select a representative set of typical transport chains.

- **Step 3**: Estimate KPI values for each and every chain included in the representative set of typical transport chains determined in Step 2.
- **Step 4**: Aggregate these values into corridor level KPIs by using weights and pre-dermined methods.

18.6.2 Estimating KPI Values

As a general rule, the reported values should be:

Consistent The methodology employed should be consistent to allow for meaningful comparisons over time. Any changes to data, system boundaries, methods or any other relevant factor in the time series has to be clearly documented.

Transparent All relevant issues need to be addressed in a factual and coherent manner. The underline assumptions, calculation methodologies and data sources used have to be disclosed.

Accurate Ensure that uncertainties are reduced as far as practicable. Values reported should be of sufficient accuracy to enable users to make decisions with reasonable assurance as to the integrity of the reported information.

Some KPI-specific considerations are mentioned below:

Transport Cost and Time Transport costs include the out-of-pocket costs plus either the insurance costs or any loss or damage to cargo while en-route. Average costs are reported.

The costs incurred in a transport link are usually described as a combination of a fixed cost (in € /tonne) and a variable cost (in € /tonne-km) that depends on the distance travelled. Arnold (2005) uses the graph of Fig. 18.4 to calculate the total cost for moving a cargo over a distance x_3 comprising of three links.

The sloping lines in Fig. 18.4 represent the costs incurred while transiting a link with the slope proportional to the average variable cost, c_j. The vertical lines represent the costs incurred at the node and any fixed costs associated with using the subsequent link. A variety of activities can occur at these nodes, some required and others discretionary. One required activity is the transfer of cargo between transport units where there is a change of mode, physical constraints or regulatory requirements. Another is the inspection of the vehicle and its cargo occurring at the boundaries between jurisdictions. The most common discretionary activities occurring at the nodes are storage, intermediate processing, consolidation/deconsolidation, repackaging and labelling. It is important to exclude these activities when evaluating the performance of a transport chain. The components of these costs can be presented explicitly as shown in Fig. 18.5.

For the transport chain of Fig. 18.5, the average cost that needs to be reported is given by C_5/x_3, provided that cost figures along the Y-axis are specified in € /tonne.

Transport time is defined as the time to complete all the essential activities in moving from the beginning to the end of the transport chain including delays associated

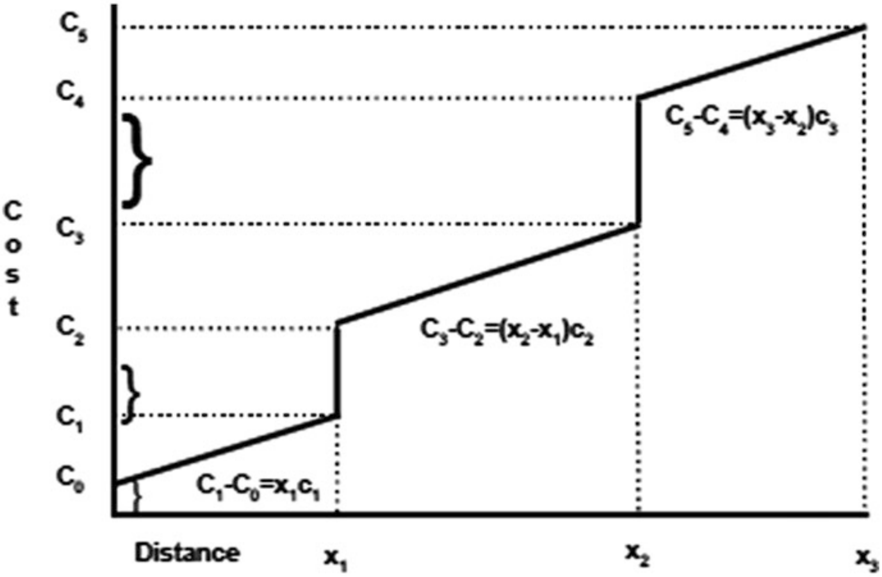


Fig. 18.4 Transit cost for a transport chain. (Source: Arnold 2005)

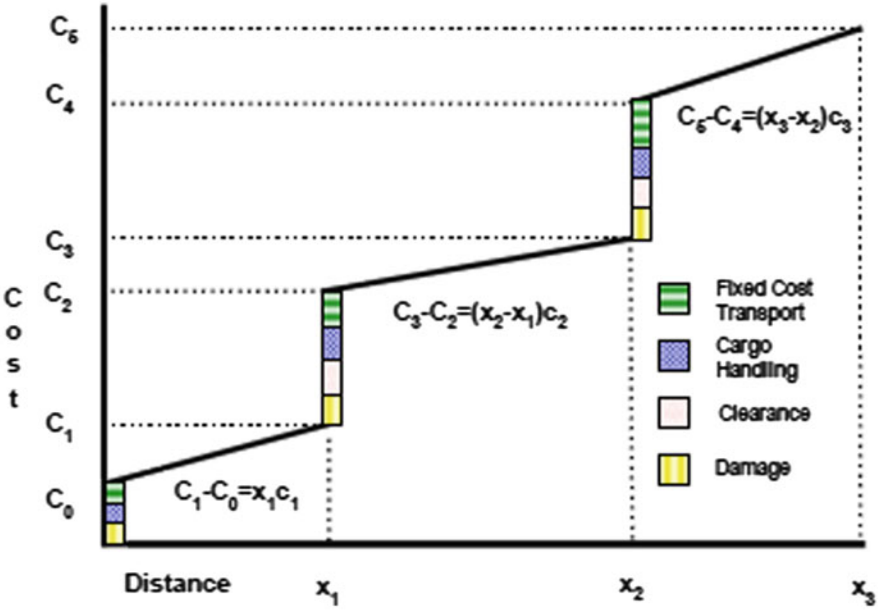


Fig. 18.5 Cost components of a transport chain. (Source: Arnold 2005)

