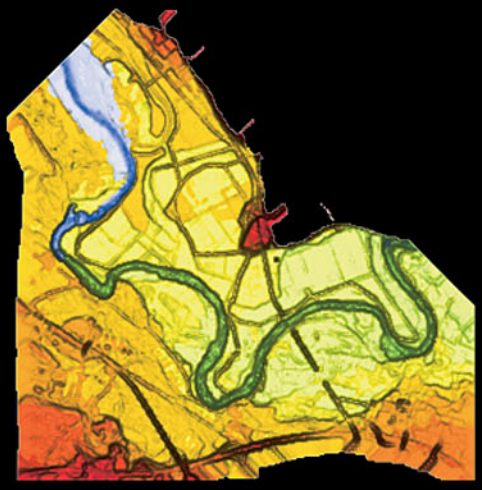
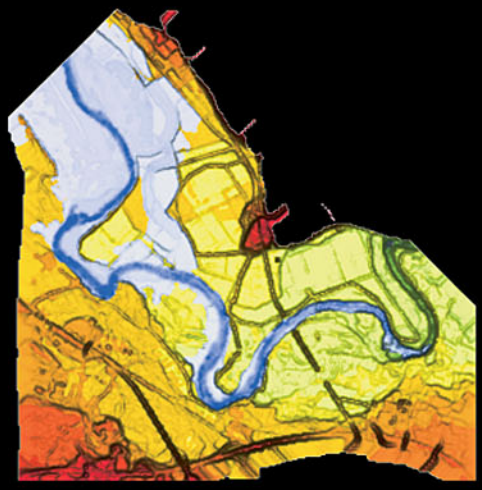
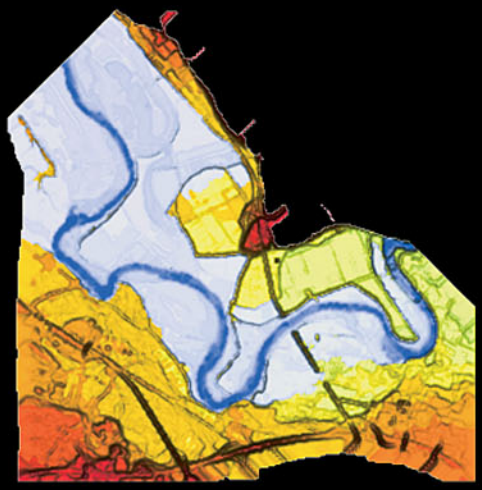






Numerical Modelling of Hydrodynamics for Water Resources

Pilar García-Navarro
Enrique Playán
EDITORS

Taylor & Francis
Taylor & Francis Group

NUMERICAL MODELLING OF HYDRODYNAMICS FOR WATER RESOURCES



BALKEMA – Proceedings and Monographs
in Engineering, Water and Earth Sciences

PROCEEDINGS OF THE INTERNATIONAL WORKSHOP ON NUMERICAL MODELLING OF HYDRODYNAMICS FOR WATER RESOURCES, ZARAGOZA, SPAIN, JUNE 18–21, 2007

Numerical Modelling of Hydrodynamics for Water Resources

Editors

Pilar García-Navarro
University of Zaragoza, Spain

Enrique Playán
E.E. Aula Dei, CSIC, Zaragoza, Spain



Taylor & Francis
Taylor & Francis Group
LONDON / LEIDEN / NEW YORK / PHILADELPHIA / SINGAPORE

Sponsors:



Universidad de Zaragoza



Consejo superior de Investigaciones Científicas



Gobierno de Aragón



Plan Nacional de I + D + I

This edition published in the Taylor & Francis e-Library, 2007.

"To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk."

Taylor & Francis is an imprint of the Taylor & Francis Group, an informa business

© 2008 Taylor & Francis Group, London, UK

All rights reserved. No part of this publication or the information contained herein may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, by photocopying, recording or otherwise, without written prior permission from the publishers.

Although all care is taken to ensure integrity and the quality of this publication and the information herein, no responsibility is assumed by the publishers nor the author for any damage to the property or persons as a result of operation or use of this publication and/or the information contained herein.

Published by: Taylor & Francis/Balkema
P.O. Box 447, 2300 AK Leiden, The Netherlands
e-mail: Pub.NL@tandf.co.uk
www.balkema.nl, www.taylorandfrancis.co.uk, www.crcpress.com

ISBN 0-203-93217-X Master e-book ISBN

ISBN 13: 978-0-415-44056-1 (hbk)
ISBN 13: 978-0-203-93217-9 (ebook)

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Table of Contents

Foreword ix

Preface xi

Keynote Lectures

Modelling overbank flows in rivers – data, concepts, models and calibration 3
D.W. Knight

Optimal design and management of water distribution systems 25
D. Savic, Z. Kapelan, R. Farmani & O. Giustolisi

The influence of source terms on stability and conservation in 1D hyperbolic equations: Application to shallow water on fixed and movable beds 47
J. Murillo, P. García-Navarro, P. Brufau & J. Burguete

Data-model integration – An approach to assimilation of sea level anomaly data into an oceanographic model 67
V. Babovic

Data-driven knowledge discovery: Four roads to vegetation-induced roughness formulae 77
V. Babovic

Modeling subsurface water flow and solute transport with HYDRUS and related numerical software packages 95
J. Šimůnek, M.Th. van Genuchten & M. Šejna

Limitations of depth-averaged modelling of shallow flows with recirculations 115
P.K. Stansby

Depth averaged turbulence models and source terms 127
L. Cea & M.E. Vázquez-Cendón

Application of distributed hydrologic models 153
R.F. Vázquez & J. Feyen

Erosion and sediment transport processes in Mediterranean mountain basins 175
S. Beguería, N. Lana-Renault, D. Regüés, E. Nadal-Romero, P. Serrano-Muela & J.M. García-Ruiz

Fundamentals of mathematical modelling of morphodynamic processes. Application to geomorphic flows 189
R.M.L. Ferreira

Numerical approach of sediment transport problems by high order finite volume methods 211
M.J. Castro, E.D. Fernández-Nieto, A.M. Ferreiro & C. Parés

Modeling and real time flash flood forecasting in a Mediterranean basin 233
I. Villanueva

Water quality in urban hydrology 251
J. Puertas & J. Suárez

Control of irrigation canals: Why and how? 271
P.O. Malaterre

A computer model for pipe flow irrigation problems 293
R. Aliod & C. González

Technical Contributions

Parallel computation of unsteady inundation shallow-water flow 317
B. Latorre, J. Murillo & P. García-Navarro

A 2D finite volume multiblock flow solver applied to flood extension forecasting 321
S. Erpicum, P. Archambeau, S. Detrembleur, B. Dewals & M. Pirotton

RKDG2 scheme for the 1D open channel flow 327
G. Kesserwani, J. Vazquez, R. Ghostine, R. Mosé & A. Ghenaim

Use of a nested modeling framework to study multiscale hydrological processes 333
J.E. Hunink, W. Bouten & E.E. van Loon

Furrow fertigation simulation model, calibration, validation and application 339
N. Zapata, J. Burguete, E.P. Santos, E. Playán & P. García-Navarro

Numerical methods for dispersion in water distribution pipes and networks 345
V.G. Tzatchkov

3D Modeling of hydrodynamics processes of lagoon-open sea interaction 351
D. Bellafiore & G. Umgiesser

On the wave-front shape and the advancing of the wetting front of a dam-break flood over an inclined plane of arbitrary bottom slope 355
P. Bohórquez

Roll waves in floods on inclines 361
P. Bohórquez

Effects of the irregularity in the cross sectional distribution of velocity on the 1D shallow-water model 367
J. Burguete, P. García-Navarro & J. Murillo

Conservative numerical boundary conditions 373
J. Burguete, P. García-Navarro & J. Murillo

A coupled 1D-2D modelling study of the 1907 flood in the Ebro delta 379
J. Raso & J. Hunink

Numerical solution of transient flow problems in soils: The nonlinear conductivity model 383
A. Arrarás, L. Portero & J.C. Jorge

Author index 389

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Foreword

This volume includes the lectures for the International Workshop Numerical Modelling of Hydrodynamics for Water Resources, held in Zaragoza, June 18–21, 2007, as well as the four page contributions presented by some of the participants. The purpose of the Workshop was to present current knowledge on steady and unsteady flow modelling in hydrodynamic systems for surface, subsurface and pressurized flows.

The editors of this volume trust that, in spite of some unavoidable difficulties in co-ordinating the work of several authors, the presented material will be an important contribution to professionals and researchers dealing with the problems of numerical modelling of hydrodynamic systems.

Pilar García-Navarro
Enrique Playán

Editors

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Preface

Numerical Modelling of Hydrodynamics for Water Resources is a generic title covering wide area of problems and rather wide field of interest. The present volume contains papers presented during an international workshop of 4 days at University of Zaragoza in Spain: published papers that a reader will find in the prefaced volume obviously do not cover whole gamut of problems and aspects announced in the generic title. At the same time the reader should be informed from the beginning that actual scope of presented papers is larger than hydrodynamics and that some of them are application paper while others deal with research activities fundamental for modelling.

Current situation on the field of numerical simulation modelling may be compared to that of the field of automobile: everybody knows that the machines we drive are not that efficient, that they are polluting, that they are only rough approximation to what is needed or possible and that they use unsustainable resource. Nevertheless most of us use them everyday and when commonly used models are modified, the modifications concern such things as more aesthetic profiles, better absorption of shocks during accident (your life may be saved but not your spine), new GPS linked to the road map, etc. Along the same line numerical simulation models became an everyday engineering tool and majority of those who apply them have no clear ideas about the limits of validity or possible errors and miscalculations that could result. As with the car drivers, most of the users see only GUI (Graphical User Interface) or "user-friendliness" of the software but not the engine. And, like most of the drivers with their cars, many (most?) of the users trust their software without knowledge of principles of the engine, not to speak of details.

Since 4th generation of the modelling software conquered the engineering market (i.e. ca 1990) most of the "engines" (i.e. equations and algorithms to solve them) embedded in commercial software did not change much. The requirements, however, presented to engineers by decision makers (investors, contractors, NGOs, public, administrations, etc., etc.) have been continuously increasing encouraged by general idea that everything can be modelled by computers, the latter being (of course) always right and, hence, the engineers have only "to simulate" to obtain the answers concerning impacts and consequences of more and more important works and projects. This evolution of social needs, as well as sheer scale of projects pushed engineers to apply software that is not always able to simulate the phenomena and situations resulting from this evolution. In turn, this pushed developers to produce new software without concurrent development of knowledge, theory and mathematical formulation. Further along this line it made some companies selling software, or studies based on simulation software, to produce results and reports that are not to be trusted but that nevertheless are trusted because, as said above, the decision makers have acquired "computer is right" convictions.

Just to give a couple of examples:

- 1D most often used models are applicable only to the situations for which De Saint-Venant hypotheses can be accepted; they cannot to simulate real life more complex phenomena depending upon transversal exchanges in complex river beds (as is clearly shown in Knight's paper of the prefaced volume).
- 3D modelling systems, becoming fashionable (because "more complete than 1D or 2D"), raise number of question marks concerning boundary conditions, but also validity of turbulence theories used as well as of limitations of hydrostatic pressure repartition hypothesis. Reading Stansby's and Knights papers one can see how many questions can be asked about spatial flows and turbulence simulation in much simpler cases than full 3D models…

- Sediment movement modelling for real-life cases is in sad condition because there is not enough knowledge about physics to formulate mathematically the phenomena (erosion, deposition, link with shear stresses, with velocities, with turbulence, with secondary flows, etc., etc.) alone, not to talk about related numerics and algorithmics.

There is also another important problem: a kind of wall that separates researchers working on numerical methods or on physical phenomena from the software developers. As observed above, the latter do not introduce last progress in numerics into commercial software, among other reasons because the "hearts" of the simulation systems are not discussed any more in scientific journals and, worse, they are sometimes even unknown to developers themselves!. Thus no criticism, comparison or progress is possible. In most papers at various Conferences one can read only about results of such or such system. The present volume is not 100% free of that, but there are also papers such as by Castro et al. or Aliod & González, giving the details of algorithms and showing how much they are important! This transparency in information concerning algorithms has led in time (1960s–1970s) to elaboration of engines of most of currently market-available simulation systems.

In a nutshell: it is of utmost importance that workshops such as the one in Zaragoza be organised and their content published. It can be hoped that under the pressure of information thus made available about the progress realised by applied research the decision makers ask more from simulation software industry, this leading to new qualitative jump in engineering tools. Because of that view it is interesting to read application papers published in present volume, such as by Babovic, Malaterre, Puertas & Suárez. They are presented at the same Conference as abovementioned papers informing on algorithmic details thus facilitating discussion and understanding of their approaches.

In recent years there have occurred a number of the meetings, conferences and publications in the domain of numerical modelling. Concurrently a number of scientific (and not only application to engineering projects) publications in that domain appeared and important progress on the subject of existence of solutions in function of boundary and initial conditions has been made. It seems essential that the community interested in this progress exchanges constantly the last ideas, developments and publications – hence the importance of present volume.

Jean A. CUNGE

Keynote Lectures

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Modelling overbank flows in rivers – data, concepts, models and calibration

D.W. Knight
The University of Birmingham, UK

ABSTRACT: The modelling of overbank flow is a technically challenging task due to the complexity of the 3-D flow structures that are present in most flooding rivers. As a result, there is a certain level of uncertainty in flood levels predicted by many numerical models. The development of a simple method for estimating the stage-discharge relationship for straight reaches of any river with a prismatic shape is described. It begins with initial concepts, based on data from experimental studies in the UK Flood Channel Facility, through to mathematical formulation, and then to their eventual inclusion in a software package, known as the Conveyance Estimation System. The method involves integration of the depth-averaged velocity field across the river, and includes the effects of bed shear, lateral shear and secondary flows. The model's potential is illustrated by its application to a number of rivers world-wide.

1 INTRODUCTION

The development of a depth-averaged velocity model for fluvial engineering is described. In the development of any model there are normally specific stages, beginning with the concepts that arise from observations, either in the laboratory or the field, leading to identification of key flow mechanisms, based on sound physical principles and amenable to mathematical formulation, through to solution methodology and eventual incorporation into some software. In this particular example, the ensuing equations are integrated over the depth, solved and incorporated into a dedicated software package, known as the Conveyance Estimation System (CES). It is available on the web for practitioners and researchers alike, and deals with the lateral distribution of velocity and boundary shear stress in prismatic channels. Details of the model and its practical application may be found in Shiono & Knight (1991), Abril & Knight (2004), McGahey et al. (2006), Knight et al. (2007) and at www.river-conveyance.net.

Flows in straight compound channels have been studied by many researchers in recent years, and comprehensive reviews are provided by Ikeda & McEwan (2007), Knight & Shiono (1996) and Knight (2005). Overbank flows in meandering channels create other flow structures, as shown by Fukuoka et al. (1996), Shiono & Muto (1998) and Willetts & Rameshwaran (1996).

2 THE IMPERATIVE OF GOOD QUALITY DATA

Physical concepts about fluid flow have first to be recognized and then incorporated in some mathematical way into a model. Recognition usually, but not always, comes through observation and analysis of data. Data is thus crucial at the developmental stage of the modelling process, as it is also at the calibration stage, when certain parameters are used in tuning the model. In river modelling there is typically never enough data for adequate calibration at suitable spatial or temporal scales. Table 1 summarizes the sources of experimental data used in developing this model and Figs 1–19 illustrate the types of apparatus often used in studying overbank flow.

Table 1. University of Birmingham database (www.flowdata.bham.ac.uk).

Name	Date of PhD	No. of exps	Facility	Type of channel/duct
Abril	1997	many	∼	modelling FCF data
Alhamid	1991	38	Bham, 22 m flume	trapezoidal, heterogeneous
Atabay	2001	50	Bham, 18 m flume	rect compound, rigid & mobile
Ayyoubzadeh	1997	25	Bham, 18 m flume	small scale, rigid & mobile
Brown	1997	26	FCF, HR Wallingford	FCF, large scale, mobile
Lai	1986	61	Bham 11 m wind tunnel	compound duct, variable geometry
Mohammadi	1998	10	Bham 15 m & 9 m flumes	V-shaped channels + others
Patel	1984	66	Bham 11 m wind tunnel (variable width)	rect & compound duct, smooth & rough
Rhodes	1991	48	Bham 25 m wind tunnel (see Figs 3 & 4)	rect and compound very wide cases, immense detail
Sterling	1998	24	Bham 22 m flume (re configured)	circular part full and with flat beds (∼ deposited sediment)
Yuen	1989	75	Bham 22 m flume (see Figs 1 & 2)	trapezoidal & compound (narrow) at critical flow
Knight	1970–85	150	numerous flumes	rect simple & compound
Knight	1987–89	74	FCF, HR Wallingford	compound (15 Volumes)
Knight	1990–96	many	FCF, HR Wallingford	meandering simple/compound
Knight	1980–04	many	∼	collected sets from elsewhere



Figure 1. Small wind tunnel (11 m) with roughened floodplains, University of Birmingham.

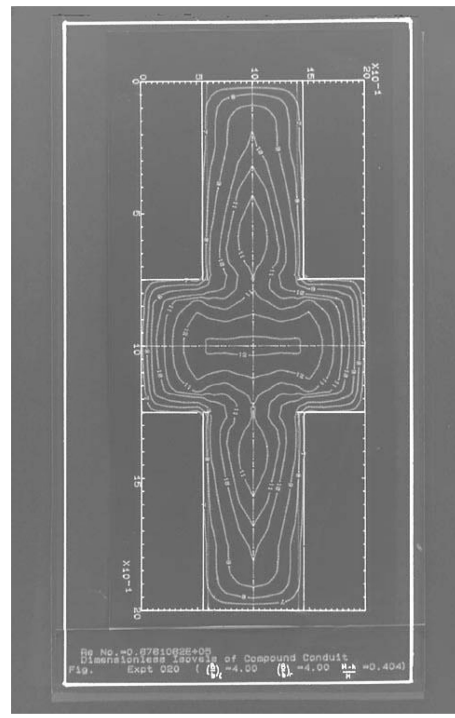


Figure 2. Isovels within small wind tunnel for whole cross-section (Dr = 0.404).

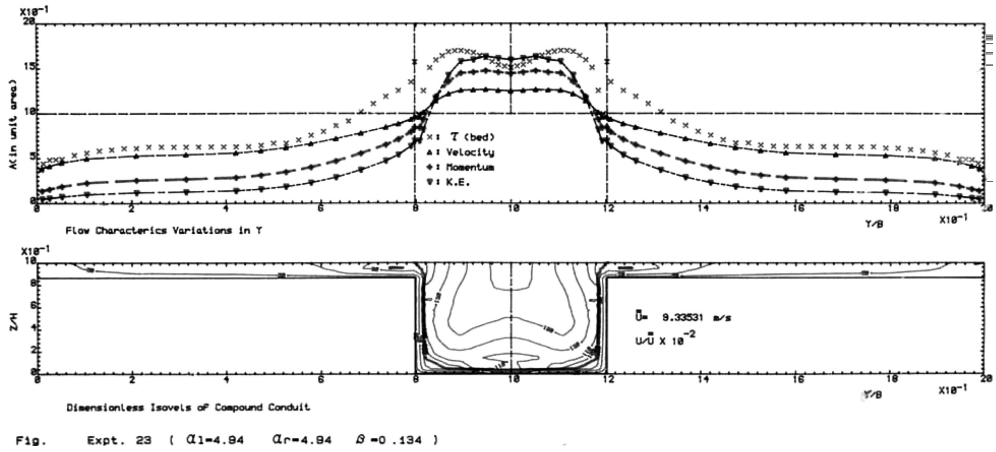


Figure 3. Half section velocity & boundary shear stress data for Dr = 0.134 (small wind tunnel).

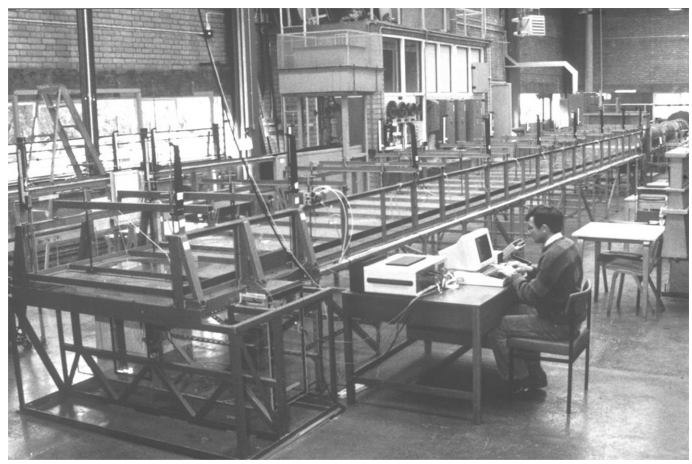


Figure 4. Large wind tunnel (25 m), University of Birmingham.

Figs 1–3 illustrate the use of a small wind tunnel to study issues related to flooding problems. The cross-shaped duct, 11 m long and with a working cross section 400 mm × 200 mm, had a smooth centre region and two roughened regions. Movement of one whole side of the wind tunnel altered the gap width, which may be interpreted as a 'depth' change. When data from the whole cross section, as illustrated in Fig. 2 for a high relative depth of 0.404, is split in half and rotated through 90°, the half-section mimics a river with two symmetric floodplains, as shown in Fig. 3 for a low relative depth of 0.134. In this case the true free surface conditions are violated, but significant lessons may be learnt about the behaviour of secondary flows at re-entrant corners, the distribution of energy and momentum across the duct/channel, the proportion of total shear force acting on different boundary elements and momentum transfer mechanisms.