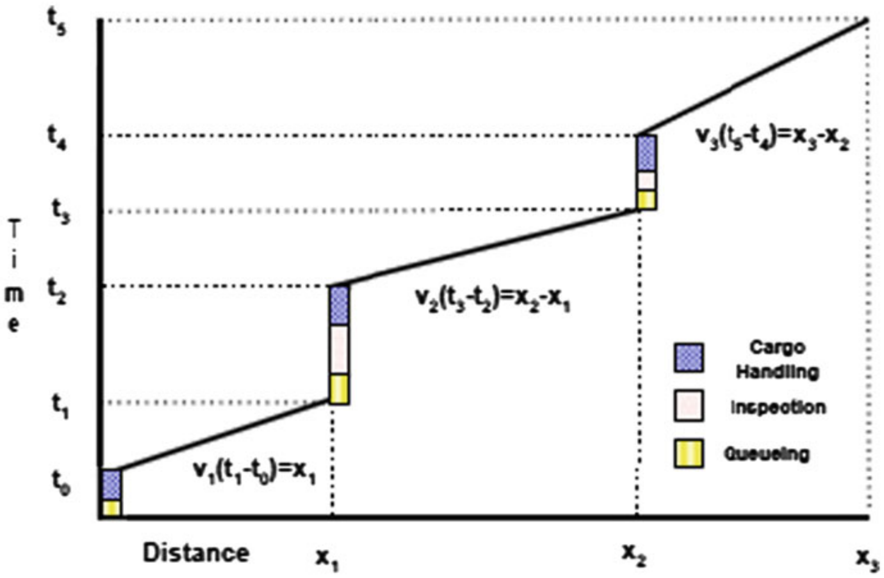


Fig. 18.6 Time components of a transport chain. (Source: Arnold 2005)

with the frequency of services and congestion at the nodes. Transport time can be presented as a function of distance along the chain using a graph of the type shown in Fig. 18.6. As with costs, the graph can display components like cargo handling, inspection and queueing.

Reliability and Frequency of Service Because of increasing attention to the timeliness of shipments and the importance of order fulfilment as a component of competitive advantage, it is necessary to consider not only the average time and cost for movement through a corridor but also the reliability in meeting delivery times. For the purposes of this discussion, reliability is measured in percentage of consignments delivered within an acceptable time window that has been defined by the corridor management a priori.

Delays are due to a combination of controllable factors, such as condition and availability of equipment, coordination of sequential activities, and labour productivity and uncontrollable market and environmental factors such as fluctuations in demand, level of background traffic and weather conditions. Although not required for calculating the KPI, knowing the reasons for the delays is vital for their mitigation.

As with reliability, the frequency of the various services in a transport chain results directly from surveys among the relevant service providers.

CO_2 and SOx Emissions When it comes to emissions, the definition of system boundaries is crucial in fulfilling all three criteria mentioned above (consistency, transparency and accuracy). Swahn (2010) defines four system boundaries (refer to Fig. 18.7):

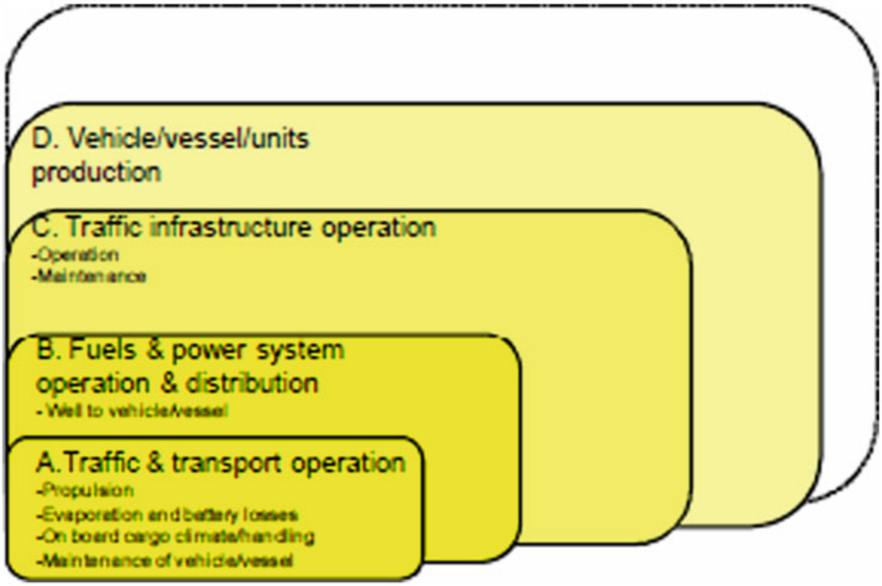


Fig. 18.7 Definition of system boundaries. (Source: Swahn 2010)

- **System boundary A** includes traffic and transport related activities regarding engine operation for the propulsion and equipment for climate control of goods, as well as losses in fuel tanks and batteries. This includes the traffic-related terminal handling, i.e. when goods do not leave their vehicle/vessel.
- **System boundary B** includes in addition the supply of energy from energy source to the tank, battery and electric motor (trains). This is the minimum required system boundary for performance of comparisons between different modes of transport.
- **System boundary C** includes in addition to the above traffic infrastructure operation and maintenance.
- **System boundary D** includes in addition to the above vehicle, vessel, load units production and scrapping (life cycle approach).

Although the introduction of the Life Cycle Assessment (LCA) methodology in decision making happens to be one of the policy recommendations that resulted from the SuperGreen project, it is essential to keep things as simple as possible in the early stages of a green corridor development. It is for this reason that the system boundary B is recommended to begin with. Later on, the boundary can be expanded to reach level D.

Another comment relates to the type of carbon emissions measured. In discussions of emissions, lots of terms are used—carbon emissions, carbon dioxide, greenhouse gases (GHG). In fact, climate change is caused by a range of gases, known collectively as 'greenhouse gases'. Of these, the most common is carbon dioxide (CO_2), which is why it's the most talked about. However, other greenhouse gases are emitted from vehicle exhausts (i.e. nitrogen dioxide and methane), and their reporting is also valuable. The choice between CO_2 and CO_2-eq (where the 'eq' stands for 'equivalent' simply meaning a unit for all GHGs expressed as if they had the same climate change effects as CO_2) depends on the availability of data and/or the capabilities of the emissions calculator used.

In general, a specialised emission calculator is needed for estimating the emission KPIs. In SuperGreen we have used the web-based tool EcoTransIT World[2] but, as long as certified footprint calculators are not available, any other model could be used in its position, provided that a relevant qualification escorts the results. User specified inputs are preferred to the model's default values, only when they are adequately verified and there is consistency across all chains examined. Otherwise, it is safer to use the default values of the model.

It is important to note that in a multi-load multi-drop vehicle trip the allocation of emissions to specific loads becomes quickly almost unworkably complex, requiring far more data than is likely to be available. A simplification is suggested by DEFRA (UK) (2009) according to which, emissions are allocated on the basis of the number of EDUs (Equivalent Delivery Units) transported for each customer. Generally speaking, the choice of EDU should reflect the limiting factor on the loading of the vehicle. If the load is typically limited by volume, then a volume-based EDU such as pallets or cube should be used. If the load is more often limited by weight, then a weight-based EDU such as tonnes will be more appropriate and provide more accurate results.

Finally, it is noted that graphs such as those for cost and time (Figs. 18.5 and 18.6) can be used for combining emissions generated while transiting a link with those produced at a node.

18.7 How do Green Corridors Relate With the TEN-T?

In December 2013, the European Parliament and the Council adopted a legislative package defining a new policy framework for the Trans-European Transport Network (TEN-T), which was proposed by the European Commission back in October 2011. The package includes a Regulation on the new Union guidelines for the TEN-T development with a time horizon extending to 2050 (EP&C 2013a) and a Regulation for establishing the Connecting Europe Facility (CEF), which will govern EU funding until 2020 (EP&C 2013b).

The TEN-T Guidelines, as the first component of the package, establish the policy basis by defining network plans including infrastructure standards, objectives and priorities for action. A dual layer network structure has been introduced, consisting of a comprehensive and a core network. The comprehensive network constitutes the

---

[2] For more information on EcoTransIT World, visit: http://www.ecotransit.org/.

basic layer of the TEN-T and is, in large part, derived from the corresponding national networks. The core network, on the other hand, overlays the comprehensive network and contains its strategically most important parts. The core network is the result of a genuine European network planning methodology that combines geographical and economic criteria. It builds on the key nodes of political, economic, cultural and transport-related importance and links them through all available transport modes.

The functions of the comprehensive and the core network complement each other: whereas the purpose of the comprehensive network is to serve accessibility functions and ensure a balanced infrastructure endowment throughout the Union, the core network pioneers the development of a sustainable mobility network. It shall be completed as a priority, by 2030. The new policy basis provides more clarity with regard to the identification of a broad range of "projects of common interest" (including the closing of missing physical links, infrastructure upgrading to target standards, ITS or innovative equipment).

To facilitate implementation of the core network, the Guidelines introduce the instrument of "core network corridors"—a coordination tool aiming at coherent project implementation and at promoting technological, operational and governance-related innovation. The core network corridors also aim to strengthen a "systems" approach that links transport infrastructure development with related transport policy measures. Eventually, this approach seeks to promote higher resource efficiency to achieve the EU carbon emissions' reduction objectives in the transport sector. Due to the broad range of measures addressed with the new Guidelines, many different actors will have to contribute to their implementation. The proposed corridor governance structures intend to foster cooperation of the various actors. Existing activities such as the rail freight corridors introduced with Regulation No 913/2010 will form an integral part of core network corridor developments.

Vis-à-vis the TEN-T guidelines, the Connecting Europe Facility (CEF) as the financing instrument sets out funding priorities in transport, energy and digital broadband for the period 2014–2020, as well as the corresponding rules. Regarding transport, it defines a geographical basis for the corridor approach and pre-identifies the most mature projects along those corridors.

Annex I to the CEF Regulation lists the 9 core network corridors of Fig. 18.8, which form the basic part of the TEN-T core network.

How do these corridors relate to the SuperGreen ones? Figure 18.9 depicts the land part of the core network plotted against the nine SuperGreen corridors. The geographic overlap is impressive, even after accounting for the fact that the priority projects of the TEN-T were taken into consideration, among several other criteria, when selecting the SuperGreen corridors in June 2010[3].

---

[3] To be exact, the TEN-T core network corridors of Fig. 18.9 are those proposed by the European Commission in October 2011. A few differences exist between them and the final set of Fig. 18.8 (as adopted by the European Parliament and the Council in December 2013) but they are not sufficient to alter the general picture of remarkable overlap with the SuperGreen corridor set.