Figs 4–11 illustrate a similar approach, but using a larger wind tunnel, equipped with an automatic data gathering system for measuring boundary shear stresses and streamwise velocity in great detail. One of the main objectives here was to investigate secondary flows at the main channel/floodplain interface for different 'river bank' side slopes using as 'wide' a main channel as

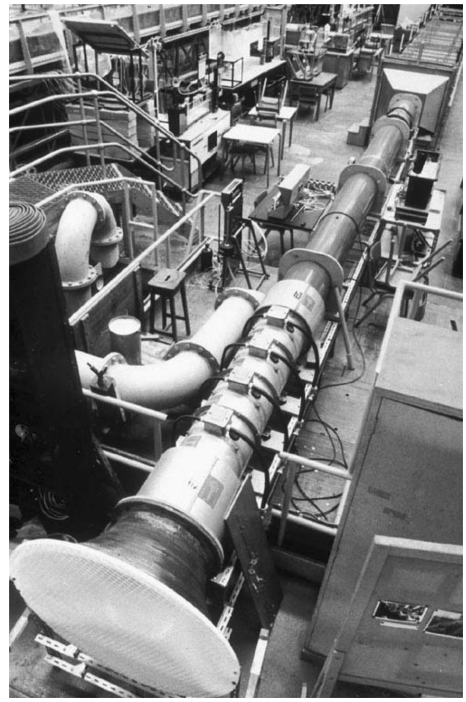


Figure 5. Blower fans (25 m wind tunnel).



Figure 6. Cross-section of wind tunnel.

possible, as shown by the cross-section in Fig. 6 and the results in Fig.7. Some lateral distributions of depth-averaged velocity and boundary shear stress are shown in Figs 8 & 9. Shear layer widths, eddy viscosities and boundary conditions in models (Knight, Omran & Abril, 2004) were also drivers in this study. Ancillary experiments were conducted in rectangular ducts with aspect ratios as high as 50 in order to investigate shear stresses in corners, as illustrated in Fig. 10 & 11. See Rhodes & Knight (1994a&b, 1995) for further details of these studies.

Figs 12–19 illustrate a variety of channels and flumes, ranging from the 10 m wide and 56 m long Flood Channel Facility (FCF), shown in Fig. 12, to the narrow flume shown in Fig. 13 which

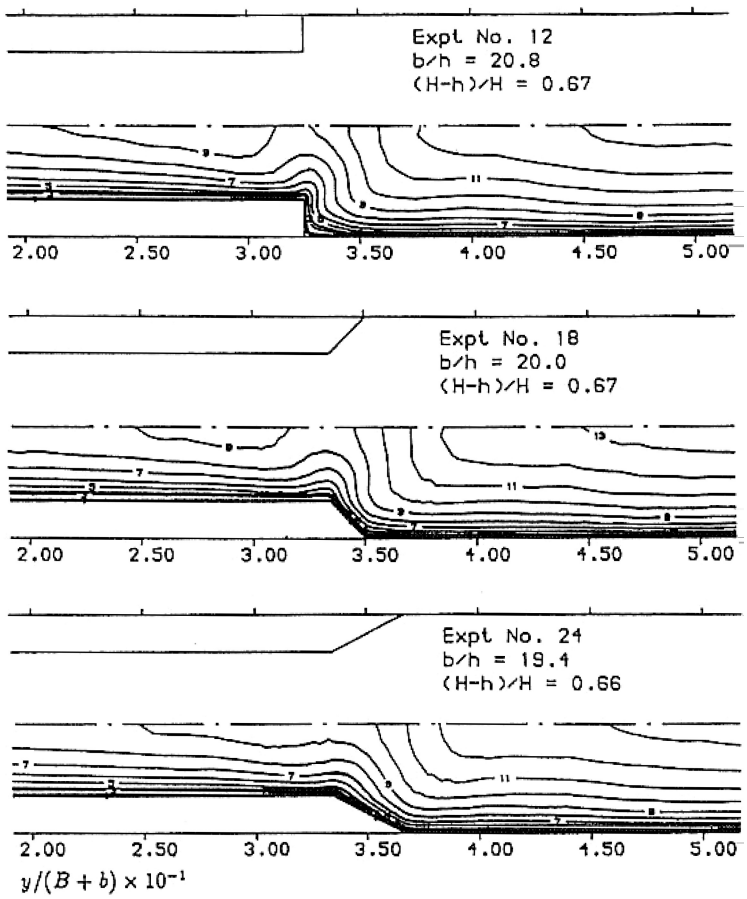


Figure 7. Isovels at the interface between a main channel and its floodplain for different main channel side slopes, $s = 0$, 1 & 2 (Rhodes & Knight, 1994b).

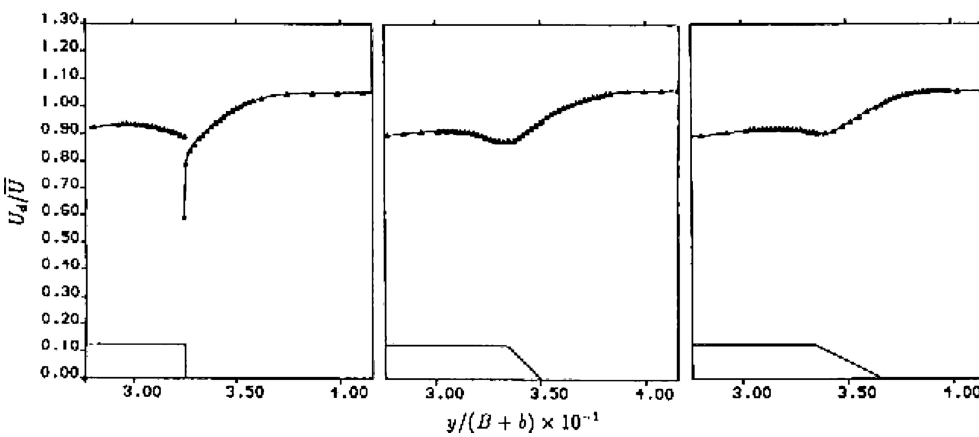


Figure 8. Depth-averaged velocity distributions for $s = 0$, 1, & 2 (Rhodes & Knight, 1994b).

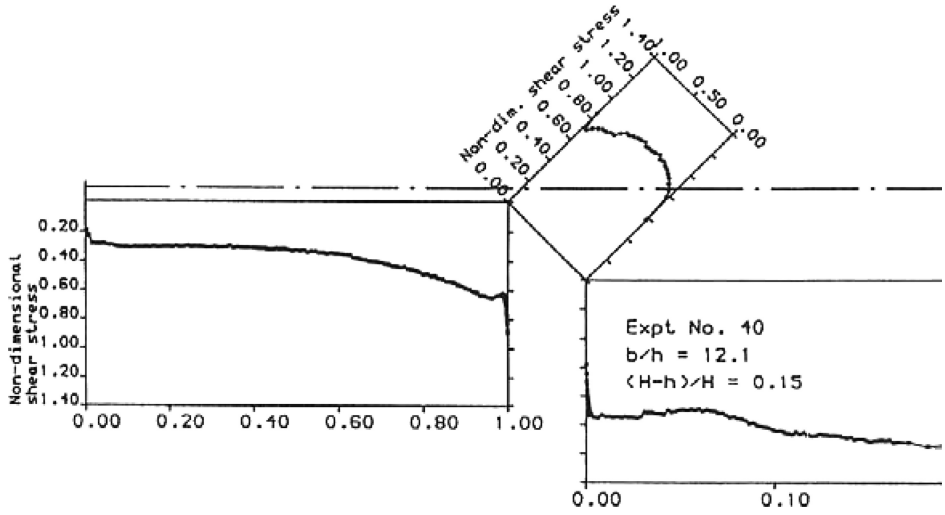


Figure 9. Boundary shear stresses, showing shear layer widths (Rhodes & Knight, 1994b).

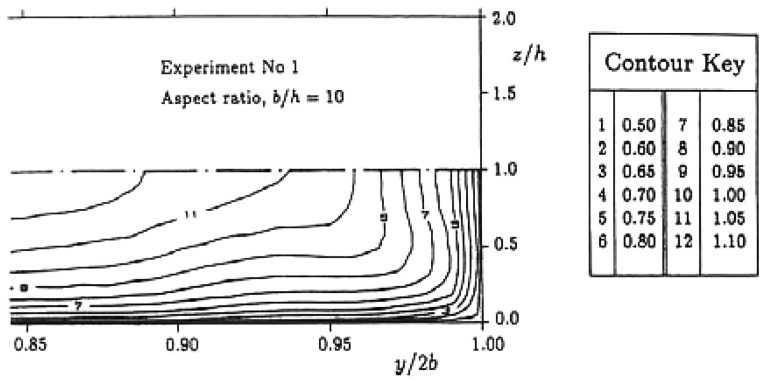


Figure 10. Isovels in the corner of a rectangular channel (Rhodes & Knight, 1994a).

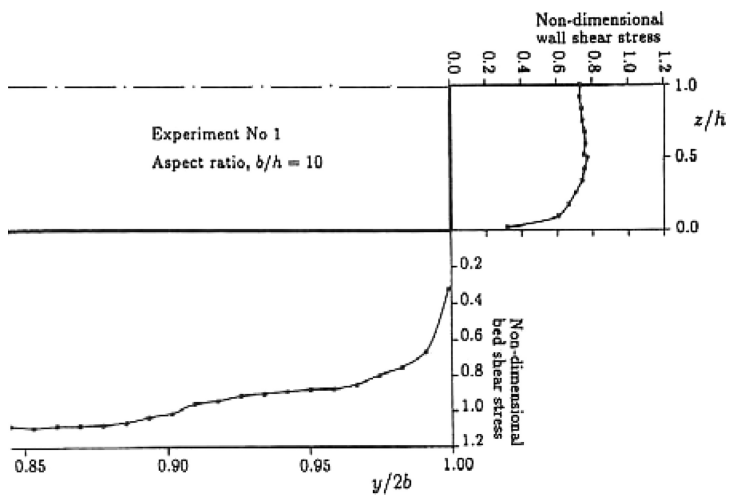


Figure 11. Boundary shear stresses in the corner of rectangular duct (Rhodes & Knight, 1994a).

(a)



(b)



(c)



Figure 12. Flood Channel Facility (FCF) at HR Wallingford (rigid boundary channels) (a) straight channel; (b) skew channel; (c) meandering channel.



Figure 13. Simple trapezoidal channel, University of Birmingham.



Figure 14. Compound trapezoidal channel, University of Birmingham.



Figure 15. Asymmetric compound channel with re-circulating sediment system.



Figure 16. Symmetric compound channel with bridge crossing at 90°.

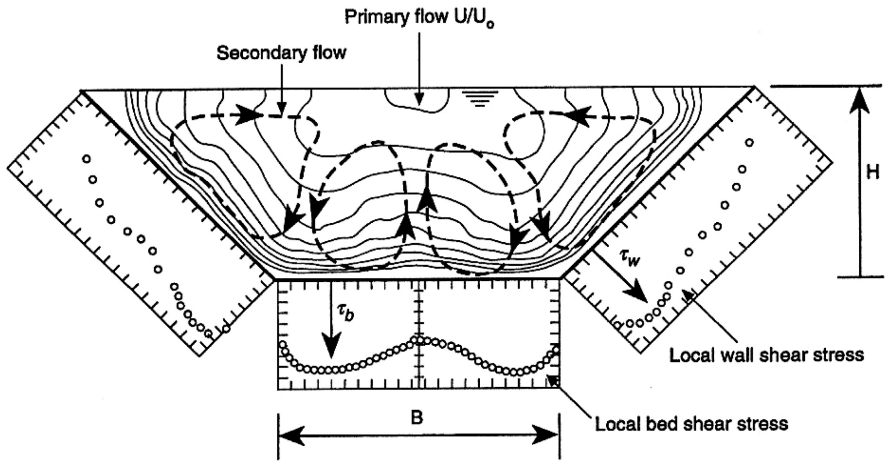


Figure 17. Isovels and boundary shear stresses in a trapezoidal channel.



Figure 18. Investigating regime channels and analytical regime theories.

was 0.6 m wide and 22 m long. These were all constructed with specific experimental objectives in mind, and therefore had different features. For example the FCF was designed as a fixed bed slope flume for large scale experimental work on straight, skewed and meandering channels, with either rigid or loose boundaries. The small flumes were designed for variable bed slope work in sub and supercritical flow (rigid bed cases) or sediment work in either re-circulating or direct sediment feed systems for regime or sediment transport studies (loose boundary cases). Figs 15, 18 & 19 illustrate some loose boundary work in straight channels.

Much of the data obtained from these studies may be found at www.flowdata.bham.ac.uk for laboratory based work, including that on the FCF, and at www.river-conveyance.net for field based studies. Alternatively the field data are given in McGahey (2006) or summarised in McGahey et al. (2006). The University of Birmingham laboratory data contains over ~600 complete sets of primary velocity (U) and boundary shear stress (τ_b) data for flow in straight prismatic open channels or ducts with various cross sections and roughness distributions. A typical example is shown in Fig. 17, which was derived from experiments in the flume shown in Fig. 13. The database is intended

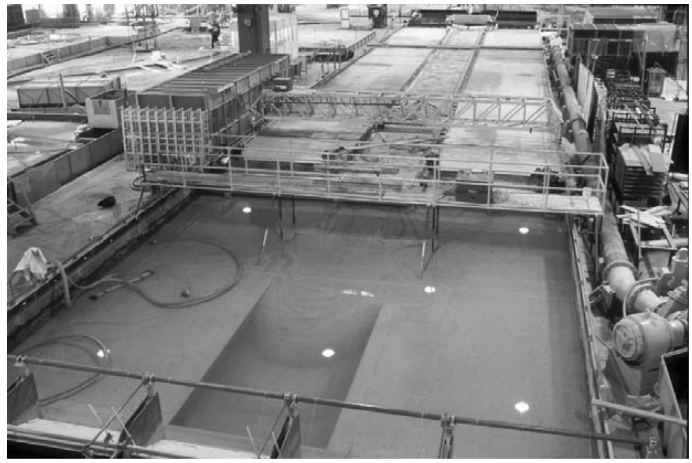


Figure 19. Flood Channel Facility (FCF) at HR Wallingford (loose boundary channel).

primarily for researchers who wish to obtain experimental data of a detailed nature, together with information about how it was measured, for use in conjunction with their own experimental studies or in validating numerical models.

3 CONCEPTS

Experimental data are not sufficient in themselves to produce models, unless intelligent use is made of them in statistical or conceptual terms, and appropriate deductions made from the observations. For example, Fig. 20 illustrates some of the key mechanisms thought to be present in overbank flow in straight prismatic channels. Fig 21 shows the appearance of plan form vorticity in a laboratory channel with one flood plain, a phenomenon also observed in large engineered rivers with wide floodplains (Knight et al., 2006). Fig. 22 shows the same channel at a slightly larger relative depth ratio, Dr, of 0.344, where Dr = floodplain depth/depth in main river channel. The streamwise vorticity at the main channel/floodplain interface is now sufficiently strong to dominate the flow in that particular region, masking the large scale planform vorticity seen in Fig. 21 at the lower relative depth of Dr = 0.180.

Fig. 23 shows the effect of these planform vortices on the level of the free surface, a topic that has been studied in some detail because of its importance in overbank sediment transfers and flood analysis (e.g. Prooijen, 2004; Bousmar & Zech, 2002). The free surface perturbations are caused by periodic lateral variations in the water surface slope, which drive water onto and off the flood plains, as shown by Prooijen (2004). At the main channel edge with the floodplain, it is the combination of planform and streamwise vorticity that influence the flow field, as already seen in Figs 21 & 22.

The streamwise vorticity, or secondary flow cells, are dealt with in the depth-averaged model in the way shown in Fig. 24. In this Figure there is an anti-clockwise cell that gives rise to the transverse velocity distribution shown, $V = f\{z\}$, with V being negative at the free surface according to the axes chosen. In the streamwise direction, x, there will be a longitudinal velocity distribution, $U = f\{z\}$, which in a wide channel would typically be logarithmic. When the product of UV is determined as a function of z, and then depth-averaged, it will yield a single value at a particular position, y, depending on the sense of the rotation of the secondary flow cell. In Fig. 24, with an anti-clockwise rotation, $(UV)_d$ will be negative at the cell centerline. At the edges of each cell V tends to zero, thus making the value of $(UV)_d$ also zero at these locations. It is then assumed that

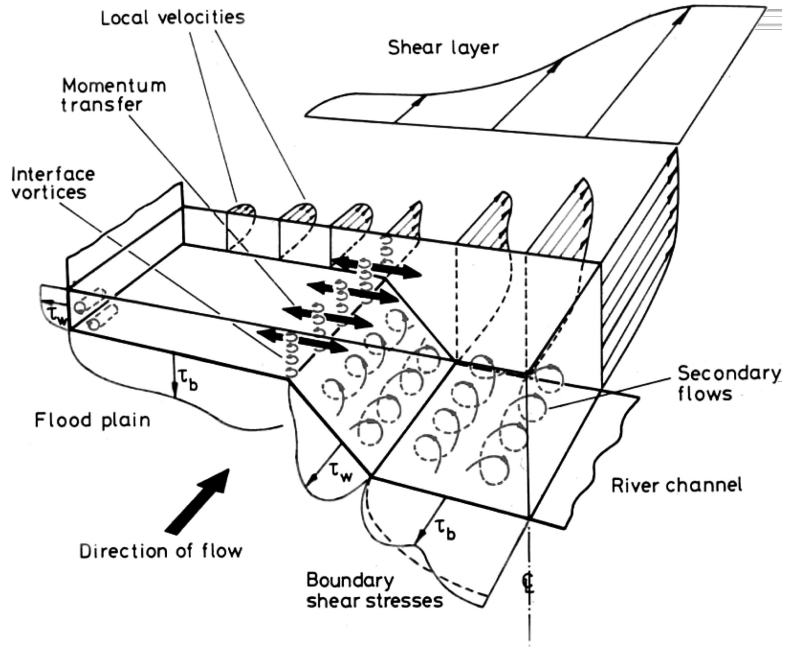


Figure 20. Hydraulic aspects of overbank flow (after Knight & Shiono, 1996).

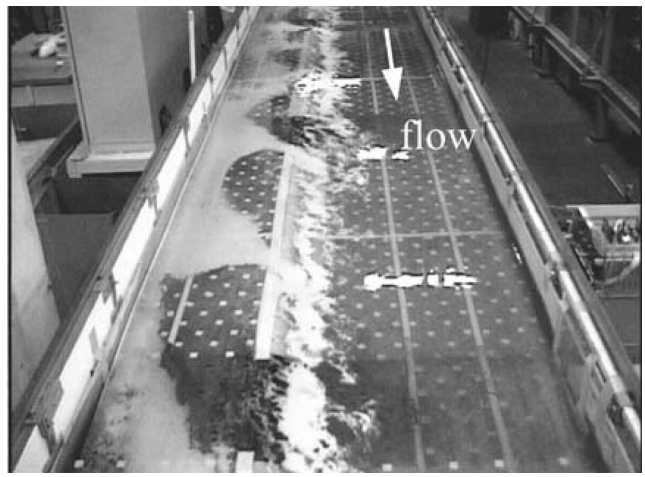


Figure 21. Low relative depth, Dr = 0.180.

there is a linear variation in $(UV)_d$, from a maximum at the cell centre to zero at the extremities, leading to a gradient term which herein is labelled as Γ. Thus there will be an alternating pattern of Γ values across a channel, linked to the number and position of flow cells. This is illustrated in Fig. 25, taken from Omran & Knight (2006), in which there are 3 cells in one half of a trapezoidal channel, as deduced for that particular depth and discharge from the experimental data of Tominaga et al. (1989). In this case 6 panels are then required to model the secondary flow structures, although 4 panels (P1-P4) may be adequate for most purposes. When precise shear stress predictions are

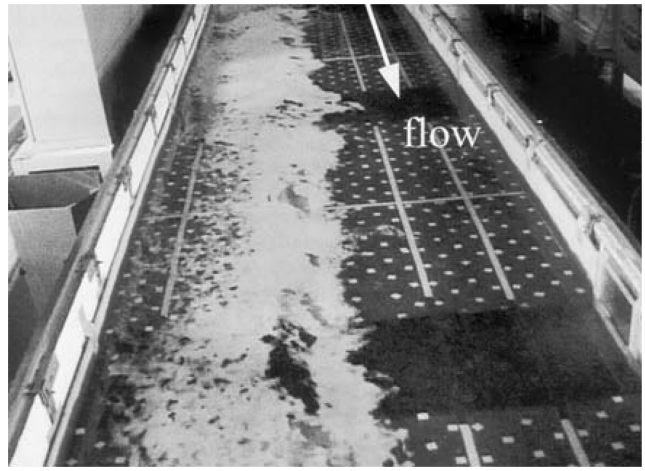


Figure 22. High relative depth, Dr = 0.344.

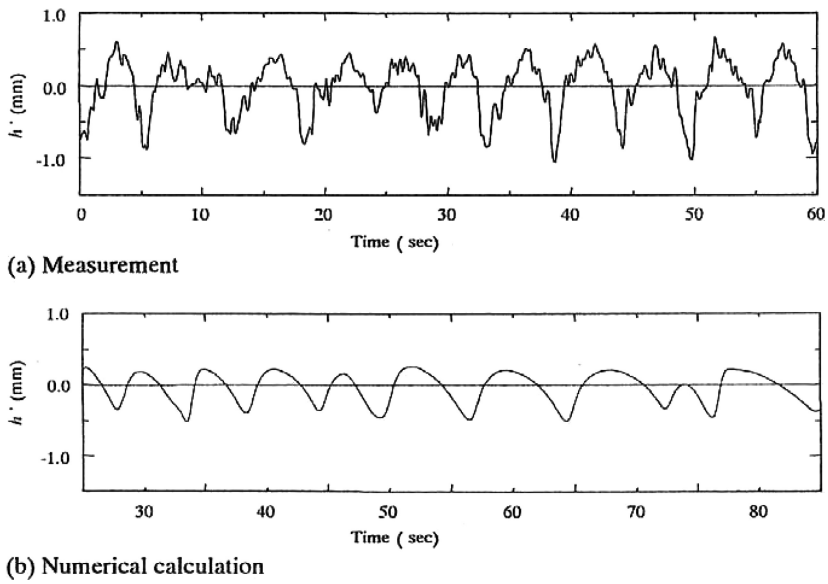


Figure 23. Free surface oscillations produced by planform vorticity (Ikeda & McEwan, 2007).

needed in the corner region, two additional panels (P2a & P3a) will improve the accuracy in that region.