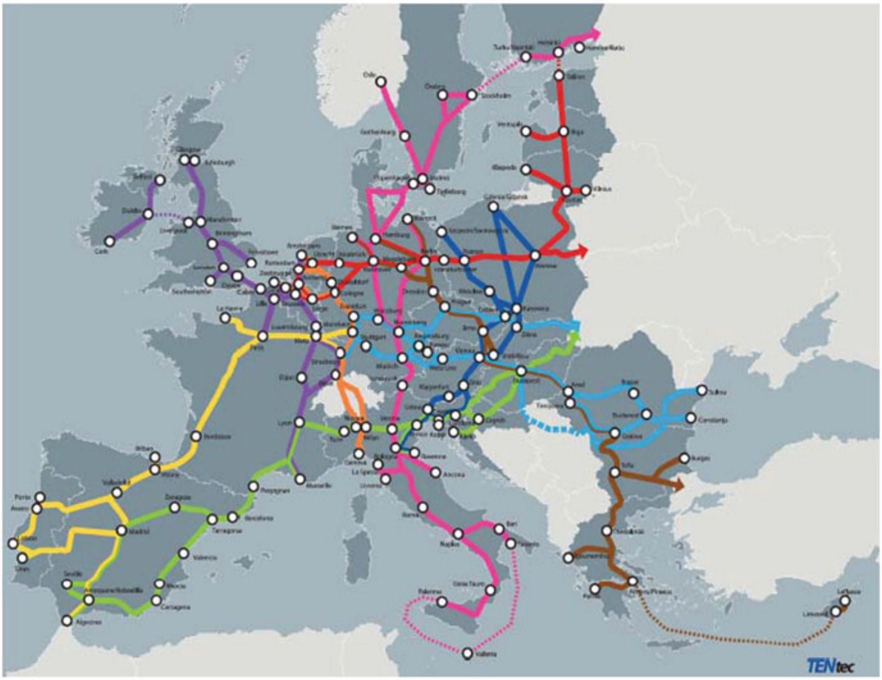


Fig. 18.8 Indicative map of the core network corridors of the TEN-T Regulation

With regard to the relation between these two sets of corridors, a key question to address is whether the TEN-T corridors exhibit the green characteristics identified in Sect. 2:

Reliance on Co-Modality Although the term co-modality is not mentioned, the Guidelines include several references to multimodality. In fact, there is an entire section (Sect. 6) devoted to the 'infrastructure for multimodal transport' that refers to the comprehensive network and includes logistic platforms. When it comes to the core network, Article 42 is crystal clear:

> … In order to lead to resource-efficient multimodal transport, … core network corridors shall be focused on modal integration, interoperability, and a coordinated development of infrastructure.

Adequate transhipment facilities The TEN-T Guidelines provide for:

- the connection of rail freight terminals with the road infrastructure or, where possible, the inland waterway infrastructure of the comprehensive network (Article 12);
- the connection of inland ports with the road or rail infrastructure (Article 15);

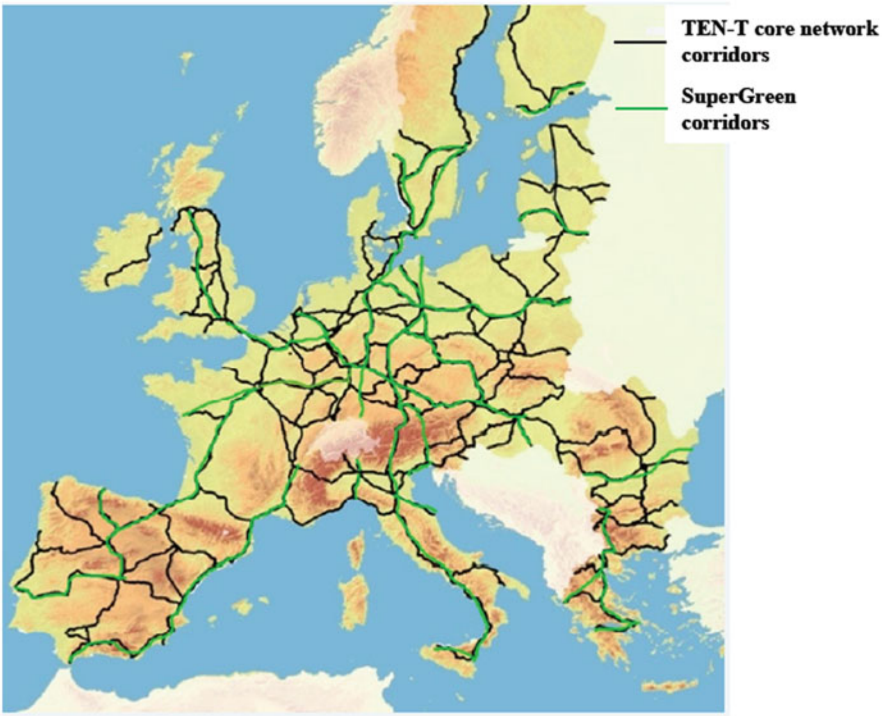


Fig. 18.9 The SuperGreen and TEN-T core network corridors

- the connection of maritime ports with railway lines or roads and, where possible, inland waterways of the comprehensive network, except where physical constraints prevent such connection (Article 22);
- multimodal interconnections between airports and infrastructure of other transport modes (Article 26);
- seamless connection between the infrastructure of the comprehensive network and the infrastructure for regional and local traffic and urban freight delivery, including logistic consolidation and distribution centres (Article 30).