4 MODEL DEVELOPMENT

Having established some of the dominant physical ‘concepts’, these flow structures then need to be formulated mathematically and incorporated into the model. An example of the movement from ‘concepts’ to ‘model development’ is shown in the work of the ASCE Task Force on River Width Adjustment (Thorne et al., 1998; Darby et al., 1998).

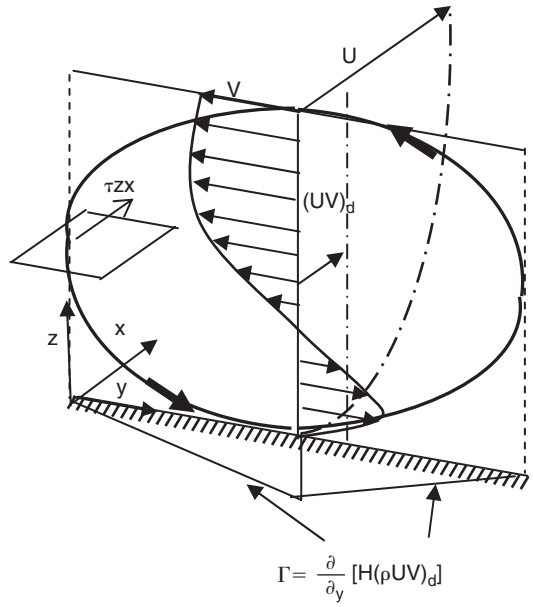


Figure 24. Depth-averaged secondary flow term for a single anti-clockwise cell.

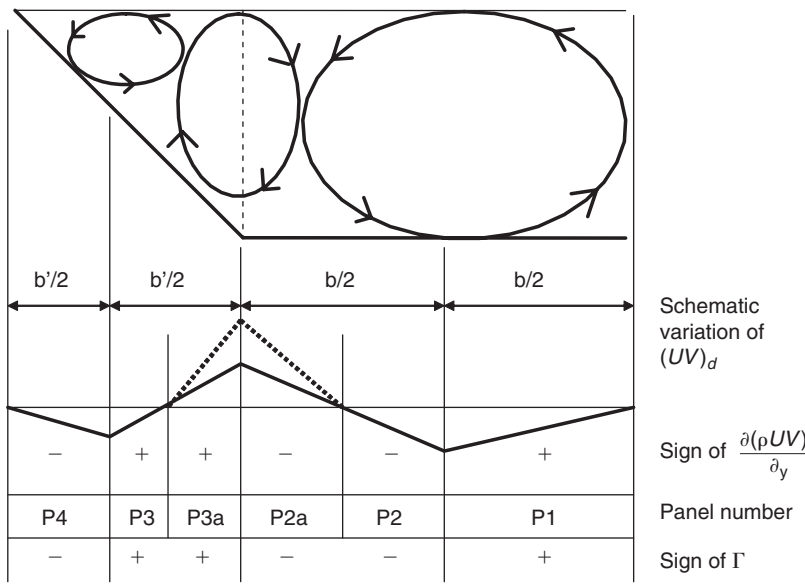


Figure 25. Number of panels used in a trapezoidal channel simulation and signs of secondary current term (after Omran & Knight, 2006).

There are of course many different types of model, each suited to the ends for which they were developed (e.g. hydrodynamics, flood routing, sediment motion, pollution, wave phenomena, etc.). Depending on the particular hydraulic problem being investigated, models may be developed for different spatial dimensions (1-D, 2-D & 3-D), have different levels of turbulence closure (e.g. using RANS models with mixing length, k-ε, Reynolds stress, SSG, k-ω, etc.), and may be based on the

Reynolds Averaged Navier-Stokes (RANS) approach, the Large Eddy Simulation (LES) approach or even undertaken by Direct Numerical Simulation (DNS).

Since turbulence itself is still not yet precisely formulated mathematically (see Jakirlic & Hanjalic, 2002), the level of turbulence closure, i.e. the degree of complexity admitted in several governing equations, has to be chosen carefully. The engineer has therefore to blend pragmatically theoretical and empirical knowledge together in order to develop any type of model at all. Furthermore, the numerical procedures may vary between models even of the same type. It is thus a matter of selecting the right 'tool' for the right 'job', being aware of the shortcomings and uncertainties in the chosen methodology.

The model described herein is one aimed at answering the following technical questions that are frequently posed by river engineers:

- How does one determine the stage-discharge relationship at a gauging station?
- How does one rationally extend the HvQ relationship to predict higher flood levels?
- How does one predict the depth-averaged velocity distribution across a river?
- How does one estimate the boundary shear stresses around the wetted perimeter?

The model is therefore a quasi 1-D one, but it includes some of the key 3-D flow structures identified earlier that occur in overbank flows. A depth-averaged equation can be obtained by integrating the RANS momentum equation over the water depth, H, to give:

$$\frac{\partial[H(\rho UV)_d]}{\partial y} = \rho g H S_0 + \frac{\partial H \bar{\tau}_{yx}}{\partial y} - \tau_b \sqrt{1 + \frac{1}{s^2}} \tag{1}$$

in which τ_b is the bed shear stress and s is the side slope (1:s, vertical: horizontal), and $\bar{\tau}_{yx}$ is a depth-averaged Reynolds stress on a vertical plane. By using a standard resistance law, but defined for local conditions, equation (1) may be developed to give the depth-averaged velocity, U_d, as a function of y in the form:

$$\rho g H S - \rho \frac{f}{8} U_b^2 \sqrt{1 + \frac{1}{s^2}} + \frac{\partial}{\partial y}\left[\rho \lambda H^2 \sqrt{\frac{f}{8}} U_d \frac{\partial U_d}{\partial y}\right] = \frac{\partial}{\partial y}[H(\rho UV)_d] \tag{2}$$

In this model, the Reynolds stress term, $\bar{\tau}_{yx}$, is dealt with by using a simple eddy viscosity approach, with λ as the dimensionless eddy viscosity coefficient ($\bar{\varepsilon}_{yx} = \lambda U_* H$) and U_* as the local shear velocity ($= \tau_b/\rho)^{1/2}$. Experimental evidence suggests that the secondary flow term on right hand side of equation (2) may be approximated by constant values for a given panel or region, as already indicated in Fig. 24. This allows one to account for all secondary flow cell strengths for inbank flows, through adjustment of the Γ values, and also to deal with planform vorticity for overbank flows. Further details of the background to this model may be found in Shiono & Knight (1991), Knight & Abril (1996), Abril & Knight (2004), McGahey et al. (2006) and Knight, Omran & Tang (2007).

5 MODEL CALIBRATION

It is clear from equation (2) that the model depends upon schematising the cross-section into a number of regions or panels, and then assigning appropriate values to the three calibration parameters f, λ and Γ that govern bed resistance, lateral shear and secondary flows respectively. Knowledge about these 3 parameters has been formulated by Knight & Abril (1996) and Abril & Knight (2004), based mainly on the experimental studies conducted in the large scale Flood Channel Facility (FCF). Each parameter is known to vary with discharge, depth and channel geometry, and equations that describe these variations have been proposed.

Details of how secondary flows affect the distribution of boundary shear stress around the wetted perimeter of channels may be found in Knight et al. (1994), Lai & Knight (1988), Knight et al.

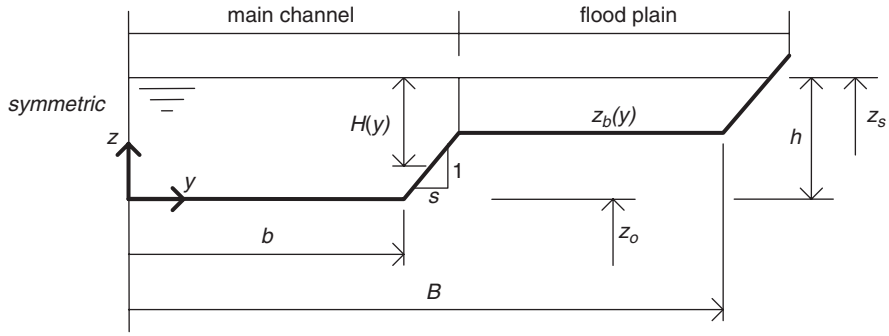


Figure 26. Schematised cross-section of a compound channel, as used in FCF experiments.

(1992) and Atabay & Knight (2006). Since τ_b is linked to f by $\tau_b = \rho f\ (U_d^2)/8$, where f is a 'local' friction factor, experimental measurements of U_d and τ_b suggests that f remains constant in certain regions for overbank flows. Consequently the ratio of floodplain resistance to main channel resistance may be expressed in simple mathematical terms, as shown by Shiono & Knight (1991) and Abril & Knight (2004). Defining the ratio of floodplain depth to main channel depth as Dr, then these might be of the form

$$f_{\rm fp} = f_{\rm mc}(a{\rm Dr}^b)\ ;\lambda_{\rm fp} = \lambda_{\rm mc}(-0.2 + 1.2{\rm Dr}^{-1.44}) \qquad \text{(3a \& 3b)}$$

where the subscripts refer to the floodplain (fp) and main channel (mc) respectively. In order to use such equations in a model, then it is only necessary to know or estimate the values of f and λ for the main river channel. Typically f may be estimated from the bankfull discharge, and λ may be taken as 0.07 unless other data are available.

Since most of the calibration data were obtained from schematised compound channels with a cross-sectional shape as shown in Fig. 26 (see www.flowdata.bham.ac.uk), it was noted that $\tau_{\rm b}$ is mainly influenced by the lateral variation of the Reynolds shear stress $\partial\overline{\tau}_{yx}/\partial y$, the secondary flow contribution $\Gamma^* = \partial(\rho\overline{UV})_d/\partial y$ and the gravitational term $\rho g H S_0$, where Γ^* is

$$\partial\left(\rho UV\right)_d/\partial y = \Gamma^* = \Gamma/H \qquad (4)$$

From both the experimental data (e.g. see Knight & Shiono (1996), Fig. 5.19) and from numerical simulations (e.g. see Knight & Abril, 1996, Fig. 2) it is known that $\overline{\tau}_{yx}$ only varies significantly within the shear layer region at the main channel/flood plain interface, as shown in Fig. 28. This also agrees with numerical results obtained in three dimensional large eddy simulations by Thomas & Williams (1995). Consequently, outside those regions, assuming average lateral values of the boundary shear $\tau_{\rm avg}$ and of the secondary flow contribution $\Gamma^*_{\rm avg}$ (e.g. later- ally-averaged within the main channel or flood plain) Equation (1) may be written as

$$\rho g H S_0 - \tau_{\rm avg} = \frac{\partial}{\partial y}\left[H\left(\rho\overline{UV}\right)_d\right] = H\Gamma^*_{\rm avg} \qquad (5)$$

Substituting $\tau_{\rm avg} = k\,\tau_{2d} = k\,\rho g H S_{\rm o}$ into Equation (5) gives

$$H\Gamma^*_{\rm avg} = \rho g H S_0 - k\tau_{\rm 2d} = \rho g H S_0(1-k) \qquad (6)$$

or

$$\Gamma^*_{\rm avg} = \rho g S_0(1-k) \qquad (7)$$

Equation (7) indicates that there is a link between the influence of the secondary flow term and the bed shear stress distribution. Fig. 27 indicates how the k values in Equation (5) depend

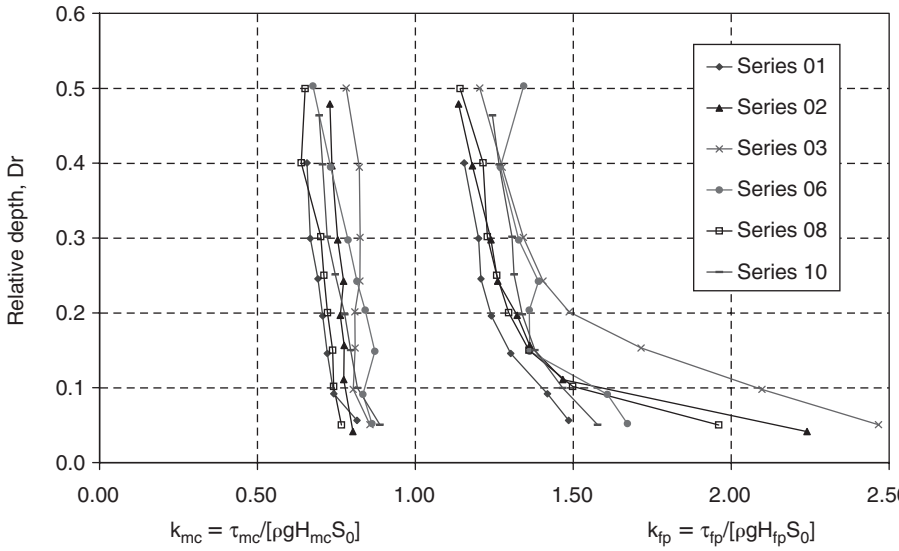


Figure 27. Main channel and flood plain bed shear stresses ratios (FCF series 1–10).

on both the relative depth and the geometry of the flood plains for each Flood Channel Facility (FCF) series. In the main channel k_{mc} values remain fairly constant with depth ($k_{mc} = 0.85$, giving $\Gamma^*_{mc} = 0.15\ \rho g S_o$), making the calibration relatively straight-forward. On the flood plain the influence of vorticity, both streamwise secondary flow and planform vorticity, becomes more important for lower relative depths. This is not surprising, since the lateral transfer of momentum induced by the difference of velocities between the two regions is stronger. Taking average values $k_{mc} = 0.85$ and $k_{fp} = 1.25$, the secondary flow terms may be roughly estimated as

$$\Gamma^*_{mc} = 0.15\ \rho g S_o \quad \text{and} \quad \Gamma^*_{fp} = -0.25\ \rho g S_o. \tag{8}$$

giving the ratio between them as $\Gamma^*_{fp}/\Gamma^*_{mc} = -0.25/0.15 = -1.6$ for relative depths Dr greater than 0.2, as used in Abril & Knight (2004).

In a fully 3-D model then the distributions of other parameters such as Reynolds stress, turbulence intensities and energy spectra should also be considered. Figs 28 & 29, taken from Shiono & Knight (1991), illustrate some values of Reynolds stress measured in overbank flows. Fig. 28 shows that the peak values of $\overline{\tau}_{yx}$ occur at the main channel/floodplain interface ($y = 0.9$ m), which is often taken as a dividing line between panels or regions. Fig. 29 shows that in this region the distribution of τ_{zx} is highly non-linear (e.g. see $y = 0.81$ m), indicating that the streamwise velocity distribution does not follow the conventional logarithmic law.

6 MODEL APPLICATION

In order to test the calibration rules put forward by Abril & Knight (2004), based on laboratory work, a number of detailed simulations were undertaken on 18 rivers worldwide using both inbank and overbank flow data. The 18 river sites were selected according to the specific properties and scales that were of interest. They covered a reasonably wide range of hydraulic conditions such as slope, roughness and size, and are summarised in tabular form by McGahey (2006). Key information concerning 14 of these rivers is also given in McGahey et al. (2006).

The model described earlier forms part of the 'Conveyance Estimation System' (CES), which was developed as a software 'tool' for the Environment Agency, UK, to assist practising engineers

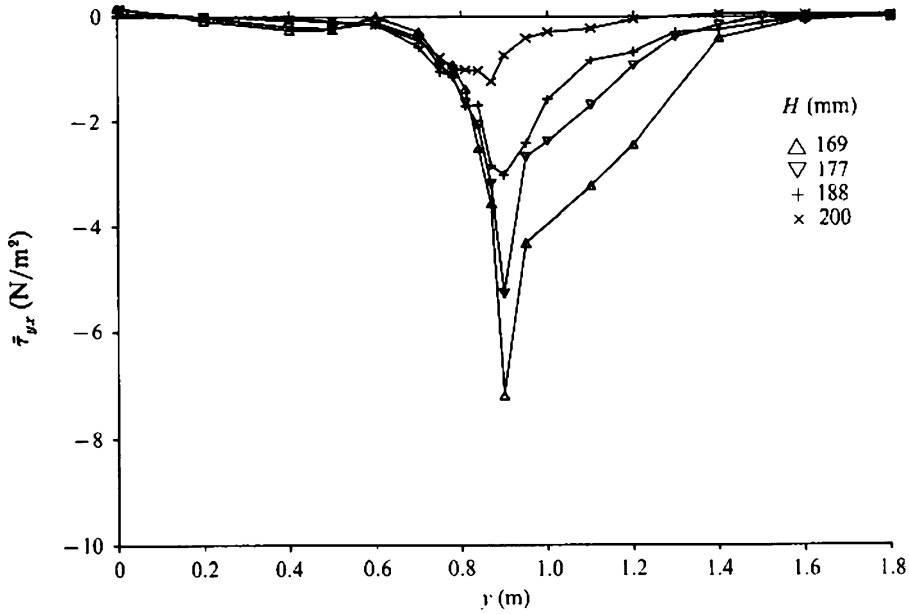


Figure 28. Reynolds stress distributions, $\overline{\tau}_{yx}$ versus y (FCF data).

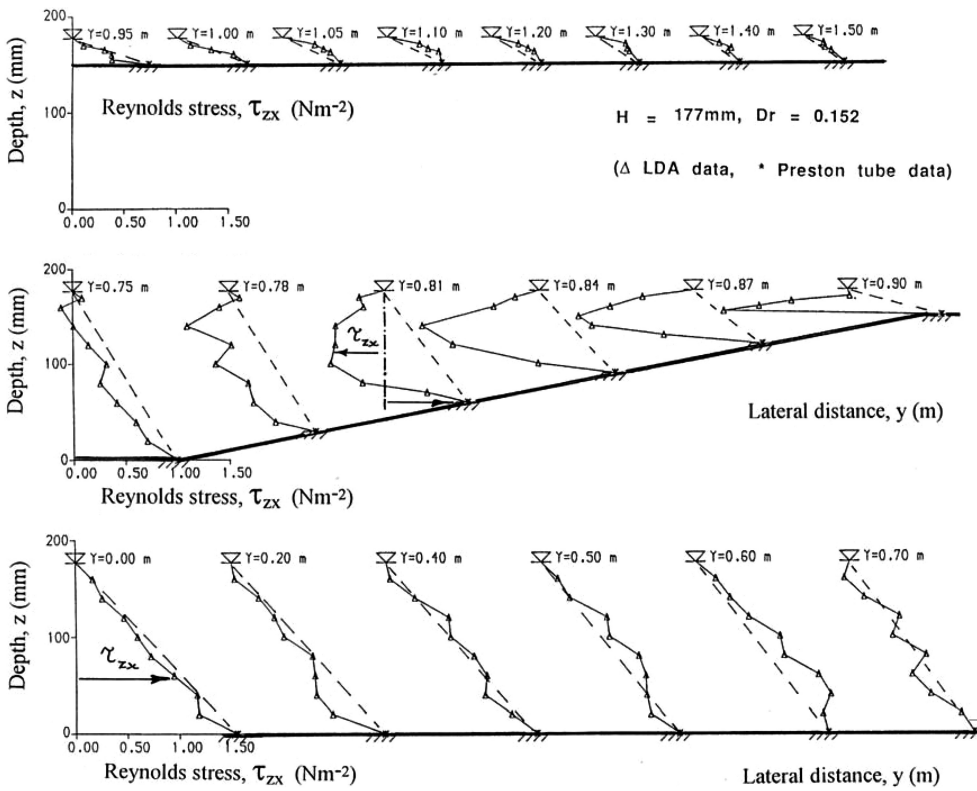


Figure 29. Variation of Reynolds stress, τ_{zx} with z & y (FCF data).

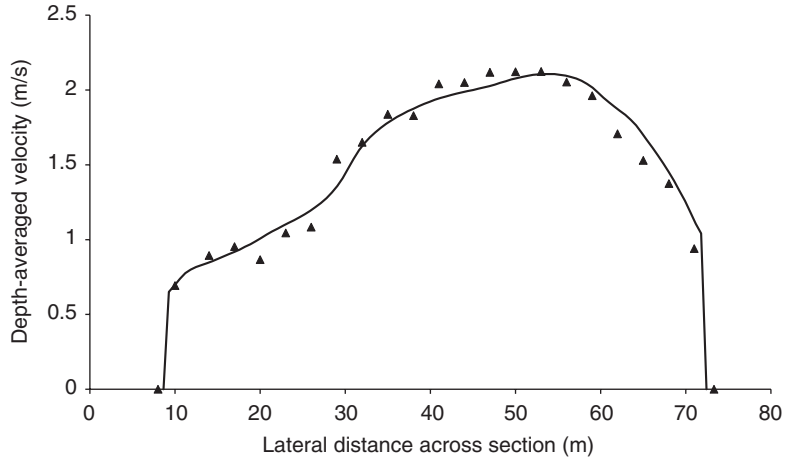


Figure 30. Velocity U_d for depth 2.494 m; $Q_{data} = 181.98\ m^3 \cdot s^{-1}$ $Q_{CES} = 183 \cdot 11\ m^3 \cdot s^{-1}$ (McGahey, 2006).

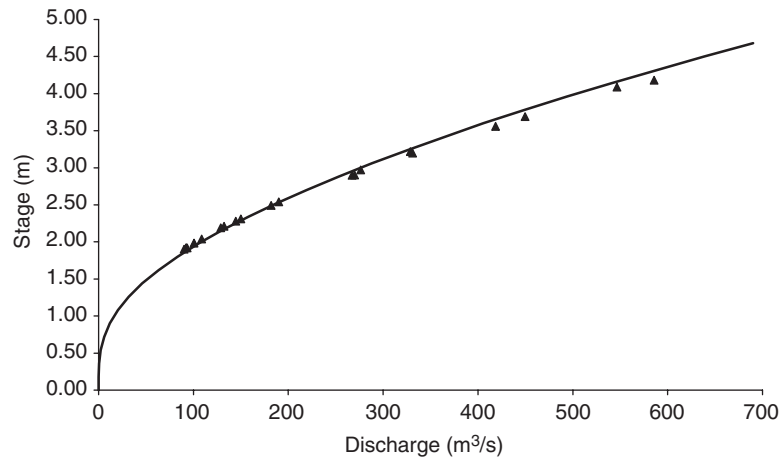


Figure 31. Stage-discharge prediction and data for Colorado River, Argentina (McGahey, 2006).

involved in estimating the conveyance capacity of rivers and drainage channels under flood flow conditions (see www.river-conveyance.net). It was developed by HR Wallingford and contains ancillary software dealing with roughness (particularly that related to vegetation), bridge afflux and uncertainty. Fig. 30 illustrates one prediction for the velocity distribution in the Colorado River, Argentina for a depth of 2.50 m. When laterally integrated, with the appropriate areas, the total discharge is within 0.6% of the measured value of $181.98\ m^3 s^{-1}$.

Using the calibration rules for the 3 parameters over a range of depths from 1.8 m to 4.0 m yielded the discharges shown in Fig. 31. The agreement between the measured and predicted stage-discharge curve is seen to be quite good.

A similar exercise was conducted for all the rivers, using a few measured distributions of U_d versus y to determine the calibration parameters in each panel. This was then followed by numerical predictions of the complete HvQ relationship for each river. The results of these predictions are shown in Fig. 32. The high R^2 value indicates the efficacy of the model used to obtain these results. Further work is in progress, exploring the physical basis of the particular parameters used in each river, and extending the range of discharges up to 25,000 $m^3 s^{-1}$.

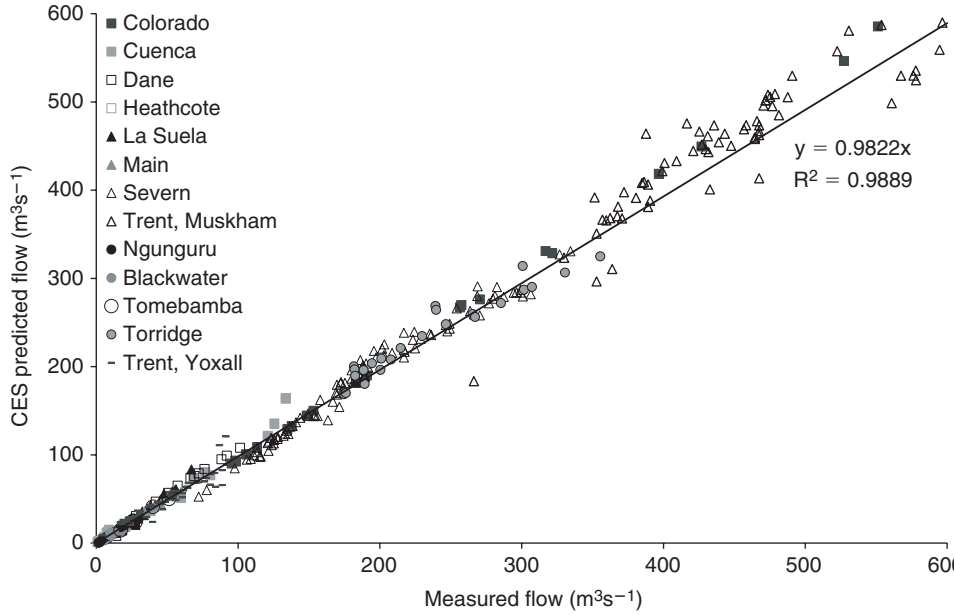


Figure 32. Discharge predictions using the CES (after McGahey et al., 2006).

7 CONCLUDING COMMENTS

The various stages in the development of a relatively simple 1-D river model have been described. The importance of data in grounding the model in sound science has been shown to have benefits in extending the 1-D approach. It has allowed a number of key 3-D flow processes that are commonly found in overbank flows in straight prismatic channels to be included. The data also led to rules for estimating values for the 3 key calibration parameters f, λ and Γ.

It has been demonstrated that good quality laboratory data have been useful in developing a model that is applicable at full scale. The model appears to provide realistic predictions of the variation of depth-averaged velocity across many rivers, and to be capable of providing reasonable estimates of the stage-discharge relationships for discharges up to $600\,m^3s^{-1}$.

Although the model was originally based on the Reynolds Averaged Navier-Stokes (RANS) equations, it contains one process that is normally excluded from such a starting point – i.e. the formation of large coherent unsteady planform vortices that often form at the edge of a floodplain. Since planform vorticity is *a priori* excluded from all RANS equation models, but is known to be present in most overbank flows at certain depths, it behoves those using RANS based models to be aware of this limitation (see Tominaga & Knight, 2004).

As with most models, empirical coefficients have to be selected for calibration purposes. These are often based on flows other than in rivers, and in some cases on flows not amenable to experimental scrutiny. See Nezu & Nakagawa (1993) and Ikeda & McEwan (2007) for further details of turbulence phenomena in open channels. In addition to the turbulence, various coherent structures may also be present, as shown by Ashworth et al. (1996) and Willetts & Rameshwaran (1996). The use of resistance coefficients to account for roughness effects is itself problematic, as indicated by Yen (1991) and Morvan et al. (2007). The engineer is thus faced with some considerable difficulty in blending theoretical and empirical knowledge together in a pragmatic manner in order to simulate river processes at all (Knight, 1996; Wang 2005).

Given the nature of flooding world-wide and the damage they cause (see Knight et al. 2006; Knight and Samuels, 2007) there is an increasing need to engage in fundamental R & D on flood

related issues. The studies described herein indicate the stages in developing a model for use by practicing river engineers, from investment in experimental facilities, to identifying key concepts, to formulation and eventual inclusion of these into some software, with associated tools.

REFERENCES

Abril, J.B. & Knight, D.W. 2004. Stage-discharge prediction for rivers in flood applying a depth-averaged model, *Journal of Hydraulic Research,* IAHR, 42(6): 616–629.

Ashworth, P.J., Bennett, S.J., Best, J.L. & McLelland, S.J. 1996. *Coherent flow structures in open channels*. 1–733. J Wiley.

Atabay, S. & Knight, D.W. 2006. 1-D modelling of conveyance, boundary shear and sediment transport in overbank flow, *Journal of Hydraulic Research,* IAHR, 44(6): 739–754.

Bousmar, D. & Zech, Y. 2002. Periodical turbulent structures in compound channels, *River Flow 2002*, [Eds D. Bousmar, D. and Y. Zech], Vol. 1: 177–185. Balkema.

Darby, S., et al. 1998. River width adjustment. Part II : Modelling, *Journal of Hydraulic Engineering,* ASCE, Task Committee on Hydraulics, Bank Mechanics and the Modelling of River Width Adjustment, Sept., 124(9): 903–917.

Fukuoka, S., Miyazaki, H, Ohgushi, H., & Kamura, D., 1996, Flow and bed topography in a meandering compound channel with phase difference between the alignment of the main channel and the levee, *Annual Journal of Hydraulic Engineering*, JSCE, 40: 941–946.

Ikeda, S. & McEwan, I.K. 2007. Flow and sediment transport in compound channels, *IAHR Monograph*, (in press).

Jakirlic, S. & Hanjalic, K. 2002. A new approach to modeling near-wall turbulence energy and stress dissipation, *J Fluid Mechanics,* 249: 139–166.

Knight, D.W. 1996. Issues and directions in river mechanics – closure of sessions 2, 3 and 5. In *'Issues and directions in hydraulics',* [Ed T Nakato & R Ettema], An Iowa Hydraulics Colloquium in honour of Professor John F Kennedy, Iowa Institute of Hydraulic Research, Iowa, USA: 435–462. Balkema.

Knight, D.W. 2005. River flood hydraulics: theoretical issues and stage-discharge relationships. In *River Basin Modelling for Flood Risk Mitigation* [Eds DW Knight & A Y Shamseldin], Chapter 17: 301–334. Taylor & Francis.

Knight, D.W. & Abril, B. 1996. Refined calibration of a depth-averaged model for turbulent flow in a compound channel, *Proc. Instn. of Civil Engineers, Water, Maritime and Energy Division*, London, 118(3), Sept..: 151–159.

Knight, D.W. & Samuels, P.G. 2007. Examples of recent floods in Europe, *Journal of Disaster Research*, Fuji Technology Press, Tokyo, Japan, 2(3): in press.

Knight, D.W. & Shiono, K. 1996. River channel and floodplain hydraulics. In *Floodplain Processes*, (Eds Anderson, Walling & Bates), Chapter 5: 139–181. J Wiley.

Knight, D.W., Al-Hamid, A.A.I. & Yuen, K.W.H. 1992. Boundary shear in differentially roughened trapezoidal channels. In *Hydraulic and Environmental Modelling:Estuarine and River Waters* (Eds R.A. Falconer, K. Shiono & R.G.S. Matthew): 3–14. Ashgate Press.

Knight, D.W., Omran, M. & Abril, J.B. 2004. Boundary conditions between panels in depth-averaged flow models revisited. In *River Flow 2004, Proc. 2nd Int. Conf. on Fluvial Hydraulics, 23–25 June, Napoli, Italy* [Eds M. Greco, A. Carravetta & R.D. Morte], Vol. 1: 371–380. Balkema.

Knight, D.W., Omran, M. & Tang, X. 2007. Modelling depth-averaged velocity and boundary shear in trapezoidal channels with secondary flows, *Journal of Hydraulic Engineering*, February, ASCE, 133(1): 39–47.

Knight, D.W., Tang, X. & McGahey, C. 2007. A method for analysing stage-discharge, velocity and boundary shear stress data for rivers in flood, *Proc. 5th Int. Conf. on Urban Watershed Management and Mountain River Protection and Development*, Chengdu, China, April 3–5: 1–15.

Knight, D.W., Yuen, K.W.H. & Alhamid, A.A.I. 1994. Boundary shear stress distributions in open channel flow. In *Physical Mechanisms of Mixing and Transport in the Environment*, (Eds K. Beven, P. Chatwin & J. Millbank), Chapter 4: 51–87. J. Wiley.

Knight, D.W., Cao, S., Liao, H., Samuels, P.G., Wright, N.G., Liu, X. and Tominaga, A. 2006. Floods – are we prepared? *Journal of Disaster Research*, Fuji Technology Press, Tokyo, Japan, 1(2): 325–333.

Lai, C.J. & Knight, D.W. 1988. Distributions of streamwise velocity and boundary shear stress in compound ducts, *Proc. 3rd Int. Symp. on Refined Flow Modelling and Turbulence Measurements*, Tokyo, Japan, July: 527–536.

McGahey, C. 2006. A practical approach to estimating the flow capacity of rivers. PhD thesis, *Open University*, Milton Keynes, UK.

McGahey, C., Samuels, P.G. & Knight, D.W. 2006. A practical approach to estimating the flow capacity of rivers – application and analysis, In *RiverFlow 2006*, September, Lisbon, Portugal, Vol.1: 303–312. Taylor & Francis.

Morvan, H., Knight, D.W., Wright, N.G., Tang, X. & Crossley, M. 2007. The concept of roughness in fluvial hydraulics and its formulation in 1-D, 2-D & 3-D numerical simulation models, *Journal of Hydraulic Research*, IAHR, (in press).

Nezu, I. & Nakagawa, H. 1993. Turbulence in Open-Channel Flows, *IAHR Monograph*, Balkema, Rotterdam.

Omran, M. & Knight, D.W. 2006 Modelling the distribution of boundary shear stress in open channel flow, In *RiverFlow 2006*, September, Lisbon, Portugal, Vol. 1, [Eds Ferreira, Alves, Leal & Cardoso], : 397–404. Taylor & Francis.

Prooijen, B.M. 2004. Shallow mixing layers, *PhD Thesis*, Delft University of Technology Report 04-1, ISSN 0169–6548.

Rhodes, D.G. & Knight, D.W. 1994a. Distribution of shear force on boundary of smooth rectangular duct, *Journal of Hydraulic Engineering*, ASCE, 120(7): 787–807.

Rhodes, D.G. & Knight, D.W. 1994b. Velocity and boundary shear in a wide compound duct, *Journal of Hydraulic Research*, IAHR, 32(5), October: 743–764.

Rhodes, D.G. & Knight, D.W. 1995. Lateral shear in a wide compound duct, *Journal of Hydraulic Engineering*, ASCE, 121(11): 829–832.

Shiono, K. & Knight, D.W. 1991. Turbulent open channel flows with variable depth across the channel, *Journal of Fluid Mechanics*, 222: 617–646 (and Vol. 231, Oct., p 693).

Shiono, K. & Muto, Y. 1998. Complex flow mechanisms in compound meandering channels with overbank flow, *Journal of Fluid Mechanics*, Vol. 376: 221–261.

Thomas, T.G. & Williams, J.J.R. 1995. Large eddy simulation of turbulent flow in an asymmetric compound open channel, *Journal of Hydraulic Research,* 33(1): 27–41.

Thorne, C., et al.1998. River width adjustment : Processes and mechanisms, *Journal of Hydraulic Engineering,* ASCE, Task Committee on Hydraulics, Bank Mechanics, and Modelling of River Width Adjustment, 124(9): 881–902.

Tominaga, A. & Knight, D.W. 2004. Numerical evaluation of secondary flow effects on lateral momentum transfer in overbank flows. In *River Flow 2004, Proc. 2nd Int. Conf. on Fluvial Hydraulics, 23–25 June, Napoli, Italy* [Eds Greco, Carravetta &Morte], Vol. 1: 353–361. Balkema.

Tominaga, A., Nezu, I., Ezaki, K. & Nakagawa, H. 1989. Turbulent structure in straight open channel flows. *Journal of Hydraulic Research*, IAHR, 27(1): 149–173.

Wang, S. 2005. On the reliability and accuracy of computational models for Hydroscience and Engineering investigations, *US-China Workshop on Advanced Computational Modelling in Hydroscience and Engineering*, Oxford, Mississippi, USA, 19, September.

Willetts, B.B. & Rameshwaran, P. 1996. Meandering overbank flow structures, In *Coherent Flow Structures in Open Channels* (Eds Ashworth, Bennett, Best & McLelland), Chapter 29, J Wiley: 609–629.

Yen, B.C. 1991. Channel flow resistance: centennial of Manning's formula, Water Resources Publications, Colorado, USA: 1–453. Water Resources Publications.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Optimal design and management of water distribution systems

D. Savic, Z. Kapelan & R. Farmani
Centre for Water Systems, University of Exeter, Exeter, Devon, UK

O. Giustolisi
Technical University of Bari, Taranto, Italy

ABSTRACT: Water distribution systems are designed to deliver water from a source in the required quantity, quality and at satisfactory pressure to individual customers. These systems must be designed and operated so as to: minimise energy use; limit water losses (leakage and not-accounted-for water); minimise the number and consequences of pipe failures; make most effective use of the existing assets; and still meet customers' needs in terms of water quality and quantity. Optimization of such systems has been receiving quite some attention in the literature and a large number of mathematical techniques have been developed and applied to water distribution system optimization problems. However, despite early claims of optimization, there has been limited evidence of wide acceptance of water distribution optimization tools. By presenting three case studies in which optimization was used in a multi-objective context (trading off costs against benefits), this papers goes some way to demonstrate the advances in optimal design and management of water distribution systems achieved in the last decade or so. The first case study shows how optimal design can include pipe decisions, tank sizing, tank sitting, and pump operation schedules, by incorporating both capital and operational costs and considering benefits (reliability) as a second objective. The second study shows why risk optimization is preferred to reliability optimization for water distribution systems. Finally, in the third study the benefits of burst reduction achieved by mains rehabilitation were evaluated by a multi-objective optimization model over a short-term planning horizon (taken to be one year in this study).

1 INTRODUCTION

The main function of an urban water system is to supply sufficient quantity of potable water and to conduct, treat and dispose of wastewater. The importance of safe water supply and disposal of wastewater is so great that a recent poll carried out by the British Medical Journal concluded that sanitation is the greatest medical milestone of the last century and a half (Guardian 2007). Life magazine goes even further by saying that drinking water filtration and the use of chlorine is "probably the most significant public health advance" of the last millennium (Life Magazine 1997). Design of water and wastewater systems has, therefore, a major impact on wellbeing of people in modern conurbations. However, drinking water and wastewater systems in many countries face increasing challenges over the next several decades in maintaining and replacing their pipes, treatment plants, and other infrastructure. The cost of operating an urban water system may be substantial (due to maintenance, repair, water and wastewater treatment, energy costs, etc.), but still one of the main costs is that of the assets (pipelines, treatment works, pumping stations, etc.) themselves. For example, the total book value of all water distribution and wastewater pipes in the UK (over 750,000 km) is estimated at over £200 billion.

A drinking water distribution system is part of an urban water system and is usually referred to as "clean water system" as opposed to "dirty water systems" associated with drainage and wastewater

networks. Water distribution systems are designed to deliver water from a source in the required quantity, quality and at satisfactory pressure to individual customers. It typically consists of a network of pipes, pumps, valves and storage reservoirs with external supplies from boreholes, rivers and lakes (usually via a treatment plant). These systems must be designed and operated so as to: minimise energy use; limit water losses (leakage and not-accounted-for water); minimise the number and consequences of pipe failures; make most effective use of the existing assets; and still meet customers' needs in terms of water quality and quantity. The above list of objectives is by far not exhaustive, but it indicates that the task of managing water distribution systems is extremely complex.

This paper introduces some modelling and optimization ideas, which have redefined the way engineers and scientists are designing and managing water distribution systems.

2 TO OPTIMIZE OR NOT?

With the advent of personal computers and computer simulation packages capable of modelling behaviour of water distributions systems, it has now become possible to simulate more complex network components (e.g. pumps, tanks, PRVs, etc) of the water distribution system as well as to investigate more complex issues associated with their design and operation (Ormsbee 2006). Such additional applications have included simulation (extend period simulations, transient analysis, water quality analysis, etc.) and system optimization (optimal model calibration and design, optimal operations, optimal sampling plans, etc.). Although optimization can be used for many different problems related to water distribution systems, the following discussion will focus on water distribution system design.

The word "optimal", according to several on-line dictionaries, means "most favourable/desirable/satisfactory" or "best". Mathematical techniques have been developed to determine optimal solutions for a wide variety of mathematical problems. Among many science and engineering problems, water distribution system optimization has been receiving quite some attention (Lansey 2006) and a large number of mathematical techniques have been developed and applied to system optimization problems listed by Ormsbee (2006). It is only logical to conclude that using computer systems to assimilate information and provide advice about water distribution system problems will be the most encouraging opportunity in the future. Goulter (1992) identifies the need for the development of such information systems in the context of design and uses the name "decision support systems" (DSS). The term "support" is important here because both practitioners and researchers have contributed to misunderstanding of what role optimization methodologies and DSS could play in water distribution system optimization. The main issue here is of decision making vs. decision support, or whether the intention of these methodologies is to provide an "optimal" solution to a problem or to explore the solution space and provide a number of good-quality solutions for the decision maker to choose from. Since any "optimal" solution obtained through the modelling process is clearly dependent on the assumptions embedded in the model (be it simulation and/or optimization), the "optimal" solution refers to the best solution to the mathematical model. This could be quite different from the actual optimal solution to the problem being analysed due to modelling assumptions and simplification. Matching the model to the real problem is a complex and difficult task and the authors believe is one of the main reasons why optimization has not enjoyed its anticipated success in water distribution system design.

Goulter (1992) identifies the following four reasons for the lack of wide acceptance of optimization: (1) The models do not work, i.e., they cannot derive solutions for practical problems or they do not give sensible solutions; (2) the solutions they provide offer no improvements over established methods and professional judgement; (3) they are too difficult to use; or (4) practitioners are not comfortable with the overall optimization approach. In addition to these, Goulter (1992) identifies reliability, or the lack of agreement to what it represents, as one of the stumbling blocks in optimal design of water distribution systems. The term reliability, in water distribution systems, mainly refers to the ability of the network to provide consumers with an adequate and high quality supply,

under normal and abnormal conditions. The reliability of water systems can be studied considering two types of failure, mechanical failure and hydraulic failure (Farmani et al. 2005). Mechanical failure usually refers to failures of system components, such as pipe breakage or pump being out of service. Mechanical failure has been studied widely in the past and large number of models have been developed that account for this type of reliability. However, the question still remains that what is the best measure of this type of reliability and what level of reliability is acceptable. Hydraulic failure, on the other hand, refers to uncertainties, such as consumption (nodal demand), pipe roughness and reservoir and tank levels. Using good engineering judgement, one can estimate demand peaks under existing conditions with an error of approximately 4–6% (Obradovic & Lonsdale 1998). While this error does not appear to be too excessive, the situation becomes much worse when one needs to predict the consumption for a time horizon of 10 or 20 years into the future. However, due to the complexity and diversity of the reliability definitions, no single definition takes the reliability of water distribution networks into account.

Walski (1995) identifies other issues that have contributed to the lack of wide acceptance of "pipe-sizing optimization" in engineering practice, including the fact that there is not one single pipe-sizing problem, but rather several significantly different problems. Those issues, according to Walski, include long-term master planning, preliminary design, subdivision layout (or zoning) and rehabilitation, which are also related to problems of tank sizing and location, tank-overflow level, pressure zone boundary, pump selection, pump operation and control-valve setting. In his later paper Walski (2001) also criticises cost minimisation as a goal for water distribution system design. He names the following reasons for this: (1) It is difficult to identify true benefits and constraints due to uncertainty in future demands; (2) there will be many alternatives with virtually the same net benefits; and (3) actual demands tend to be controlled (to a certain extent) by pipe sizes – "if you build it (pipes of certain sizes), they will come (customers, hence demand). Walski (2001) concludes that "the state-of-the-art" in water distribution system design is stuck in 'local optima' in the researchers have pushed cost minimization as far as it will go".