Integrated Logistics Concepts It is worth mentioning that the general objective of the TEN-T is to "... strengthen the social, economic and territorial cohesion of the Union and contribute to the creation of a single European transport area which is efficient and sustainable, increases the benefits for its users and supports inclusive growth" (Article 4).

Furthermore, one of the criteria for identifying 'projects of common interest,' which compise the building blocks of the TEN-T, is the demonstration of 'European added value' (Article 7) which, in turn, is defined as '... the value of a project which, in addition to the potential value for the respective Member State alone, leads to a significant improvement of either transport connections or transport flows between the Member States which can be demonstrated by reference to improvements in efficiency, sustainability, competitiveness or cohesion . . . ' (Article 3).

Reliance on Advanced Technology There are numerous references to advanced technology applications including ICT. The following is an indicative list:

- "[TEN-T contributes to efficiency through] . . . cost-efficient application of innovative technological and operational concepts" (Article 4);
- "The TEN-T shall be planned, developed and operated in a resource-efficient way, through . . . the deployment of new technologies and telematic applications, where such deployment is economically justified" (Article 5);
- "In the development of the comprehensive network, general priority shall be given to measures that are necessary for . . . implementing and deploying telematic applications and promoting innovative technological development" (Article 10);
- "Telematic applications shall, for the respective transport modes, include in particular ERTMS (for railways), RIS (for inland waterways), ITS (for road transport), VTMIS and e-Maritime services (for maritime transport) and the SESAR system (for air transport)" (Article 31);
- "In order for the comprehensive network to keep up with innovative technological developments and deployments, the aim shall be in particular to support and promote the decarbonisation of transport through transition to innovative and sustainable transport technologies" (Article 33):
- "The core network corridors shall support the comprehensive deployment of interoperable traffic management systems and, where appropriate, the use of innovation and new technologies" (Article 42).

Energy Efficiency Relevant references include:

- "In the development of the comprehensive network, . . . particular consideration shall be given to measures that are necessary for . . . ensuring fuel security through increased energy efficiency, and promoting the use of alternative and, in particular, low or zero carbon energy sources and propulsion systems" (Article10);
- "Member States shall pay particular attention to projects of common interest which both provide efficient freight transport services that use the infrastructure of the comprehensive network and contribute to reducing carbon dioxide emissions and other negative environmental impacts, and which aim to stimulate resource and carbon efficiency, in particular in the fields of vehicle traction, driving/steaming, systems and operations planning" (Article 32).

Use of Alternative Clean Fuels The TEN-T Guidelines provide direct references to alternative fuels for all transport modes:

- "Member States shall ensure that the railway infrastructure, save in the case of isolated networks, is fully electrified as regards line tracks and, to the extent necessary for electric train operations, as regards sidings" (Article 12);

- "Projects of common interest for motorways of the sea … may also include activities … for improving environmental performance, such as the provision of shore-side electricity … and alternative fuelling facilities …" (Article 21);
- "In order for the comprehensive network to keep up with innovative technological developments and deployments, the aim shall be in particular to make possible the decarbonisation of all transport modes by stimulating energy efficiency, introduce alternative propulsion systems, including electricity supply systems, and provide corresponding infrastructure" (Article 33).

As for the core network, Article 39 stipulates full electrification of the line tracks and selective sidings for the railways, while alternative clean fuels should be available for the road, inland waterway and maritime transport infrastructures. For air transport, the relevant requirements are reduced to the "… capacity to make alternative clean fuels available."

Development of Innovative Logistics Solutions The promotion of innovative solutions is mentioned several times in the guidelines:

- "In the development of the comprehensive network, general priority shall be given to measures that are necessary for… promoting the efficient and sustainable use of the infrastructure …" (Article 10);
- "When developing the comprehensive network in urban nodes, Member States shall, where feasible, aim to ensure promotion of efficient low-noise and low-carbon urban freight delivery" (Article 30);
- "Member States shall pay particular attention to projects of common interest which … aim to promote the deployment of innovative transport services …" (Article 32);
- "Projects of common interest relate to all directly concerned stakeholders, … [who may contribute to] … the promotion of sustainable transport solutions, such as enhanced accessibility by public transport, telematic applications, intermodal terminals/multimodal transport chains, low-carbon and other innovative transport solutions and environmental improvements" (Article 50).

Collaborative Business Models Although no direct reference to business models can be found in the guidelines, there are several ones relating to the need for enhanced cooperation among stakeholders including provision of information:

- "The… core network corridors, is a strong means of realising the respective potential of stakeholders, of promoting cooperation between them and of strengthening complementarity with actions by Member States" [Preample (50)];
- "Member States shall ensure … that freight terminals and logistic platforms, inland and maritime ports and airports handling cargo are equipped for the provision of information flows within this infrastructure and between the transport modes along the logistic chain" (Article 28);
- "Telematic applications shall be such as to enable traffic management and the exchange of information within and between transport modes for multimodal transport operations and value-added transport-related services, improvements

in safety, security and environmental performance, and simplified administrative procedures" (Article 31);

- "Member States shall pay particular attention to projects of common interest which ... aim to promote the deployment of innovative transport services ... through ... the establishment of relevant governance structures" (Article 32);
- "Member States shall pay particular attention to projects of common interest which ... aim to facilitate multimodal transport service operations, including the necessary accompanying information flows, and improve cooperation between transport service providers" (Article 32).

The above references lead to the conclusion that all green characteristics of a corridor that have been identified in Sect. 2 are met more or less by the TEN-T core network corridors, as they have been introduced in the new Guidelines. A direct implication of this conclusion is that the governance structure proposed for the TEN-T core network corridors can be applied for all other green corridors after accounting for the fact that the latter do not involve passenger traffic which is included in the TEN-T corridors.