There is, however, a plethora of evidence that researchers and practitioners have moved on from the time Goulter (1992) wrote his state-of-the-art review of optimization for water distribution system design. The introduction of Genetic Algorithms (GA) to water distribution system optimization in the early nineties of the last century has provided the ability of optimizing large and complex systems. The uptake of the technology by the software industry (e.g., MWH Soft, KYPIPE, Haestad Methods, Wallingford Software, etc) has brought the idea of DSS closer to reality. The following examples demonstrate some new developments and trends in water distribution system optimization.

3 OPTIMAL WATER DISTRIBUTION SYSTEM DESIGN

Water distribution systems are inherently complex because they exhibit many attributes of complex systems. They are large-scale and spatially extensive; they are laid out with multiple pipe loops to maintain satisfactory levels of redundancy for system reliability; their hydraulic properties, such as aqueous flows and pressures are governed by nonlinear hydraulic equations; their water quality properties are governed by advection, convection and diffusion; they contain complex hydraulic devices such as various valves and pumps; they are impacted by pumping and energy requirements; they are complicated by numerous layout, pipe sizing, and pumping alternatives and influenced by analysis of trade-offs between capital investment, operations and maintenance costs during the design process.

Water distribution system design optimisation is one of the most heavily researched areas in the hydraulics profession. However, water network design is a challenging optimization problem because it requires evaluation of network hydraulics, which can be highly non-linear due to governing hydraulic equations and pump characteristics, or exhibit discontinuous behaviour due to devices such as pressure-reducing valves. Complexity of the problem is also affected by water quality modelling, which is required to account for dynamic changes in the quality of the water

transported through the system. The design is also considered a combinatorial problem because decision variables are discrete, either as commercially available diameters or pumps and other control devices being switched on and off at discrete time intervals. Finally, uncertainty in demand loadings adds another complexity dimension to the problem of optimal water distribution system design. The following examples show how this complex problem can be tackled in three different circumstances using modern optimization tools.

3.1 *Design for total cost and reliability considerations (Farmani et al. 2005)*

The Anytown water distribution system, which was set up by Walski et al. (1987) as a realistic benchmark with which to compare and test network design optimization software, has features and problems typical of those found in many real systems. The Anytown problem was originally tackled by participants in the Battle of the Network Models workshop and has since been examined by Murphy et al. (1994) and Walters et al. (1999). All participants in the original workshop used optimization models to size the piping system while manually choosing the location and size of tanks. The difference between the various methods originally used lies essentially in the pipe optimization models, which were based on linear programming, partial enumeration, or nonlinear programming techniques. No attempt was made to optimize the provision of tanks or pumps, except by use of expert judgment and experience. Most of the participants were able to develop a solution that would work at peak loading but would not have adequate capacity to fill tanks at off-peak loadings (Walski et al. 1987).

Murphy et al. (1994) later obtained a better solution to the problem using a standard genetic algorithm, which was able to handle the tanks and pumps as additional design variables with network deficiencies incorporated into the objective function as penalties, using a set of weightings, one for each type of constraint violation. Walters et al. (1999) introduced the structured messy genetic algorithm for solving the Anytown water system. The problem was posed as a multi-objective optimization problem for the first time, with minimum cost and maximum benefit as dual objectives. The design variables were the addition of new pipes, reconditioning and replacement of existing pipes, sitting and sizing of new tanks, and provision of new pumps. The method produced designs cheaper than any previously published.

The multi-objective optimization of water distribution is investigated here through the analysis of the Anytown water supply system (Fig. 1). Pipes in the central city are thick solid lines, pipes represented by thin solid lines are in residential areas, and new pipes are represented by dashed lines. The objective of the Anytown network problem was to determine the most economically effective design to reinforce the existing system to meet projected demands, taking into account pumping costs as well as capital expenditures. The town is formed around an old centre situated to the southeast of pipe 34, where excavations are more difficult to undertake and consequently are more expensive. There is a surrounding residential area, with some existing industries near node 17 and a projected new industrial park to be developed to the north. Options include duplication (in a range of possible diameters) of any pipe in the system, addition of new pipes, selecting the operation schedule of pumping stations, and provision of new reservoir storage at any location. Water is pumped into the system from a water treatment works by means of three identical pumps connected in parallel. The link and node data, the average daily water use at each node as well as different loading conditions and the variation in water use throughout the day are available from CWS (2007).

The presence of pumps requires that both the design and operation of the network should be considered in the optimization. Inclusion of water storage tanks in the network is to provide equalisation and emergency storage. Where there is more than one tank, the problem of designing and operating the system becomes much more complicated. Numerous decisions must be made in the design of a storage tank, including size, location, type, and expected operation (Walski 2000). The Anytown problem has features and problems typical of those found in many real systems and thus is very complicated. As an example, tank design can cause problems of water stagnation in tanks oversized due to fire flows and emergency storage requirements. Sequences of demands describing daily operations also increase computational effort.

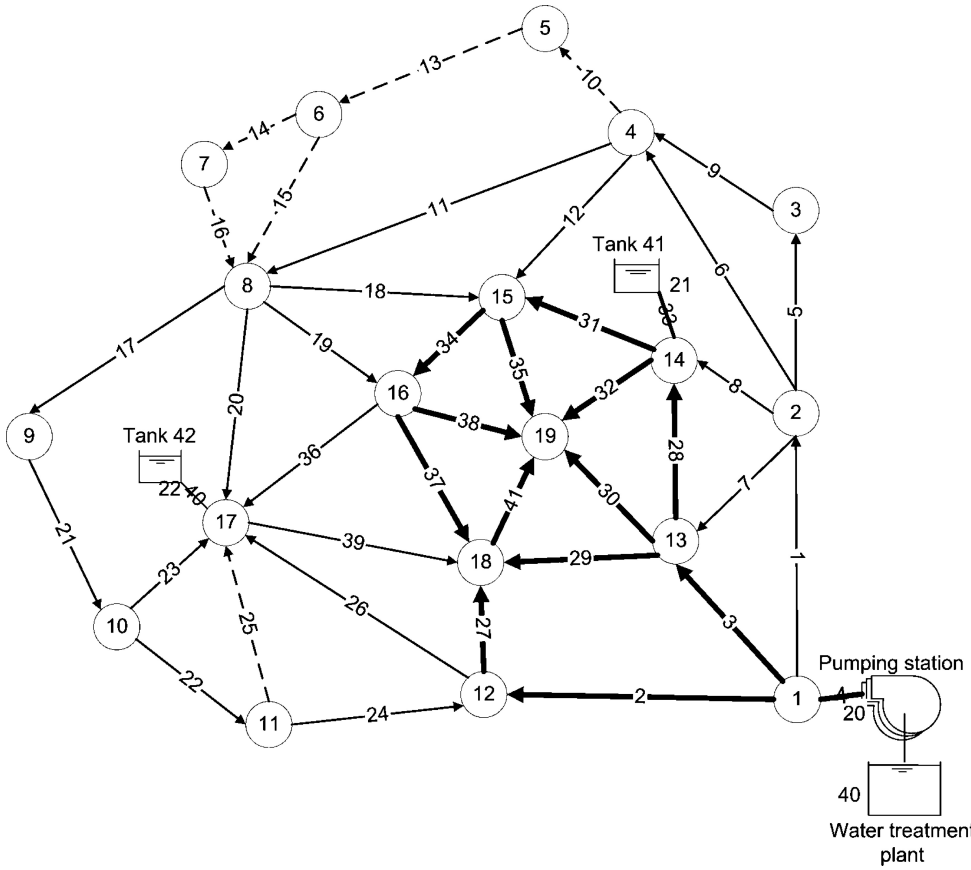


Figure 1. Anytown network (Farmani et al., 2005).

To achieve fully feasible designs, the following approach is considered to model the problem. The location, overflow elevation and minimum normal day elevation, diameter, and bottom of tank from minimum normal day elevation are treated as independent design variables for the tanks, thereby defining completely the design of the storage. However, storage will initially not match the requirements of the network, and penalties are necessary to produce feasible solutions. Considering tank characteristics as explicit design variables will result in fast convergence of the algorithm in feasible solutions, contrary to algorithms that consider only tank volume as a design variable. In the latter case, adjusting the operational levels and tank diameters is a challenging task as they have secondary effects on the optimization process.

3.1.1 *Decision variables*

Thirty-five existing pipes are considered for duplication or cleaning and lining (giving 35 pipe options and 35 diameter variables). There are 6 additional new pipes. All nodes are considered as potential sites for new tanks, except those that are already connected directly to the existing tanks. Therefore the design variables for tanks are 2 possible locations for the new tanks in the network, overflow and minimum normal day elevations, bottom of tank from minimum normal day elevation, and diameter of each tank (8 variables). Tanks are connected to a node by a short pipe, known as a riser, of known length (30.7848 m or 101 ft), but of variable diameter, giving a total of 2 diameter variables for the risers and a total of 78 design variables for links in the network. Given a 24 h operation cycle and 1 h time step, the control variables for the three pumps give a further 24 design variables and a total of 112 design variables for pipes, tank volumes, and pumps.

For pipe status the following integer values were used as "genes" (GA terminology): 0 for "leave," 1 for "clean and line," and 2 for "duplicate" the existing pipes. In the pipe and riser sizes, 0 to 10 correspond to the 10 available discrete pipe diameters and one extra possible decision, which is the "do nothing" option. The number of pumps in operation during each hour is represented by integer values 0, 1, 2, and 3 for the corresponding number of pumps. Every node in the network is a possible location for new tanks except those already connected to existing tanks and, therefore, there are 16 possibilities.

3.1.2 *Constraints*

Supplying demand to consumers at an adequate pressure is the main constraint in the design of water distribution systems. The presence of tanks and pumps in the network adds extra difficulty to the optimization as these are the two most difficult components to model. Tanks store water for daily operations to smooth the pumping demands throughout the day, making the system more cost effective, and also store water for emergency conditions. In this work, the tank characteristics are treated as independent design variables. In particular, the design top and bottom water levels for all tanks are specified for each trial solution. All tanks should empty and fill over their operational ranges during the specified average demand day, leaving the specified emergency volumes untouched. The actual maximum and minimum water levels are then identified for each tank during a 24 h simulation. There will in general be a mismatch between the top and bottom water levels specified and those resulting from the simulation, and also a mismatch between initial and final water levels. The accumulated sum of the mismatch in levels is used as the operating level difference (TLD) constraint.

The performance of each candidate design solution is evaluated through simulation of the network flows. An extended period hydraulic network solver (EPANET2, Rossman 2000) is used to determine the heads at all the system nodes for all design conditions, and the accumulated sum of the nodal pressure shortfalls (NPS) is used as the head deficit constraint. The design constraints are given in Table 1. The cost of the solution includes the capital costs of pipes, pumps, and tanks as well as the present value of the energy consumed during a specified period. The unit costs for pipe laying, cleaning, and lining are available from CWS (2007). Pump station operating costs are based on a unit cost for energy, constant throughout the 24 h, equal to \$0.12/kWh. The present worth of energy costs is based on an interest rate of 12% and an amortisation period of 20 years. Tank costs are considered a function of volume and are from CWS (2007). Intermediate tank sizes are considered and the corresponding costs are interpolated linearly from standard sizes and costs. This work sets the problem as a multi-objective optimization problem under multiple loading conditions. Although designing under multiple loading conditions will introduce redundancy to a network, to further increase the hydraulic reliability and the availability of water during pipe failure, maximization of the resilience index is used as an objective.

3.1.3 *Objectives*

The presence of pumps requires that both the design and operation of the network should be considered in the optimization. The total cost is the sum of pump operating costs, tank capital

Table 1. Summary of design costs (\$million), Minimum individual surplus head (m), and Reliability (Farmani et al., 2005).

Description	1	2	3	4	5	6	7	8
Pipes	6.489	6.617	6.623	7.968	9.283	10.009	10.394	11.682
Tanks	0.675	0.674	0.677	0.708	0.572	0.579	0.582	0.577
Pumps	6.184	6.22	6.222	6.154	6.18	6.222	6.17	6.217
Total	13.348	13.511	13.522	14.83	16.035	16.81	17.146	18.476
Minimum individual surplus head	1.0	0.9	1.0	3.2	2.3	0.8	0.3	1.3
Resilience index	0.182	0.183	0.184	0.198	0.2	0.203	0.208	0.214

costs, and pipe network capital cost. However, cost-optimization tends to reduce cost by reducing the diameter of or completely eliminating some pipes, thus leaving the system with insufficient capacity to respond to pipe breaks or demands that exceed design values without violating required performance levels. To improve the performance of a water distribution network under failure conditions, this study considered reliability as the second objective in the optimization problem. Therefore, two objectives, minimising total cost (operations and capital) and maximising reliability were considered.

Reliability can be included in the design of water distribution networks using two approaches. The first are surrogate-based measures (deterministic modelling) and the second involve stochastic analysis of uncertainty (probabilistic modelling). Deterministic modelling has been studied by a large number of researchers in the past (Tung 1985, Wagner et al. 1988, Goulter 1992; Tanyimboh & Templeman 1993). Tung (1985) defined network reliability as the probability that flow can reach all the demand nodes in the network. Wagner et al. (1988) defined similar reliability measures as the probability that a demand node is connected to a source (connectivity) or that all demand nodes are connected to a source (reachability). Goulter (1992) and Tanyimboh & Templeman (1993) implemented the concept of entropy the expected shortfall in case of a pipe breakdown. Probabilistic modelling has been considered by some researchers (Xu & Goulter 1999, Tolson et al. 2004; Babayan et al. 2005, 2006, 2007, Kapelan, et al. 2004, 2005, 2006) in recent years for reliability analysis of water distribution systems.

This study considers a surrogate-based, multi-objective optimization method to account for network reliability. First some definitions of surrogate measures for reliability of water distribution networks are given. The amount by which the minimum available head ($h_{\text{ava},i}$) exceeds the minimum required head ($h_{\text{req},i}$) at the critical node in the system over all loading conditions can be considered an indicator of benefits or minimum surplus head benefit (I_s) as follows:

$$I_s = \min\left(h_{\text{ava},i} - h_{\text{req},i}\right) \quad i = 1, ..., N_{\text{node}}$$

The benefit can also be measured in terms of flow provided through a system. The concept of resilience was introduced by Todini (2000) to account for the fact that water distribution networks are designed as looped systems to increase hydraulic reliability and availability of water during pipe failure. Water distribution networks are usually designed to deliver water to each node, satisfying the demand in terms of design flow and head. However, in the case of change in demand or a pipe failure, the water flow will change and the original network will be transformed into a new one with higher internal energy losses. This might make it impossible to deliver the desired flow rate at a minimum delivery pressure. Providing more power (energy per unit time) than that required at each node could be one of the possible ways to avoid this problem. This will add sufficient surplus to be dissipated internally in case of failure. This surplus has been used by Todini (2000) to characterize the resilience of looped networks; although it does not involve statistical consideration of failures, its increase will lead to improved network reliability.

The power dissipated in the pipes (P_{int}) plus the power delivered to the user in terms of flow (q_{i}) and head ($h_{\text{ava,i}}$) at each node (P_{ext}) is equal to the total available power.

$$P_{\text{tot}} = P_{\text{int}} + P_{\text{ext}}$$

$$P_{\text{ext}} = \gamma \sum_{i=1}^{N_{\text{node}}} q_i h_{\text{ava},i}$$

where, γ is the specific weight of water.

The total available power includes the power at the entrance to the water distribution network as well as the power introduced into the network by the pumps (P_{k}).

$$P_{\text{tot}} = \gamma \sum_{j=1}^{N_{\text{reservoir}}} Q_j H_j + \sum_{k=1}^{N_{\text{pump}}} P_k$$

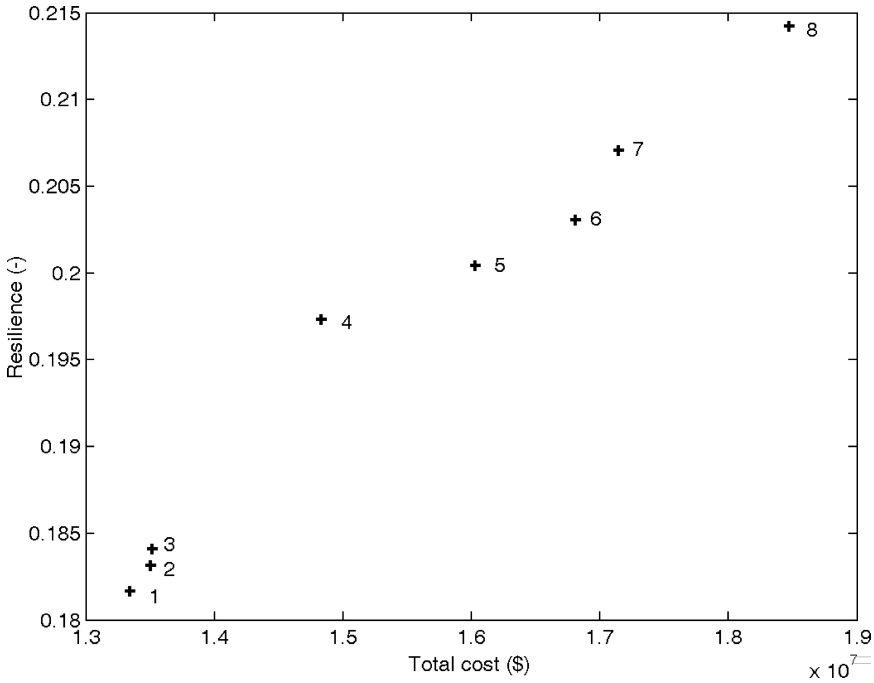


Figure 2. Total cost vs. resilience payoff (Farmani et al., 2005).

where Q_j and H_j are the discharge and the head, respectively, at each reservoir.

The resilience index of a looped network is defined as

$$I_r = 1 - \left(\frac{P_{\text{int}}}{P_{\text{max, int}}} \right)$$

where P_{int} is the amount of power actually dissipated in the network while satisfying the total demand

$$P_{\text{int}} = P_{\text{tot}} - \gamma \sum_{i=1}^{N_{\text{node}}} q_i h_{\text{ava},i}$$

and $P_{\text{max,int}}$ is the maximum power that could be dissipated internally while satisfying the total demand.

$$I_r = \frac{\sum_{i=1}^{N_{\text{node}}} q_i \left(h_{\text{ava},i} - h_{\text{req},i} \right)}{\left(\sum_{j=1}^{N_{\text{reservoirs}}} Q_j H_j + \sum_{k=1}^{N_{\text{pumps}}} \frac{P_k}{\gamma} \right) - \sum_{i=1}^{N_{\text{node}}} q_i h_{\text{req},i}}$$

Thus, the reliability objective was expressed as:

Minimize I_r

3.1.4 *Solutions*

The twin objectives of maximizing resilience and minimizing cost under multiple loading conditions leads to a set of non-dominated (Pareto) solutions. Figure 2 gives the payoff characteristic between the total cost and the resilience index as a surrogate measure for reliability for the Anytown network obtained by using the NSGAII algorithm (Deb et al. 2000). In this method a fast, non-dominated

sorting approach with a selection operator that creates a mating pool by combining the parent and offspring populations and selecting the best solutions (with respect to fitness and spread) is presented. The next generation is populated starting with the best non-dominated front, progressing through the rest until the population size is reached, and, if in the final stage there are more individuals in the non-dominated front than available space, the crowded distance-based niching strategy is used to choose which individuals of that front are entered into the next population (Deb et al. 2000).

The limited number of solutions on the payoff curve (Fig. 2) shows that the problem is highly constrained, and the distances between the solutions reflect the discrete nature of the optimization problem. To determine the payoff characteristic between the total cost and resilience index, the multi-objective algorithm was run with a population size of 100 sample solutions and allowed to run for 5,000 generations.

Details of the design costs, minimum individual surplus head, and resilience index for all the solutions on the Pareto front are given in Table 1. It is evident that the main differences in the solution costs are from the pipe costs. This illustrates that to provide higher reliability in the network, the network capacity has been increased to respond to pipe breaks or demands that exceed design values. Increase in the resilience index did not have a huge effect on pump operation schedules and tank sizing or sitting. Pressure requirements were checked and found to be satisfied for all the specified design loadings for all the solutions on the Pareto set every minute over the 24 h design period.

To assess the quality of the solutions on the Pareto front in response to pipe failure, for each fixed configuration the possibility of single pipe failure for each pipe is considered in turn. The effect of each pipe failure on the network is studied by closing the failed pipe and running a 24 h system simulation, with the time at which the first pressure deficiency occurs being recorded. Deficiencies occurring after 24 h are ignored, as this is considered sufficient time to effect a repair. The number of pipes in the network that, when individually closed, will cause pressure deficiencies to occur under normal demands within 24 h is calculated for each solution on the Pareto front. The cumulative number of pipe closures that cause failure during the 24 h time horizon is also calculated. From these results, it is difficult to identify a clear trend, with solutions 6 and 7 having a high number of failures despite relatively high resilience indices and costs. Detailed inspection of the network solutions on the Pareto front shows that although each demand node is connected to at least two links, there is not enough surplus head at each demand node to satisfy demands when pipes fail.

It is clear that the resilience index alone does not give an adequate representation of network robustness when pipe failures are considered. Hence a third objective, maximizing the minimum surplus head at any demand node, is introduced and a three-objective optimization is run. Figure 3 gives the payoff characteristic between the total cost and the resilience index for the Anytown network, extracted from the Pareto surface generated by the three-objective optimization.

Details of the design costs, minimum individual surplus head, and resilience index for all the solutions on the Pareto front are given in Table 2. The solutions on the Pareto front are all fully feasible and satisfy all design and operational criteria in terms of tank sizing and sitting and pump operational schedules. Again the main difference in the solution costs comes from the pipe costs.

3.2 *Design for cost and risk considerations (Kapelan et al. 2006)*

While including reliability in optimal design of water distribution systems has proven its usefulness it is also important to know the consequences of failures and not just the failure likelihoods (probabilities). Risk is usually defined as a measure of the probability and severity of adverse effects. Unreliability, on the other hand, is only a measure of the probability that the system does not meet its intended functions. In other words, unreliability does not include the consequences of failures, whereas risk is inclusive and thus more representative. Clearly, not all failures can justifiably be prevented at all costs. Thus, system reliability cannot constitute a viable metric for resource allocation, unless an a priori level of reliability has been determined (Haimes 2001). Therefore, risk should be a consideration when optimizing for water distribution system design.

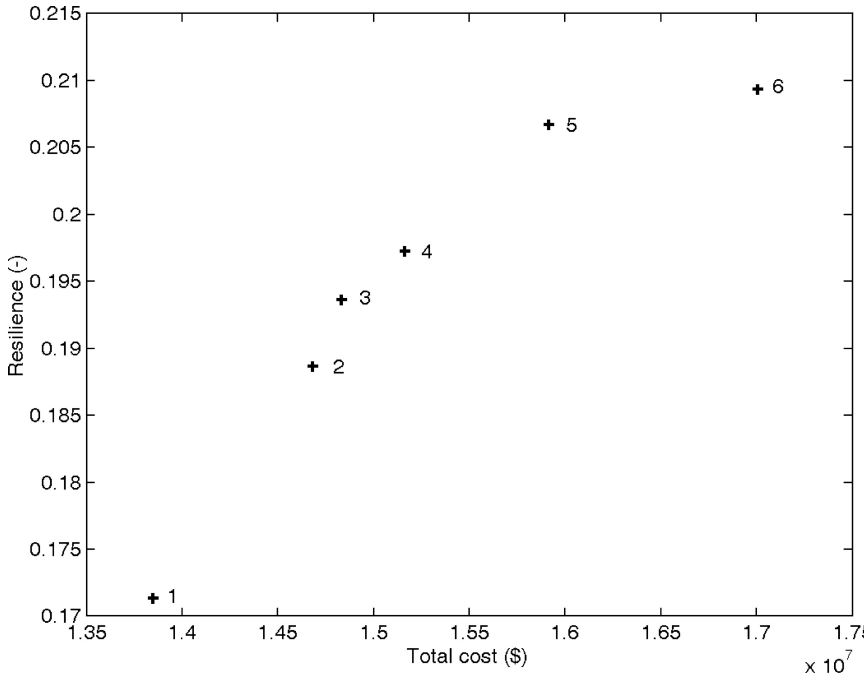


Figure 3. Cost versus resilience payoff generated by three-objective optimization (Farmani et al., 2005).

Table 2. Summary of design costs ($million), Minimum individual surplus head (m), and Reliability (Three-objective optimisation).

Description	1	2	3	4	5	6
Pipes	7.273	7.885	8.219	8.515	9.400	10.418
Tanks	0.553	0.575	0.542	0.545	0.546	0.549
Pumps	6.025	6.224	6.071	6.106	5.972	6.041
Total	13.851	16.684	14.832	15.166	15.918	17.008
Minimum individual surplus head	1.9	2.3	3.1	3.2	2.3	2.7
Resilience index	0.171	0.189	0.194	0.197	0.207	0.209

Two alternative, multi-objective, problem formulations are considered here when solving the optimal water distribution system design under demand uncertainty: (1) reliability-based approach and (2) risk-based approach. In the reliability approach, the optimal water distribution system rehabilitation problem is formulated using the following two objectives: (1) minimise total design cost; and (2) maximise the system reliability, where reliability in this study refers to hydraulic failure, that is uncertainties due to normal fluctuations of consumption (nodal demand):

The New York tunnels problem (Schaake & Lai 1969) was used in this case study (see Figure 4). The original objective of the problem was to identify the least cost network rehabilitation solution which satisfies minimum head requirements at all nodes for a set of fixed nodal demands. The only means of rehabilitation allowed is to duplicate existing pipes with new ones. A total of 16 solutions are possible for each pipe in the network: do not duplicate that pipe or duplicate the pipe with a new pipe having one of 15 available diameters (range 91–518 cm). Therefore, even though the New York tunnels network is fairly small, the single-objective optimisation problem is quite large (total number of possible solutions, i.e. network configurations is $16^{21} = 1.9 \cdot 10^{24}$).

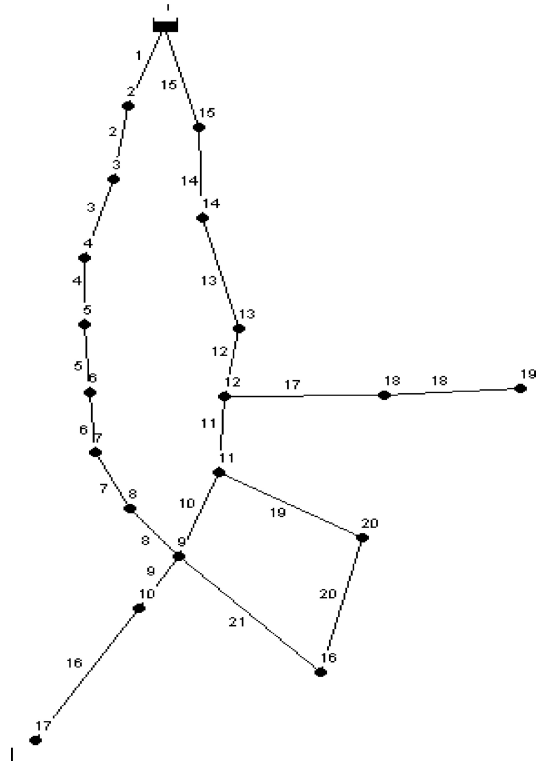


Figure 4. New York Tunnels Strengthening Problem (Kapelan et al., 2006).

A number of authors have solved the problem as a deterministic, single-objective problem. A review of these approaches can be found in Savic & Walters (1997). The objective of the analysis shown here is to identify the Pareto optimal front for the rehabilitation of the New York tunnels by solving two alternative multi-objective problems, one based on maximisation of reliability and the other one based on risk minimisation. This is achieved using the following two models.

3.2.1 *Reliability model*

In this model, the optimal water distribution system rehabilitation problem is formulated using the following two objectives: (1) minimise total design cost and (2) maximise the system reliability:

$$\textit{Minimise } C\left(D_1, D_2, ..., D_{N_d}\right)$$

$$\textit{Maximise } \textit{Rel}\left(D_1, D_2, ..., D_{N_d}\right)$$

where C – total capital cost; Rel – system reliability defined as probability $P(h_{\text{ava},i} \geq h_{\text{req},i}$, $i = 1, \ldots, N_{\text{nodes}})$ that heads (h_i) at all networks nodes $(i = 1, \ldots, N_{\text{nodes}})$ are *simultaneously* equal to or above the corresponding minimum head requirement at each node $(h_{\text{req},i})$; D_i – value of the i-th discrete decision variable (in general, the design option index); and N_d – number of decision variables. Additional constraints deal with continuity and headloss equations (implicitly satisfied by the EPANET solver) and discrete nature of decision variables (discrete set of commercially available diameters).

3.2.2 *Risk model*

Problem formulation in this approach is similar to the model above. The only difference is that instead of the system reliability maximisation in equation, total system risk (of water not being

delivered at acceptable pressure) is minimised. Therefore, the second objective is replaced with the following

$$Minimise\ Risk = \sum_{i=1}^{N_{\text{nodes}}} P_i\left(h_i < h_{\text{req},i}\right) \cdot CF_i$$

where: *Risk* – total system risk equal to sum of all nodal risks of water not being delivered to the customers at the acceptable pressure; $P_i(h_i < h_{\text{req},i})$ – probability of pressure head failure at i-th node equal to $1 - P_i(h_i > h_{\text{req},i})$, where $P_i(h_i > h_{\text{req},i})$ denotes the probability of success, i.e. head robustness at the i-th node; CF_i – consequence of pressure failure at i-th node, surrogated here as a fraction of the total mean system demand:

$$CF_i = \frac{q_i^{\text{avg}}}{\sum_{i=1}^{N_{\text{nodes}}} q_i^{\text{avg}}}$$

where q_i^{avg} is the mean demand at node i. Note that consequence of failure can be expressed in several other ways, e.g. as a number of customers affected by the pressure failure or the financial loss experienced, etc.

In the two approaches presented here it is assumed that nodal demands are the only source of uncertainty, i.e. it is assumed that all other hydraulic model inputs are deterministic variables. In addition to this, uncertain nodal demands are assumed to be independent (i.e. uncorrelated) random variables following some pre-specified probability density function (PDF). These assumptions were, again, introduced for simplicity reasons. The methodology developed and presented here can, actually, handle uncertainty in any type of model input variable/parameter (e.g. pipe roughness coefficients). Also, there is no particular reason why the methodology shown here cannot be used to handle correlated random variables.

3.2.3 *Solutions*

In the methodology presented here, a sampling technique is used to quantify the relevant uncertainties. Each time two objectives are calculated, a total of N_s sets (i.e. samples) of random nodal demands are generated. The corresponding nodal heads are then obtained by running the hydraulic solver for each of the demand samples. The reliability is estimated as the fraction (i.e. percentage) of total number of samples N_s for which the minimum head requirement condition is met simultaneously at all network nodes. Similarly, total risk value is estimated as the sum of products of nodal failures (estimated as a ratio of number of samples with the total nodal head above the minimum required and the total number of samples) and consequences of these failures (deterministic values).

When calculating the value of objective functions, the conventional approach would be to use a Monte Carlo (MC) sampling technique with a large number of samples. Rather than doing that, in the approach presented here, the Latin Hypercube (LH) technique (with a small number of samples) is used. The main advantage of the LH over the MC sampling technique is the better random sample stratification leading to more accurate estimation of the empirical, nodal head PDF tails (important here when evaluating design robustness). The drawback associated with the use of LH sampling technique when compared to the MC sampling technique is the increased computational effort required to generate the same number of samples. Fortunately, when the number of samples is small (e.g. less than 100) this difference is not significant. In addition to this, for the problem considered here, the time required for the sample generation constitutes a negligible fraction of the total time required for the hydraulic solver calculations.

In this study a modified NSGAII (Deb et al. 2000) capable of performing efficient multi-objective search under uncertainty is used. The new GA is developed by Kapelan et al. (2006) and named the robust NSGAII (rNSGAII). The main idea behind the rNSGAII is to spread solution reliability evaluation over several generations, i.e. over the effective life of the corresponding chromosome, and to breed solutions (i.e. chromosomes) which are robust enough to survive over multiple generations. This enables calculation of each chromosome's objective as the average of

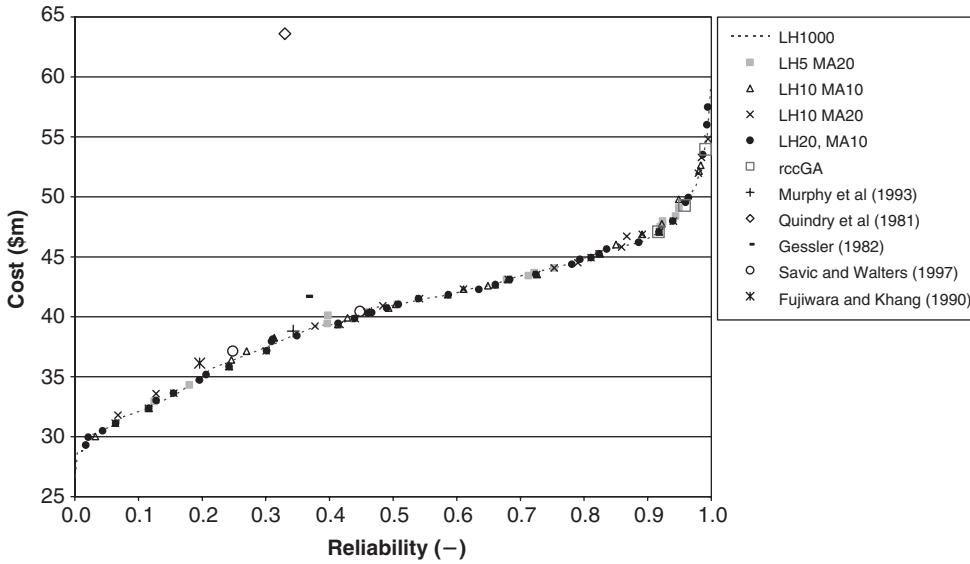


Figure 5. Total cost vs. reliability payoff (Kapelan et al., 2006).

past objective values over that chromosome's age. As a consequence, even if a small number (e.g. 5–20) of samples is used for each objective evaluation, the objectives are effectively evaluated using larger numbers of samples (e.g. 100–400 if the chromosome survived for 20 generations). This way, significant computational savings can be achieved when compared to the full sampling approach where, typically, thousands of samples must be used. More details about the algorithm can be found in Kapelan et al. (2006).

A number of deterministic solutions found in the literature are shown in Figure 5. Unlike in the deterministic approach of Quindry et al. (1981), Gessler (1985), Fujiwara & Khang (1990), Murphy et al. (1993) and Savic & Walters (1997), it is assumed in Kapelan et al. (2006) that nodal demands are uncertain variables following a Gaussian PDF with a mean equal to the deterministic demand value and a standard deviation equal to 10% of the mean value. The EPANET hydraulic solver (Rossman 2000) is used to calculate unknown heads and flows for each demand sample.

Due to stochastic nature of the genetic algorithm approach multiple rNSGAII runs are performed for both reliability and risk calculations using different numbers of LH samples ($N_s = 5$, 10 and 20), different required minimum ages of chromosomes ($MA = 5$, 10 and 20) and ten different random seeds (i.e. different initial populations). The solutions shown in the next section represent a selection of the solutions obtained in this process. In all runs a population of 200 chromosomes was used. All rNSGA runs were stopped after 500 generations. Once the optimisation runs were finished, the solution robustnesses and/or risk values were re-evaluated using 100,000 MC samples.

3.2.3.1 Reliability-based solutions

The Pareto optimal solutions obtained for the reliability-based problem formulation are shown in Figure 5. The curve "LH1000" represents the solutions obtained using the standard full-sampling approach. In this approach the standard NSGAII is used with 1000 LH samples for each fitness evaluation. The three points/solutions labelled rccGA (robust chance-constrained GA) represent the solutions obtained using the robust, single-objective water distribution system rehabilitation approach (Babayan et al. 2004, Kapelan et al. 2003).

All Pareto optimal fronts (Fig. 5) obtained using a small number of LH samples (LH5-LH20) are in a very good agreement with the full sampling approach front (LH1000). The coverage and

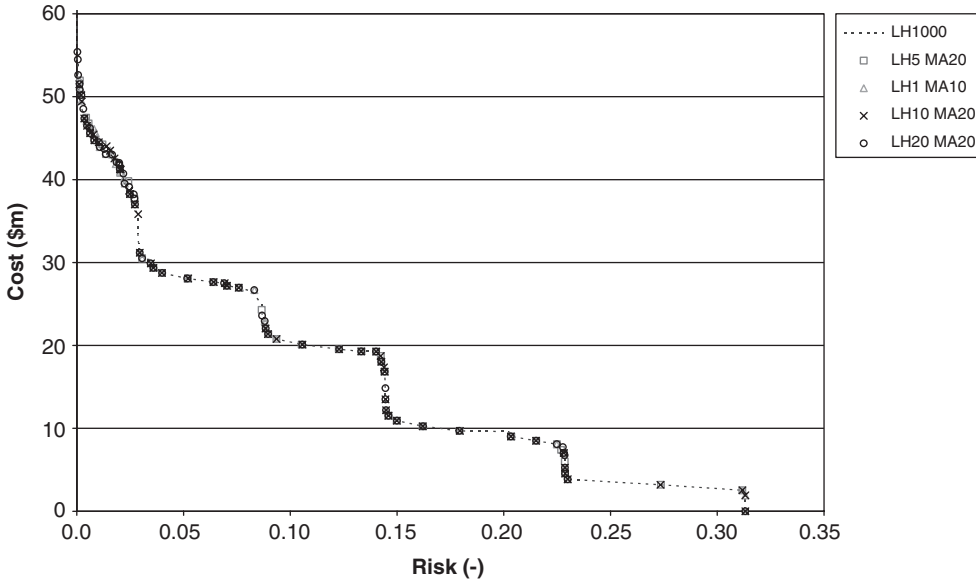


Figure 6. Total cost vs. risk payoff (Kapelan et al., 2006).

spread quality of the identified Pareto optimal fronts increases with the number of LH samples used. While the Pareto optimal front obtained for $N_s = 5$ LH samples is not very good in terms of both coverage (not uniform) and spread (neither of the front tails were identified) of the solutions obtained, the Pareto front for $N_s = 10$ samples and especially the front for $N_s = 20$ LH samples has excellent coverage and uniform spread of solutions. This is no surprise since the larger the number of samples the higher the resolution with which the solution robustness can be determined. The three optimal single-objective solutions obtained using the rccGA optimisation method (Kapelan et al. 2004) are in excellent agreement with corresponding Pareto optimal solutions. Most of the previously identified 'optimal' deterministic solutions are located close to the Pareto optimal front but are grouped in the low robustness zone (20–45%). The cost of rehabilitation solutions rises exponentially with high robustness levels (approximately 0.90 and above).

Analysis of the solutions obtained shows that those obtained using the reliability approach have higher cost than those obtained using the deterministic approach. This is expected because deterministic optimisation leaves very little redundancy in the system to cope with demand fluctuations, thus costing less than reliability based solutions, but being of lower reliability.

3.2.3.2 Risk-based solutions

The Pareto optimal solutions obtained for the risks-based problem formulation are shown in Figure 6. Again, as in the robustness-based approach, the curve "LH1000" represents the solutions obtained using the standard NSGAII with 1000 LH samples for each fitness evaluation.

The shape of the Pareto optimal front obtained is significantly different from the one obtained using the reliability-based model. This is, basically, a consequence of the definition of risk adopted. All Pareto optimal fronts obtained using small number of LH samples (LH5-LH20) have good coverage and spread of points and are in very good agreement with the full sampling approach (LH1000. The cost of rehabilitation solutions rises exponentially for decreasingly low risk values (approximately 0.02 and below). Similar to reliability-based solutions all solutions obtained using the risk-based approach have higher cost than those obtained using the deterministic approach. In addition to that, risk-based optimisation tends to lower (as much as possible) the risk at nodes with high demand due to the particular definition of risk used in this study.

4 SHORT-TERM REHABILITATION PLANNING

The economic and social costs of pipe bursts in water distribution networks are very significant. Water managers need reliable replacement plans for critical pipes, balancing investment with expected benefits in a risk-based management scenario. Thus, a robust and feasible decision support tool for water system rehabilitation is required. This kind of tool should incorporate: (i) a model to forecast pipe failures and (ii) a strategy to solve a multi-objective optimisation problem trading investment vs. benefits (Giustolisi et al. 2006).

4.1 *UK case study (Giustolisi et al. 2006)*

The data used in this case study were collected at pipe level from 1986 to 1999 and harvested for both asset information and burst records during the monitoring period. In particular, the database used for this case study refers to a single Water Quality Zone (WQZ) within a wide water distribution network (further details available in Berardi et al. 2005). Only four of the asset attributes describing pipe features have been chosen among all available information. These are age (Aep), diameter (Dp), length (Lp) and number of properties (Prp) supplied, all available at pipe level.

A statistical approach to modelling water network failures is justified for small pipes only, since the surveys required for calibrating any physical model are not economically viable. Therefore, only small diameter pipes (i.e. nominal diameter up to 250 mm) have been selected among all available assets. Evolutionary Polynomial Regression (EPR) (Giustolisi & Savic, 2006), is used as a data-mining tool. Models returned by EPR are explicit expressions, which clearly identify the most influencing factors for pipe bursts (Berardi et al. 2005). Although the returned models confirms that the age is an important factor for predicting bursts (*BR*), they also show that it is not the only factor which significantly influences the deterioration of water mains. Pipe length and diameter play an important role in determining the number of bursts. Furthermore, the equation confirms findings in scientific/technical literature on the subject that more bursts occur on small diameter pipes (Walski et al. 1986).

Due to the fact that failures recorded during the monitoring period occur on a small number of pipes, most of them experiencing more than one burst, there were insufficient data to assess individual failure rates for most of the pipes. As a remedy, previous studies suggest associating the same failure rate to pipes showing similar features (i.e. material, size, age, etc.). Thus, the whole asset database was divided next into 8 classes based on pipe diameters in the WQZ (ranging from 32 mm to 250 mm). Furthermore, each diameter class has been further split into one-year long age sub-classes. Of all the models (functional expressions) returned for these pipe classes, the following one was considered the best (selected based on its parsimony and statistical quality). Parsimony is indicated by the number of input variables selected among those available and statistical quality by the Coefficient of Determination (Giustolisi et al. 2006):

$$\mathrm{BR} = \left(6.15 \cdot 10^{-3} \frac{\mathrm{Lt}}{\mathrm{De}^2}\right) \mathrm{Ae}^2 + 0.436$$

where Lt is the total length of pipes in a class, and Ae and De are equivalent pipe age and diameter, respectively, computed as:

$$\mathrm{Ae} = \frac{\sum_{\mathrm{class}} \mathrm{Lp} \cdot \mathrm{Aep}}{\mathrm{Lt}}$$

$$\mathrm{De} = \frac{\sum_{\mathrm{class}} \mathrm{Lp} \cdot \mathrm{Dp}}{\mathrm{Lt}}$$

The obtained EPR model structure is influenced by the proposed classification scheme, i.e., it describes the propensity of pipes to fail in terms of an aggregate model. Before such a model could be applied at the pipe level (e.g. in supporting decision for individual pipe replacement) a further procedure is required, as described in Giustolisi et al. (2006).