18.8 A Maritime Green Corridor

Several of the SuperGreen corridors involved a maritime component, in whole or in part. One of them is the so-called 'Silk Way' corridor. This corridor consists of two main transport services linking the Far East with Europe. Today there are mainly two alternatives for shipping large transhipments of goods between the two regions, one being the deep sea service linking Shanghai to the Le-Havre-Hamburg region, while the other is the rail-link between Beijing and Duisburg/EU. The main goods transported in the corridor are consumer goods. Most of the description in this section is based on work in SuperGreen deliverable D2.4 (Ilves et al. 2011).

The Silk Way deep sea route has its origin in the port of Shanghai with the Le-Havre—Hamburg range as point of destination, via the Suez Canal.

For this analysis the results from the IMO Second GHG Study (Buhaug et al. 2009) were used. Since the distance between the Far East and Europe is approximately 20,000 km, the total voyage is covered in the range of 35–41 days (assuming an average speed of 20–24 knots).

In order to calculate the transport cost for shipping one TEU from the Far East to Europe, the end of fourth quarter 2009 freight rate from UNCTAD's Review of Maritime Transport (Asariotis et al. 2010), has been applied. Although freight rates may fluctuate substantially within and over the years, as well as between different container lines, the selection represents relatively updated figures and a market average of the three largest container lines covering the Asia-Europe trade. Based on the above the calculated cost per t/km follows in the table below (Table 18.1).

An alternative route of the Silk way corridor is by rail. According to the corridor description, the rail way link goes from Beijing to Hamburg but following the service provided by DB Schenker's TransEurAsia Express, the analysis for the rail link will utilise Duisburg as point of destination. Although this means that the two transport

Table 18.1 Calculation of transport cost in € /tkm for Silk Way deep sea service

Given cost in USD per container (TEU)	1422 (€ 1002[a])
Distance covered in km	20,000
Average net tonnes transported per TEU	12
Cost per ton (1002.32/12 net tonnes per TEU)	84
Cost in € /tkm	*0.004*

[a] Calculated by DnB Nor Markets currency exchange calculator

Fig. 18.10 Rail service linking Far East to Europe via Russia. (http://www.trans-eurasia-logistics.com/Products/China-Europe/index.php [Accessed 8 April 2011])

services with similar points of origin now will have different point destinations, the distance between Duisburg and Hamburg is 'only' 376 km. Thus, it is assumed that the results of the analysis will not be significantly affected.

For cargo transport the rail link between Shanghai/Beijing and Duisburg takes approximately 18 days from terminal to terminal along the route depicted below. Although such a train service is not capable of transporting in one shipment the amount of goods carried by a large container vessel, the transport time is considerably shorter than the 35–40 days of the deep sea transport (Fig. 18.10).

The service is based on a regularly scheduled transport with a fixed route and departure days. Due to differences in rail gauges between Russia and China, a block train is formed in Zabaykalsk at the Russian/Chinese border with containers coming from Shanghai/Beijing. From Zabaykalsk the train travels en-route to the EU border at Brest/Malaszewicze. From here, connections are available to Duisburg (including all gateway connections), Hamburg, Warsaw, Prague and other destinations in Europe. According to the EcoTransIT online calculation tool the total distance from Shanghai to Duisburg is approximately 11,000 km.

Similar to deep sea transport service there are a number of different publications regarding the energy efficiency of rail transport, and the tendency of fluctuating results is also evident for rail transport. As shown in the table below there are significant differences between the studies both in terms of the gCO_2/tkm ranges and the respective range average values.

Table 18.2 Overview of CO_2 emission for rail transport

gCO_2/tkm range	gCO_2/tkm (range average)	Reference	Geographical scope
10–119	41	IMO Second GHG study (2009)	USA, Europe, UK
–	18	Psaraftis and Kontovas (2010)	Trans-Siberian railway
14–148	81	Lindstad and Mørkve (2009)	USA, Europe
13–33[a]	20[b]	Geitz and Jia (2010)	Europe incl. Black Sea Region
–	17[c]	Maersk Line (2007)	Not given

[a] Calculated by the SuperGreen Project (converting CO_2 per train km to gCO_2/tkm)
[b] Estimated by the SuperGreen Project, Work package 2
[c] Diesel consumption by train

According to the results of the Geitz and Jia (2010) study, the carbon footprint of rail (expressed in gCO_2/tkm) is considerably lower than those provided by the IMO Second GHG study (2009) and Lindstad and Mørkve (2009) (refer to Table 18.2).

Furthermore, there are also considerable differences in the presented gCO_2/tkm *range average*. The reason for this is assumed to be much related to the energy mix applied in the different studies in addition to variations in geographical scope. Due to the difficulty of actually tracing the electricity production method (coal, nuclear power, natural gas, hydro, etc.), there are uncertainties in comparisons to other transport modes. It is the results of the IMO Second GHG study (2009) that have been used in the summary KPI table.

The cost per t/km for one container is calculated as follows (Table 18.3):

KPI Evaluation Results The KPI summary table for the Silk Way corridor is presented below. It should be kept in mind that, since the most important aspect of this analysis is to shed light on the energy efficiency and ability to perform transport work between the two regions, the focus has been on the rail link going from China, via Russia into Europe, and the deep sea service linking China with Europe (Shanghai-Le-Havre-Hamburg range) (Table 18.4).