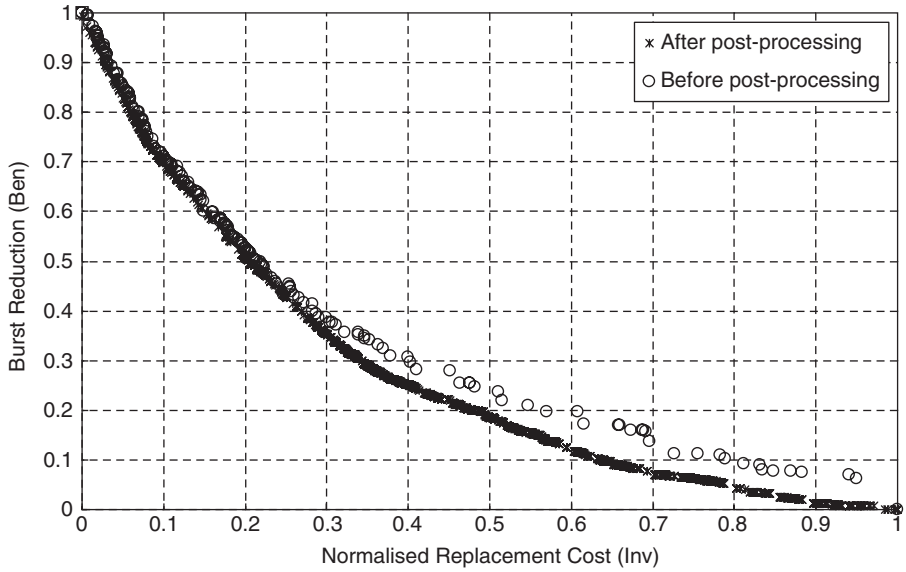


Figure 7. Investment/benefit diagram (Giustolisi et al., 2006).

4.1.1 *Rehabilitation optimization model*

For an optimised replacement plan to be developed one need to find a subset of mains for replacement which simultaneously minimizes: (i) the sum of BRp for the whole network; and (ii) the investment for pipe replacement. This problem can be defined as a typical investment/benefit optimization in which the benefits of reducing the number of bursts are balanced against the required investment for pipe replacement. Genetic algorithms have proven to be a powerful tool for solving this kind of problem. In this study a multi-objective GA, OPTIMOGA, which is based on Pareto dominance principles, has been used to solve this problem.

Two objective functions are used: (i) the pipe replacement cost (investment) and (ii) the bursts reduction (benefit). The investment for each pipe replacement is assumed to be directly dependent on diameter and pipe length and computed using an empirical formula, which in this case was determined using typical Italian rehabilitation costs as guidance. Next, every single cost is normalised with respect to the full-pipe replacement scheme expense (i.e., all pipes in the network replaced), Inv(De, Lp). Benefits (Ben) are evaluated in terms of burst reduction as the ratio between the total expected BRp after rehabilitation and the total predicted BRp without any pipe replacement. This normalises the benefit function between 0 and 1, thus making the achieved benefits clearly understandable. Moreover, because economical and social costs associated with bursts vary with diameter, a re-scaling factor for the number of bursts is assumed for different diameters. Without loss of generality the adopted scale factor is assumed to be 1 for De = 32 mm. The model is then used to forecast pipe bursts for a 1-year planning horizon. That is followed by the increase in age for all pipes in the selected WQZ, which is followed by the assessment of BRp values for all the pipes. Therefore, the evolutionary optimisation runs were performed by setting BRp = 0 for those pipes which are to be replaced. Figure 7 shows the investments/benefit diagram obtained by the evolutionary optimisation algorithm as the Pareto front of non-dominated solutions (circles in Figure 7).

4.1.2 *Post-processing of optimization results*

The Pareto front (circles in Figure 7) could be considered as a means for supporting decision making for pipe rehabilitation planning. Each point on the investments/benefit diagram corresponds to a subset of pipes for replacement. This could potentially allow planning rehabilitation at pipe level. However, this diagram is not directly useful for assessing the operational replacement decisions.

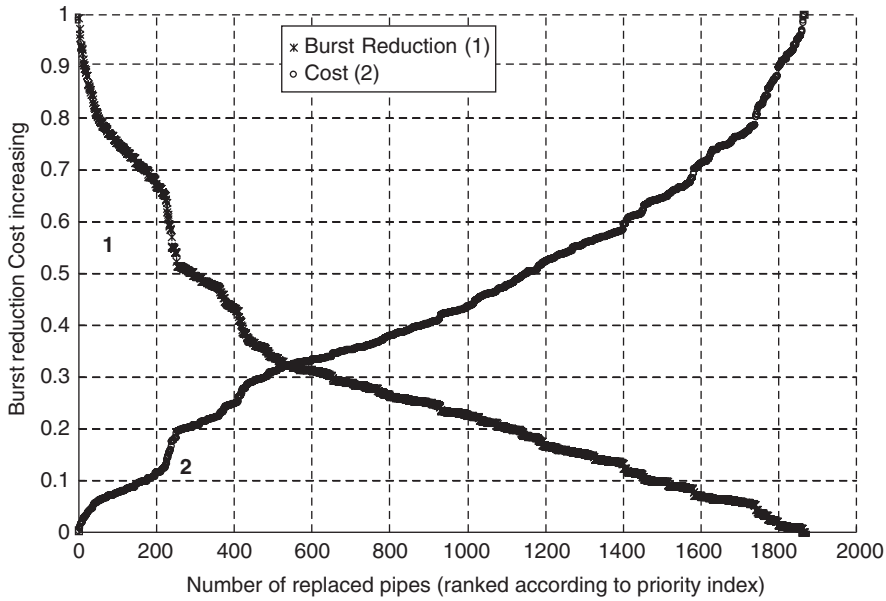


Figure 8. Decision support diagram (Giustolisi et al., 2006).

In other words, it does not provide any priority information on the pipe rehabilitation even when a single point (solution on the Pareto front) is selected, but identifies a set of replacement plans with similar benefits against a large investment variation. Moreover, it assumes that the overall rehabilitation will be carried out at the same time, which is unrealistic.

In order to obtain a better means for supporting operational decision making, a post-processing phase is proposed which starts from the Pareto front identified in Figure 7. The key issue is how to arrive at a priority list (or a priority index) for scheduling of pipe replacements. It is suggested here that the priority index is defined in terms of a probability computed as the ratio between the number of times a pipe was present in the solutions on the Pareto front and the total number of the solutions on the Pareto front (i.e., replacement subsets). Curve no. 1 in Figure 8 represents a prioritisation scheme for pipe replacement, which provides the percentage reduction in the number of predicted bursts (benefit) as a cumulative value related to the substitution of the relevant pipes (on the x-axis).

In Figure 8, the mains are ranked according to how often they appear in the solutions on the Pareto front, thus the most critical pipes are grouped on the left of the curve. Furthermore, by linking the priority information (probability of a pipe appearing in a Pareto front solution) to the cumulative renewal costs an additional diagram is provided (Fig. 4, curve no. 2). This last diagram provides the cumulative cost (along the y-axis) related to the replacement of the corresponding pipes (along the x-axis), normalised by dividing it by the renewal cost of all pipes. Each pipe of the WQZ can be identified by its own probability of replacement (priority within a rehabilitation scheme), and by its burst reduction rate (i.e., benefit, which can be considered as its contribution to the overall deterioration of the network). In this way, a new front of solutions (stars in Figure 7) is constructed according to the probability of replacement of each main, starting from the most critical one. Analysis of this new front in Figure 7 (stars) shows that the new front of solutions dominates that found by means of evolutionary optimisation.

Therefore, the proposed post-processing phase led to the development of three useful tools (diagrams) for replacement planning. By using knowledge of available annual budgets for rehabilitation, decision makers can use the normalised value of the annual investment threshold (Fig. 7, stars) to identify which pipes have to be replaced and what is the most effective replacement sequence.

Furthermore, they can decide on what level of replacement they can afford during annual planning by using the diagrams of Figure 8. For example, Figure 8 shows that if the 200 most critical pipes are replaced, the normalised benefit of 0.3 (bursts reduction of 30% with respect to a no-replacement scenario) will be achieved with a normalised replacement cost of 0.1 (10% of the full replacement expense). At the same time, doubling the investment, it is possible to get a further bursts reduction of 20% by replacing about 50 pipes, i.e. the most critical pipes with large diameters, which also have a higher mean length than the previously identified 200 pipes. Moreover, diagrams in Figure 8 highlight the presence of an investment threshold (of about 30 % of the total replacement expense) beyond which small bursts reductions require large amounts of money.

5 SUMMARY AND CONCLUSIONS

The paper presents three case studies in which optimization were used for design and management planning purposes. The case of Anytown network was considered first, where the design variables were the pipe rehabilitation decisions, tank sizing, tank sitting, and pump operation schedules. To provide flexibility, the network was designed and operated under multiple loading conditions. Inclusion of pump operation schedules requires consideration of water system operation over an extended period. The cost of the solution included the costs of pipes and tanks (capital costs) as well as the present value of the energy consumed during a specified period (operational costs). The optimization problem was solved using the non-sorting dominated genetic algorithm, NSGAII. Optimization tends to reduce costs by reducing the diameter of, or completely eliminating, some pipes, thus leaving the system with insufficient capacity to respond to pipe breaks or demands that exceed design values without violating required performance levels. A resilience index was thus considered as a second objective to increase the hydraulic reliability and availability of water during pipe failures. Sensitivity analysis of solutions on the Pareto (trade-off, payoff) front generated by twin-objective optimization shows poor performance of these networks under random pipe failure or a pump being out of service. The minimum surplus head was then added as a third objective to overcome the shortcomings of the resilience index.

The New York tunnels rehabilitation was the second problem, which was defined as a multi-objective optimisation problem under uncertainty. Two alternative problem formulations were considered. The first objective in both approaches is to minimise the total (capital) rehabilitation cost. The second objective is to either maximise the overall reliability or to minimise the total risk associated with the rehabilitation scheme. The reliability was defined as the probability of simultaneously satisfying minimum pressure head constraints at all nodes in the network. Total risk is defined as the sum of nodal risks, where nodal risk is defined as the product of the probability of pressure failure at that node and consequence of such failure. Decision variables are the alternative rehabilitation options for each pipe in the network. The only source of uncertainty considered was the future water consumption. Uncertain demands were modelled using a probability density function (PDF) assigned in the problem formulation phase. The corresponding PDFs of the analysed nodal heads are calculated using the Latin Hypercube sampling technique. The optimal rehabilitation problem was solved using the so-called robust NSGAII method which is a modification of the well-known NSGAII optimisation algorithm. In rNSGAII a small number of demand samples are used for each fitness evaluation leading to significant computational savings when compared to the full sampling approach. The results obtained demonstrate that both new methodologies are capable of identifying the robust (near) Pareto optimal fronts while making significant computational savings.

The case of a UK distribution network was considered last. Rehabilitation needs were defined in terms of pipe burst reduction and the tool used incorporated a model to forecast pipe failures and a multi-objective optimization tool that identifies non-dominated (Pareto) optimal solutions trading investment vs. benefits. The burst modelling was performed using the evolutionary polynomial regression (EPR) technique, providing a symbolic model for predicting pipe bursts. The benefits of burst reduction achieved by mains rehabilitation were evaluated by another multi-objective

optimization model over a short-term planning horizon (taken to be one year in this study). The multi-objective model was based on another genetic algorithm methodology. The procedure identifies different subsets of pipes scheduled for rehabilitation ranging from no-replacement (i.e., no reduction of the predicted number of bursts) to the complete replacement scheme (i.e., maximum reduction of the predicted number of bursts) trading cost of rehabilitation against achieved benefits. The result of the strategy is a Pareto front, which by itself does not provide any prioritisation of pipes for replacement. Thus, the study introduces a further processing step by which pipes are prioritised for rehabilitation based on the number of times each belongs to a solution on the Pareto front. By considering costs and such priority rating of each main, an improved investments/benefit diagram can be constructed.

The work presented in this paper goes some way to demonstrate the advances in optimal design and management of water distribution systems achieved in the last decade or so. Since the time of Goulter's paper on why optimization has not had anticipated impact on engineering practice (Goulter 1992), the development of new optimization techniques based primarily on evolutionary computing techniques (but also including simulated annealing, ant colony optimization, particle swarm optimization, etc) has made significant progress in addressing the four reasons for the lack of wide acceptance of optimization. The models now work with practical problems incorporating both design and operational issues. They also find solutions that offer significant improvements over established methods and professional judgement. By incorporating optimization models into commercially available simulation packages the idea of optimization DSS has become a reality as optimization is not as difficult to use as it was the case some 20 years ago when the first Battle of the Network Models (Walski 1987) was organised. Finally, practitioners are more comfortable with the overall optimization approach as evidenced by the number of organisations offering optimization on a consultancy basis (Lansey 2006).

The use of multi-objective optimization, which trades off cost and benefits, has addressed the main criticism raised by Walski (2001), which stated that the use of cost optimization does not consider benefits stripped off the system due to excessive efficiency gains (reduction in pipe diameters, storage provision, etc), and that raised by Haimes (2001) who sees single-objective (cost) optimization as misuse of optimization. The example of risk-optimization (Kapelan et al. 2006) also demonstrates that consequences of failures should be considered in the context of risk, rather than just reliability (probability of failure), as it is common in the water distribution design literature. However, this is not to say that the work required on improving optimization models is done and further research is not needed. For example, consideration of multiple-stage designs (over a period of time) should be further improved. Furthermore, the treatment of water quality and its impacts on water distribution system design has to be researched and brought into the practice. Real-time operation will also need attention as water service providers install more sensors into their systems. Calibration of water distribution systems also needs further attention as more accurate hydraulic models are needed for the use with optimization algorithms.

REFERENCES

Babayan, A.V., Kapelan, Z., Savic, D.A. and Walters, G.A. (2004). Comparison of two approaches for the least cost design of water distribution systems under uncertain demands. In 6th Annual Symposium on Water Distribution Systems Analysis, Salt Lake City, Utah, USA.

Babayan, A.V, Z. Kapelan, D.A. Savic, and G.A. Walters (2005), Least Cost Design of Robust Water Distribution Networks Under Demand Uncertainty, Journal of Water Resources Planning and Management, ASCE, Vol. 131, No. 5, pp. 375–382.

Babayan, A.V., Kapelan, Z., Savic, D.A., and Walters, G.A. (2006), Comparison of Two Approaches for the Least Cost Rehabilitation of Water Distribution Systems under Uncertainty, Engineering Optimization, 38 (3), pp. 281–297.

Babayan, A.V., D.A. Savic, G.A. Walters, and Z.Kapelan (2007) Robust Least-Cost Design of Water Distribution Networks Using Redundancy and Integrated Based Methodologies, Journal of Water Resources Planning and Management, ASCE (in press).

Berardi, L., Savic, D.A. and Giustolisi, O. (2005) Investigation of Burst-Prediction Formulas for Water Distribution Systems by Evolutionary Computing, in Proc. of the Eight International Conference on Computing and Control for the Water Industry, CCWI2005, Savic, D.A., Walters, G.A., King, R. and Khu, S-T. (eds.), University of Exeter, UK, Vol. 2, pp. 275–280.

Centre for Water Systems, CWS (2007) http://www.projects.ex.ac.uk/cws/index.php?benchmarks (last accessed 27/04/2007).

Deb, K., Agrawal, S., Pratapm, A., and Meyarivan, T. (2000). "A fast and elitist multi-objective genetic algorithm: NSGA-II." Technical Rep. 200001, Indian Institute of Technology, Kanpur Genetic Algorithms Laboratory (KanGAL), Kanpur, India.

Farmani, R., G.A.Walters, D.A.Savic, (2005) Trade-off between Total Cost and Reliability for Anytown Water Distribution Network, ASCE Journal of Water Resources Planning and Management, Vol. 131, No. 3, pp. 161–171.

Fujiwara, O. and Khang, D.B. (1990). A two-phase decomposition method for optimal design of looped water distribution networks. Water Resources Research, 26(4), 539–549.

Gessler, J. (1985). Pipe network optimization by enumeration. In Special Conf. on Computer Applications/Water Resources, ASCE, New York, pp. 572–581.

Giustolisi, O., D Laucelli and D A Savic (2006) Development of Rehabilitation Plans for Water Mains Replacement Considering Risk and Cost-Benefit Assessment, Civil Engineering and Environmental Systems, Vol. 23, No. 3., pp. 175–190.

Giustolisi, O. and D.A. Savic (2006) A Symbolic Data-Driven Technique Based on Evolutionary Polynomial Regression, Journal of Hydroinformatics, Vol. 8, No. 3, pp. 207–222.

Goulter, I.C. (1992), "Systems analysis in water-distribution network design: from theory to practice", Journal of Water Resources Planning and Management, ASCE, Vol. 118 No. 3, pp. 238–48.

Guardian Newspaper (2007) Sanitation rated the greatest medical advance in 150 years, Friday, January 19, also available at: http://www.guardian.co.uk/medicine/story/0,,1994182,00.html?gusrc = ticker–103704 (last accessed 25/04/07).

Haimes, Y. Y. (2001). Water system complexity and the misuse of modeling and optimization. In Y. Y. Haimes, D. A. Moser, & E. Z. Stakhiv (Eds.), Risk-Based Decisionmaking in Water Resources IX. New York : ASCE.

Kapelan, Z., Savic, D. and Walters, G.A. (2003). Robust least cost design of water distribution systems using GAs. In Computer Control for Water Industry (CCWI), London, UK, pp. 147–155.

Kapelan, Z., A.V. Babayan, D.A. Savic, G.A. Walters, and S.T. Khu (2004) Two new approaches for the stochastic least cost design of water distribution systems, Water Science & Technology: Water Supply, IWA, Vol. 4, No. 5–6, pp. 355–363.

Kapelan, Z., D.A. Savic and G.A. Walters (2005) Multi-objective Rehabilitation of Water Distribution Systems under Uncertainty, Water Resources Research Vol. 41, No. 11, W11407.

Kapelan Z., Savic, D.A., Walters, G.A. and A.V. Babayan (2006) Risk and Robustness based Solutions to a Multi-objective Water Distribution System Rehabilitation Problem under Uncertainty, Water Science & Technology, IWA, Vol. 53, No. 1, pp. 61–75.

Lansey (2006) The Evolution of Optimizing Water Distribution System Applications. Proceedings of the 8th Water Distribution System Analysis Symposium, Cincinnati USA, August 27–30.

Life Magazine (1997) Life Magazine, 1997. The millennium: The 100 events headline: No. 46; water purification.

Murphy, L.J., Simpson, A.R. and Dandy, G.C. (1993). Pipe network optimization using an improved genetic algorithm. Report No R109, Department of Civil and Environmental Engineering, University of Adelaide, Australia.

Murphy, L.J., Dandy, G.C., and Simpson, A.R. (1994). "Optimum design and operation of pumped water distribution systems." Proc., Conf. on Hydraulics in Civil Engineering, Institution of Engineers, Brisbane, Australia, pp. 149–155.

Obradovic, D., and Lonsdale, P. (1998). Public water supply, E&FN Spon, London.

Ormsbee, L.E. (2006) The History of Water Distribution Network Analysis: The Computer Age. Proceedings of the 8th Water Distribution System Analysis Symposium, Cincinnati USA, August 27–30.

Quindry, G.E., Brill, E.D. and Liebman, J.C. (1981). Optimisation of looped water distribution systems. ASCE Journal of the Environmental Engineering Division, 107(EE4), 665–679.

Rossman, L.A. (2000). EPANET users manual, U.S. EPA, Cincinnati, Ohio.

Savic, D.A. and Walters, G.A. (1997). Genetic algorithms for the least-cost design of water distribution networks. Journal of Water Resources Planning and Management, ASCE, 123(2), 67–77.

Schaake, J. and Lai, D. (1969). Linear programming and dynamic programming applications to water distribution network design. Report No 116, Dept. of Civil Engineering, MIT, Cambridge, MA, USA.

Tanyimboh, T.T., and Templeman, A.B. (1993). “Optimum design of flexible water distribution networks.” Civ. Eng. Syst., 10(3), 243–258.
Todini, E. (2000). “Looped water distribution networks design using a resilience index based heuristic approach.” Urban Water, 2(3), 115–122.
Tolson, B.A., Maier, H.R., Simpson, A.R., and Lence, B.J. (2004). “Genetic algorithms for reliability-based optimization of water distribution systems.” J. Water Resour. Plan. Manage., 130(1), 63–72.
Tung, Y.-K. (1985). “Evaluation of water distribution network reliability.”Hydraulics and hydrology in the small computer age, W. R. Waldrop, ed., ASCE, New York, pp. 359–364.
Wagner, J.M., Shamir, U., and Marks, D.H. (1988). “Water distribution reliability: Analytical methods.” J. Water Resour. Plan. Manage., 114(3), 253–275.
Walski, T.M., T.M., Brill, E.D., Gessler, J., Goulter, I.C., Jeppson, R.M., Lansey, K., Lee, H.L., Liebman, J.C., Mays, L.W., Morgan, D.R., and Ormsbee, L.E. (1987). “Battle of networks models: Epilogue.” J.Water Resour. Plan. Manage., 113(2), 191–203.
Walski, T.M. (1985). “State-of-the-Art: Pipe Network Optimization.” Computer Applications in Water Resources, Torno, H., ed., ASCE, New York, New York.
Walski, T.M. (2000). “Hydraulic design of water distribution storage tanks.” Water distribution systems handbook, L. Mays, ed., Chap. 10, pp. 10.1–10.20.
Walski, T.M. (2001). “The Wrong Paradigm—Why Water Distribution Optimization Doesn’t Work.” Journal of Water Resources Planning and Management, ASCE, 127(4): 203–205.
Walters, G.A., Halhal, D., Savic, D., and Ouazar, D. (1999). “Improved design of Anytown distribution network using structured messy genetic algorithms.” Urban Water, 1(1), 23–38.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

The influence of source terms on stability and conservation in 1D hyperbolic equations: Application to shallow water on fixed and movable beds

J. Murillo, P. García-Navarro & P. Brufau
Fluid Mechanics, CPS, University of Zaragoza, Zaragoza, Spain

J. Burguete
E.E.A.D., C.S.I.C., Zaragoza, Spain

ABSTRACT: The non conservative terms that appear in systems of conservation laws introduce an extra difficulty in practical applications. The influence of these terms in the stability region for explicit finite volume schemes is analyzed under the hypothesis of a well balanced scheme, based on the C-Property. It is demonstrated that in each particular situation the stability region is amplified or reduced, and how in some special cases it depends on the initial conditions. The time step requirements for stability are defined at the cell edges, related with the traditional Courant-Friedrichs-Lewy (CFL) condition. This is applied to the plain shallow water equations, to the shallow water equations with solute transport and to sediment transport problems.