The above analysis leads to the following concluding remarks regarding transport mode performance:

- Rail has significantly lower cargo carrying capacity compared to the average deep sea vessel deployed in such a trade (Far East—Europe).
- Rail has considerable lower transport time and is as such a competitive advantage compared to the deep sea service.
- Due to the deep sea scale effect of being able to transport a significant larger amount of TEUs in one shipment, deep sea transport achieves a better performance in terms of gCO_2/tkm and much lower cost per tkm.

Table 18.3 Calculation of transport cost in € /tkm for the Silk Way railway service. (Source: DB Schenker 2011, SuperGreen calculations, 2010)

Quotation specifics	Cost elements
Freight main haul/train (40′ HC)	$ 8,230
Cross-docking Rail Terminal China (loaded)	$ 122
Insurance main haul/train China	$ 25
Security costs Russian Federation	$ 100
Re-expedition costs	$ 35
Other administration	$ 210
Liability insurance	$ 35
Given Total cost per container (TEU)[a]	*$ 8,757 (€ 6,159)*[b]
Distance covered in km	11,000
Average net tonnes transported per TEU	12
Total transported net tonnes	1200
Cost per ton in € (6 158,60/12 net tonnes per TEU)	513
Cost in € /tkm	*a 0.05*

[a] Price quotation for the Transeurasia Express, Not included: provision of empty containers, risk surcharges, currency surcharges. Provided by SuperGreen Project Partner, DB Schenker

[b] Calculated by DnB Nor Markets Currency exchange calculator

Table 18.4 Benchmarks for the Silk Way corridor

	Rail	Road	DSS	SSS
CO_2 (g/tkm)	41[a]	–	12.5[b]	–
NOx (g/tkm)	–	–	–	–
Cost (€ /tkm)	0.05	–	0.004	–
Average speed (km/h) (Calculation is based on the distance/transit time)	26	–	20–23	–
Reliability (%)	–	–	–	–
Frequency (no per year)	–	–	–	–

[a] Block Train

[b] TEU > 8000

18.9 Conclusions

A green corridor is a concept introduced by the European Commission in 2007 to reduce the environmental and climate impact of freight logistics while increasing safety and efficiency. It is marked by a concentration of freight traffic between major hubs and by relatively long distances.

The analysis performed under the SuperGreen project resulted in the following characteristics that make an otherwise efficient corridor green:

e. Reliance on co-modality, which in turn requires:
   - adequate transhipment facilities at strategic locations; and
   - integrated logistics concepts.

f. Reliance on advanced technology, allowing:
   - energy efficiency; and
   - use of alternative clean fuels.

g. Development and demonstration capabilities of environmentally-friendly and innovative transport solutions, including advanced telematic applications.

h. Collaborative business models.

The basic impact of green corridors on the European supply chains stems from the fact that the consolidation of large volumes of freight for transport over long distances improves the competitiveness, and thus the possibilities of engagement, of modes like rail and waterborne transport, which are environmentally friendlier than trucks. The resulting shift of cargoes away from European roads will alleviate the serious congestion problem that this transport mode faces, producing positive externalities to the other users of the road network through improvements in reliability and reduction of transport time. The scale and length of such freight corridors enables further optimisation in terms of energy use that will result in additional environmental and financial (through lower operating costs) gains.

In addition, the international character of the corridors, which involve at least three Member States, addresses the fragmented nature of transport networks, especially rail, dealing with: (i) the haunting interoperability issues in geographical terms and (ii) the lack of coordination among various stakeholders that comprises probably the most persisting problem of modern logistics. At the same time, focusing on a subset of the network improves the chances of identifying workable solutions by limiting the overwhelming scale of the problem.

Yet another expected benefit of corridors that enhance the efficiency of transport modes (alone and in combination) through better utilisation of resources relates to limiting the considerable investments needed for expanding the capacity of the transport networks which, in an environment of budgetary consolidation and increasing public opposition to major transport infrastructure projects especially in the vicinity of urban areas, can be of paramount importance.

The chapter also provided examples of the top-down and bottom-up corridor development approaches employed in Europe, emphasising on the significant role of the public/private sector cooperation and the active participation of stakeholders in their success.

The corridor performance monitoring methodology developed by SuperGreen was also presented. It is based on a small number of KPIs reflecting the objectives pursued by the corridor management. The following is a list of suggested indicators:

- **Out-of-pocket costs** (excluding VAT), measured in € /tonne-km;
- **Transport time**, measured in hours (or average speed, measured in km/h, depending on the application);
- **Reliability** of service (in terms of timely deliveries), measured in percentage of consignments delivered within a pre-defined acceptable time window;
- **Frequency** of service, measured in number of services per year;
- **CO_2 emissions**, measured in g/tonne-km; and
- **SOx emissions**, measured in g/tonne-km.

The proposed methodology consists of the following steps:

- Disintegrate the corridor into transport chains
- Select a representative set of typical transport chains
- Estimate KPI values for each and every chain of the selected set
- Aggregate these values into corridor level KPIs by using proper weights and methods.

As an example, the deep sea service linking China to Europe (Shanghai—Le-Havre-Hamburg range) was compared to the trans-Siberian rail link between Beijing and Duisburg/EU. Rail transport is faster but the scale effect of deep sea shipping leads to significant advantages in terms of costs and CO_2 emissions both expressed on a per tonne-km basis.

Finally, the new transport infrastructure policy of the European Union was reviewed to investigate the relationship between green corridors and the recently introduced concept of TEN-T core network corridors. It was found that the TEN-T Guidelines include the elements necessary to promote sustainable transport in the broad sense. Its declared objective is to provide the infrastructure basis for achieving the overall European transport policy objective of meeting mobility needs while reducing GHG emissions.

Core network corridors—where the EU's coordination and funding action will be concentrated—are foreseen to pioneer such a development. The existing green corridors, initiated by some Member States can be seen as a nucleus (to be integrated into the broader context of the TEN-T Guidelines), and the benchmarking methodology developed within the SuperGreen project will be a very useful tool to optimise planning and implementation.

Acknowledgment Work in this paper was supported in part from EU project SuperGreen (grant agreement TREN/FP7TR/233573/"SUPERGREEN"),

References

Arnold, J. (2005). *Best practices in corridor management*, Trade Logistics Group, The World Bank, Washington, D.C., February 2005.

Asariotis, R., Benamara, H., Hoffmann, J., Misovicova, M., Núez, E., Premti, A., Sanderson, D., Sitorus, B., Valentine, V., & Viohl, B. (2010). *Review of maritime transport 2010*. Report by the UNCTAD Secretariat, United Nations, New York and Geneva, 2010.

Buhaug, Ø., Corbett, J.J., Endresen, Ø., Eyring, V., Faber, J., Hanayama, S., Lee, D. S., Lee, D., Lindstad, H., Markowska, A. Z., Mjelde, A., Nelissen, D., Nilsen, J., Pålsson, C., Winebrake, J. J., Wu, W., Yoshida, K. (2009). *Second IMO GHG study 2009*, International Maritime Organization (IMO), London, UK, April 2009.

Corridor A/IQ-C. (2011). *6th progress report*. Executive Board ERTMS Corridor A and International Group for Improving the Quality of Rail Transport in the North-South Corridor. August 2011.

DB Schenker. (11. January 2011). *Trans Eurasia logistics and InterRail Services GmbH enter into "Ostwind" operations partnership*, Press release, Berlin. http://www.trans-eurasia-logistics.com/News/index.php.

DEFRA (UK) (2009). Guidance on measuring and reporting Greenhouse Gas (GHG) emissions from freight transport operations. http://archive.defra.gov.uk/environment/business/reporting/pdf/ghg-freight-guide.pdf.

Engström, R. (16. June 2011). *The Swedish green corridor initiative—history, current situation and thoughts about the future*. Proceedings of the 8th SoNorA University Think Tank Conference, Szczecin, Poland.

European Commission. (18. October 2007). *Freight transport logistics action plan*. Communication from the Commission COM (2007) 607, Brussels.

European Parliament and Council. (20. December 2013a). Regulation (EU) No 1315/2013 of the European Parliament and of the Council of 11 December 2013 on Union guidelines for the development of the trans-European transport network and repealing Decision No 661/2010/EU. Official Journal of the European Union L 348/1, 20.12.2013.

European Parliament and Council. (20. December 2013b). Regulation (EU) No 1316/2013 of the European Parliament and of the Council of 11 December 2013 establishing the Connecting Europe Facility, amending Regulation (EU) No 913/2010 and repealing Regulations (EC) No 680/2007 and (EC) No 67/2010. Official Journal of the European Union L 348/129.

Geitz, W. D., & Jia, N. (2010). Benchmark of environmental emission for railway Hinterland transport from the Port of Hamburg, Project Report, Railistics GmbH.

Ilves, I., Holte, E.-A., Georgopoulou, C., Rönkkö, S., Aditjandra, P., Vanaale, E., Medda, M., & Fozza, S. (2011). Benchmarking of green corridors (Version 2). SuperGreen project Deliverable D2.4, Document number: 02-40-RD-2011-14-02-1.

Kyster-Hansen, H., Thisgaard, P., Henriques, M., & Niss, M.-K. (21. June 2011), *Green corridor manual (draft)—purpose, definition and vision for green transport corridors*, Danish Transport Authority.

Lindstad and Mørkve. (26.–29. May 2009). A methodology to assess the energy efficiency and the environmental performance of maritime logistics chains. Conference proceedings, 10th International Marine Design Conference, Trondheim.

Maersk Line. (2007). Constant care for the environment, Brochure.

Mertel, R., & Sondermann, K-U. (6. December 2007). Final report for publication. BRAVO project.

Pålsson, C., Bengtsson, N., Salanne, I., Pütz, K., Tsilingiris, P., Georgopoulou, C., Schön, A., Minsaas, A., Wahl, A-M., Holte, E-A., Panagakos, G., Eiband, A., & Psaraftis, H. (2010). Definition of benchmark indicators and methodology. SuperGreen project Deliverable D2.2, Document number: 02-22-RD-2010-16-01-6.

Panagakos, G., Psaraftis, H., Moyano, H., & Holte, E-A. (2012). *Green Corridors handbook* (Vol. II). SuperGreen project, Document number: 07-20-PP-2012-05-02-3.

Psaraftis, H. N., & Kontovas, C. A. (2010). Balancing the economic and environmental performance of maritime transportation. *Transportation Research Part D, 15,* 458–462.

Psaraftis, H. N., & Panagakos, G. (2012). Green corridors in European surface freight logistics and the SuperGreen project. Procedia—Social and Behavioral Sciences, Special Issue: Proceedings of Transport Research Arena 2012, Vol. 48, pp. 1723–1732.

Swahn, M. (June 2010). How to evaluate transport's environmental performance. Article at *GreenPort Journal*.

Index

C.-Y. Lee, Q. Meng (eds.), *Handbook of Ocean Container Transport Logistics,*
International Series in Operations Research & Management Science 220,
DOI 10.1007/978-3-319-11891-8