1 INTRODUCTION

In the basis of a one-dimensional mathematical model for debris flow, Brufau et al., 2000 presented a Godunov-type finite volume scheme based on Roe's numerical approach (Roe, 1986). Hudson and Sweby, (Hudson and Sweby, 2003) investigated the equations governing bed-load sediment transport for steady and unsteady approaches using a flux-limited version of Roe's scheme and extended it to the two-dimensional case (Hudson and Sweby, 2005) requiring an exact balance between the flux gradient and the source term, following the idea of exact conservation (C-property) introduced by Bermudez and Vazquez (Bermudez and Vazquez, 1994). Crnjaric-Zic et al, 2004 extended other Godunov-type schemes, ENO and WENO, to the one dimensional sediment transport equations, using the mathematical model presented by Hudson, (Hudson, 2001) also requiring the C-property. In their work the trust term was included in the Jacobian matrix of the flux assuming, by means of a local linearization, that the system remained hyperbolic. Also the C-property was required in the case of quiescent flow. Rosatti and Fraccarollo, 2006 presented a mathematical model for high concentration sediment transport that was solved by means of a Godunov-type scheme requiring the C-property and leading to accurate results.

In Murillo et al., 2006 and in Murillo et al., 2007 the stability region for systems of conservation laws was analyzed, making emphasis on the numerical conservation and on the preservation of the positivity property of the solution when necessary in the presence of source terms. The numerical approach was based on the approximate Roe's solver, and applied to the shallow water equations, where the source term treatment satisfied the C-property including both bed and friction terms, leading to an exact balance among fluxes and non conservative terms, not only in the case of quiescence flow over discontinuous bed topography, also in the case of steady fluid in motion.

This is a conservation law expressing that a function u varies and is transported according to both the distribution of a flux function f and to a source term s, in the form:

$$\partial_t u + \partial_x f = s(u,x) \tag{1}$$

At this point, the source term is assumed to follow $s = \partial_x \tau$, and the advection, or transport, velocity λ is:

$$\lambda = df/du \tag{2}$$

To introduce the upwind finite volume scheme, (1) is integrated in a volume Ω:

$$\frac{\partial}{\partial t}\int_\Omega u(x,y)d\Omega + \int_\Omega \partial_x(f-\tau)d\Omega = 0 \tag{3}$$

where $d\Omega$ denotes the volume boundary. If Gauss's theorem is applied to the second integral in (3):

$$\frac{\partial}{\partial t}\int_\Omega u\, d\Omega + \oint_{\partial\Omega}(f-\tau)dl = 0 \tag{4}$$

In the one-dimensional case, the volumes are actually lines and, considering a uniform mesh, $\Omega = \Delta x$ is a constant. Now, a cell-centred finite volume method can be formulated where all the dependent variables of the system are represented as piecewise constants (first order). Therefore quantities u, f and τ are uniform per cell of length Δx. In particular, the first integral in (4) can be approximated by:

$$\frac{\partial}{\partial t}\int_\Omega u\, d\Omega = \frac{\partial u_i}{\partial t}\Delta x \cong \frac{u_i^{n+1} - u_i^n}{\Delta t}\Delta x \tag{5}$$

In the second integral of (4):

$$\oint_{\partial\Omega} f = f_{i+1} - f_{i-1} = \delta f_{i+1/2} + \delta f_{i-1/2} \tag{6}$$

with $\delta f_{i+1/2} = f(u_{i+1}) - f(u_i) = f_{i+1} - f_i$.

Using (2), the linearized advection velocity $\tilde{\lambda}_{i+1/2}$ can be defined (Roe, 1986) as:

$$\tilde{\lambda}_{i+1/2} = \frac{(f_{i+1} - f_i)}{(u_{i+1} - u_i)} \tag{7}$$

Following the upwind philosophy, which discriminates the sense of propagation according to the sign of the advection velocity, the flux difference is split (Roe, 1986) as a sum of waves travelling in and out of a given cell:

$$\delta f_{i+1/2} = \delta f^-_{i+1/2} + \delta f^+_{i+1/2} = \tilde{\lambda}^-_{i+1/2}\delta u_{i+1/2} + \tilde{\lambda}^+_{i+1/2}\delta u_{i+1/2} \tag{8}$$

with $\tilde{\lambda}^\pm = (\tilde{\lambda} \pm |\tilde{\lambda}|)/2$ and $\tilde{\lambda}_{i+1/2} = \tilde{\lambda}^-_{i+1/2} + \tilde{\lambda}^+_{i+1/2}$.

In the case of the source term τ:

$$\oint_{\Delta x} \tau = \tau_{i+1} - \tau_i = \delta\tau_{i+1/2} + \delta\tau_{i-1/2} \tag{9}$$

The edge source term $\delta\tau_{i+1/2}$ is also split into in-going and out-going contributions

$$\delta\tau_{i+1/2} = (\delta\tau)^-_{i+1/2} + (\delta\tau)^+_{i+1/2} \tag{10}$$

where

$$(\delta\tau)^{\pm}_{i+1/2} = \tfrac{1}{2}(1 \pm \mathrm{sgn}(\tilde{\lambda}^{\pm}_{i+1/2}))(\delta\tau)_{i+1/2} \tag{11}$$

The updating scheme for cell i includes only the in-going contribution of flux and source term to that cell as follows

$$u_i^{n+1} = u_i^n - [\delta(f-\tau)^{+}_{i-/2} + \delta(f-\tau)^{-}_{i+/2}]\frac{\Delta t}{\Delta x} \tag{12}$$

Equation (12) is the so called unified discretization (Burguete and García-Navarro, 2001). This can also be expressed in a more compact form as:

$$u_i^{n+1} = u_i^n - [\nu^{+}_{i-/2}\delta u_{i-/2} + \nu^{-}_{i+/2}\delta u_{i+/2}], \quad \nu^{\pm}_{i\mp1/2} = \lambda^{*,\pm}_{i\mp1/2}\frac{\Delta t}{\Delta x} \tag{13}$$

with

$$\lambda^{*,\pm}_{i\mp1/2} = \lambda^{\pm}_{i\mp1/2}\theta^{\pm}_{i+1/2}, \quad \theta^{\pm}_{i\mp1/2} = 1 - \left(\frac{\delta\tau}{\tilde{\lambda}\delta u}\right)^{\pm}_{i\mp1/2} = 1 - \left(\frac{\delta\tau^{-}}{\delta f^{-}}\right)^{\pm}_{i\mp1/2} \tag{14}$$

Note that $\nu^{\pm}_{i\mp1/2}$ is a dimensionless quantity that plays the role of a local CFL number generalized to cases involving source terms. The coefficient $\theta^{\pm}_{i\mp1/2}$ expresses the discrete ratio of source term to flux differences. They will be key parameters in our discussion of stability conditions.

2.1 *Influence of $\theta^{\pm}_{i+1/2}$ on the stability condition*

In the homogeneous case $\theta^{\pm}_{i+1/2} = 1$. Then, numerical stability for scheme (13) is ensured if

$$0 \le \nu^{+}_{i-1/2} \le 1, \quad -1 \le \nu^{-}_{i+1/2} \le 0 \tag{15}$$

At the same time, the following condition for the monotonicity of the solution holds

$$u^{\min} \le u_i^{n+1} \le u^{\max} \tag{16}$$

where $u^{\max} = \max\{u_i^n, u_{i+1}^n, u_{i-1}^n\}$ and $u^{\min} = \min\{u_i^n, u_{i+1}^n, u_{i-1}^n\}$.

The aim of the following analysis is to find the criterion that preserves (16) in presence of source terms. For that reason, first it is necessary to enforce (16) by requiring

$$\tilde{\lambda}^{+,*}_{i-1/2} \ge 0, \quad \tilde{\lambda}^{-,*}_{i+1/2} \le 0 \tag{17}$$

this is possible when

$$\theta^{\pm}_{i\mp1/2} \ge 0 \tag{18}$$

Given a grid mesh and initial data values, (16) is a limit on the value of the time step to meet the stability criterion. For the sake of simplicity, assume that all the u_{i+1} and u_{i-1} values are equal but $u_i > u_{i+1}$. The size of the allowable incoming contributions is bounded by

$$|\nu^{+}_{i-/2}\delta u_{i-/2} + \nu^{-}_{i+/2}\delta u_{i+/2}| \le \lambda^{*}_{\max}\delta u_o\frac{\Delta t}{\Delta x} \tag{19}$$

where $\lambda^{*}_{\max} = \max\{|\lambda^{*,+}_{i-1+2}|, |\lambda^{*,-}_{i+1+2}|\}$ and $\delta u_o = u_i - u_{i+1}$. As the updating flux must be limited by the quantity that ensures that the final state is included between the initial values, the following is also true:

$$\lambda^{*}_{\max}\delta u_o\frac{\Delta t}{\Delta x} \le \delta u_o \tag{20}$$

being δu_o the initial difference. Under these conditions the maximum allowable time step is formulated in terms of the edge-time steps as follows:

$$\Delta t = CFL\Delta t_{\max}, \quad CFL \leq 1$$
$$\Delta t_{\max} = \min\{\Delta t_{i+1/2}\}_{k=1,N} \quad \Delta t_{i+1/2} = \frac{\Delta x}{|\lambda^*_{i+1/2}|}, \quad \theta_{i+1/2} \geq 0 \tag{21}$$

If $\theta_{i+1/2}$ is set equal to one in (21), the basic CFL stability condition for the homogeneous case is automatically recovered. Otherwise, (21) states a more general rule. It is worth remarking the relevance of the source term discretization when analysing the stability region defined by (21). If a unified formulation has been used so that in steady state

$$(\delta(f-\tau))^-_{i+1/2} = 0 \quad (\delta(f-\tau))^+_{i-1/2} = 0 \tag{22}$$

and this means that $\theta^{\pm}_{i\mp1/2} = 0$ and therefore $\lambda^{*,\pm}_{i\mp1/2} = 0$. The numerical scheme becomes unconditionally stable at steady state, and the C-Property is achieved.

Figure 1 represents the stability region of the scheme as a function of θ. The point $\theta = 1$ on the curve corresponds to the homogeneous case (no source term) and $\Delta t(\lambda^*_{i+1/2})$ is the maximum time step compatible with stability in this case or CFL condition (Courant et al. 1952).

The rest of the curve corresponds to cases with source terms. In Figure 1 the influence of the relative sign of flux and source terms is displayed. The dashed zone ($0 \leq \theta \leq 1$) is the set of situations in which it could be possible to use larger time steps than the limit of the homogeneous stability: both fluxes and source terms have the same sign so the total contribution is reduced. In the case $\theta > 1$, stability requires a reduction of the time step size over that dictated by the CFL condition, as fluxes and source terms have an opposite sign, so the net contribution increases.

On the other hand it is possible to find situations where $\theta_{i+1/2} < 0$. In that case the source term contribution dominates over the flux difference, $|\delta\tau^-|_{i+1/2} > |\delta f^-|_{i+1/2}$ and a different line of reasoning must be followed. Depending on the requirements over the solution, different strategies can be adopted. If preservation of the sign of the variable is the main objective,

$$u^{n+1}_i \leq 0 \text{ when } u^n_i, u^n_{i\pm1} \leq 0, \quad u^{n+1}_i \geq 0 \text{ when } u^n_i, u^n_{i\pm1} \geq 0 \tag{23}$$

the contributions between cells are also limited by the initial values stored in the cells so that the stability condition is determined also by the initial condition:

$$\Delta t_{i+1/2} = \gamma\frac{\Delta x}{|\lambda^*_{i+1/2}|}, \quad \gamma = \frac{\min\{|u_i|, |u_{i+1}|, |\delta u_{i+1/2}|\}}{|\delta u_{i+1/2}|}, \quad \theta^{\pm}_{i+1/2} < 0 \tag{24}$$

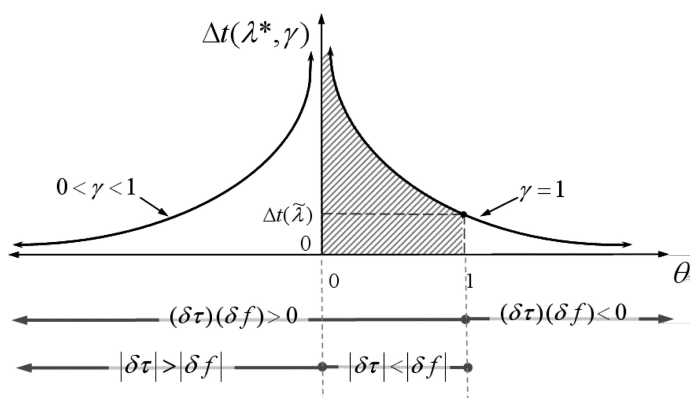


Figure 1. Stability region. First order scheme in presence of source terms.

If u is a gradually varied function, the coefficient γ is 1 and the time step limit in (24) reduces to (21). Otherwise, $0 \leq \gamma < 1$ and an actual reduction in the time step is required. In the special case $\gamma = 0$ the local time step $\Delta t_{i+1/2}$ would be zero according to (24). This is absurd and must be interpreted as a condition of no information crossing that cell edge. In practice, a threshold value equal to the machine accuracy is defined for the minimum value of γ before imposing the condition that no information crosses the edge.

3 SYSTEMS OF CONSERVATION LAWS WITH SOURCE TERMS

The numerical method is extended in this chapter to solve hyperbolic non-linear systems of equations with source terms, of the form:

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = \mathbf{S}(\mathbf{U}, x) \tag{25}$$

It will be first assumed that the source term $\mathbf{S}$ can be expressed as:

$$\mathbf{S}(\mathbf{U}) = \partial_x \mathbf{T} \tag{26}$$

so (25) becomes:

$$\partial_t \mathbf{U} + \vec{\nabla}(\mathbf{F}(\mathbf{U}) - \mathbf{T}(\mathbf{U})) = 0 \tag{27}$$

The mathematical properties of the hyperbolic system of equations include the existence of a Jacobian matrix, $\mathbf{J}$, of the flux $\mathbf{F}$ defined as

$$\mathbf{J} = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} \tag{28}$$

From its eigenvectors, two matrices $\mathbf{P}$ and $\mathbf{P}^{-1}$ can be constructed with the property that they diagonalize the Jacobian $\mathbf{J}$,

$$\mathbf{J} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}^{-1} \tag{29}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix with eigenvalues in the main diagonal.

The equivalent to (3) for the system is

$$\frac{\partial}{\partial t}\int_{\Omega} \mathbf{U} d\Omega + \int_{\Omega} \nabla(\mathbf{F} - \mathbf{T}) d\Omega = 0 \tag{30}$$

If Gauss's theorem is applied to the second integral in (30):

$$\frac{\partial}{\partial t}\int_{\Omega} \mathbf{U} d\Omega + \oint_{\partial\Omega} (\mathbf{F} - \mathbf{T}) \mathrm{d}l = 0 \tag{31}$$

Now, a cell-centred finite volume method can be formulated where all the dependent variables of the system are represented as piecewise constants (first order). Therefore the vector quantities $\mathbf{U}$, $\mathbf{F}$ and $\mathbf{T}$ are uniform per cell of length Δx. In particular, the first integral in (31) can be approximated by:

$$\frac{\partial}{\partial t}\int_{\Omega} \mathbf{U}(x, y) d\Omega = \frac{\partial \mathbf{U}_i}{\partial t}\Delta x \cong \frac{\mathbf{U}_i^{n+1} - \mathbf{U}_i^n}{\Delta t}\Delta x \tag{32}$$

In the second integral of (31):

$$\oint_{\partial\Omega} \mathbf{F} = \mathbf{F}_{i+1} - \mathbf{F}_{i-1} = \delta\mathbf{F}_{i+1/2} + \delta\mathbf{F}_{i-1/2} \tag{33}$$

with $\delta\mathbf{F}_{i+1/2} = \mathbf{F}(\mathbf{U}_{i+1}) - \mathbf{F}(\mathbf{U}_i) = \mathbf{F}_{i+1} - \mathbf{F}_i$.

Due to the non-linear character of the flux $\mathbf{F}$, the definition of an approximated flux Jacobian, $\tilde{\mathbf{J}}_{i+1/2}$ (Roe, 1986) allows for a local linearization and is exploited here. Roe suggested that the following conditions should be imposed on the approximate matrix $\tilde{\mathbf{J}}_{i+1/2}$:

$$\begin{aligned}
&\text{(i) } \tilde{\mathbf{J}}_{i+1/2} = \tilde{\mathbf{J}}_{i+1/2}(\mathbf{U}_i, \mathbf{U}_{i+1}) \\
&\text{(ii) } \mathbf{F}_{i+1} - \mathbf{F}_i = \tilde{\mathbf{J}}_{i+1/2}(\mathbf{U}_{i+1}. - (\mathbf{U}_i) \\
&\text{(iii) } \tilde{\mathbf{J}}_{i+1/2} \text{ has real eigenvalues and a set of eigenvectors} \\
&\text{(iv) } \tilde{\mathbf{J}}_{i+1/2} = \tilde{\mathbf{J}}_{i+1/2}(\mathbf{U}_{i+1}) = \tilde{\mathbf{J}}_{i+1/2}(\mathbf{U}_i) \text{ if } \mathbf{U}_{i+1} = \mathbf{U}_i
\end{aligned} \tag{34}$$

Matrices $\tilde{\mathbf{P}}^{-1}$ and $\tilde{\mathbf{P}}$ can be built so that they diagonalize the approximate Jacobian matrix $\tilde{\mathbf{J}}_{i+1/2}$

$$\tilde{\mathbf{J}}_{i+1/2} = (\tilde{\mathbf{P}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{P}}^{-1})_{i+1/2}, \quad \tilde{\mathbf{P}}_{i+1/2} = [\tilde{\mathbf{e}}^1, \ldots, \tilde{\mathbf{e}}^{N\lambda}]_{i+1/2} \tag{35}$$

and $\tilde{\mathbf{\Lambda}}$ is the diagonal eigenvalues matrix. From the approximate Jacobian (Roe, 1986)

$$\tilde{\mathbf{J}}_{i+1/2}\tilde{\mathbf{e}}^m_{i+1/2} = (\tilde{\lambda}\tilde{\mathbf{e}})_{i+1/2} \quad m = 1, \ldots, N\lambda \tag{36}$$

where $N\lambda$ is the number of eigenvalues, $\tilde{\lambda}^m$. Following a flux difference procedure, the difference in vector $\mathbf{U}$ across the grid edge is projected onto the matrix eigenvectors basis

$$\delta\mathbf{U}_{i+1/2} = \sum_{m=1}^{N\lambda} (\alpha\tilde{\mathbf{e}})^m_{i+1/2} \tag{37}$$

Now the contributions in (33) are written as:

$$\tilde{\mathbf{J}}_{i-1/2}\delta\mathbf{U}_{i-1/2} + \tilde{\mathbf{J}}_{i+1/2}\delta\mathbf{U}_{i+1/2} = \sum_{m=1}^{N_\lambda} (\tilde{\lambda}^m\alpha^m\tilde{\mathbf{e}}^m)_{i+1/2} + \sum_{m=1}^{N_\lambda} (\tilde{\lambda}^m\alpha^m\tilde{\mathbf{e}}^m)_{i+1/2} \tag{38}$$

Since all the variables defined at the cell are uniform, the term $\mathbf{T}$ in the second integral in (31), is approximated by

$$\oint_{\partial\Omega} \mathbf{T}\mathbf{n}\mathrm{d}l = \mathbf{T}_{i+1} - \mathbf{T}_{i-1} = \delta\mathbf{T}_{i+1/2} + \delta\mathbf{T}_{i-1/2} \tag{39}$$

The normal source difference $\delta\mathbf{T}_{i+1/2}$ can also be expressed in function of the eigenvalues and eigenvectors of $\tilde{\mathbf{J}}_{i+1/2}$, using the approximate matrix $\tilde{\mathbf{P}}_{i+1/2}$ in order to reach a unified formulation:

$$\delta\mathbf{T}_{i+1/2} = \sum_{m=1}^{N_\lambda} (\beta\tilde{\mathbf{e}}^m)^m_{i+1/2} \tag{40}$$

where the different β^m coefficients are computed as

$$\boldsymbol{\beta}_{i+1/2} = \tilde{\mathbf{P}}^{-1}_{i+1/2}\delta\mathbf{T}_{i+1/2} \tag{41}$$

with $\boldsymbol{\beta}_{i+1/2} = [\beta^1, \ldots, \beta^m]^T_{i+1/2}$.

In order to discriminate the sense of advection linked to the sign of the different eigenvalues, two matrices $\tilde{\mathbf{\Lambda}}^{\pm}$ are defined:

$$\tilde{\mathbf{\Lambda}}^{\pm} = (\tilde{\mathbf{\Lambda}} \pm |\tilde{\mathbf{\Lambda}}|)/2 \tag{42}$$

The flux difference across each edge $i+1/2$ is split into contributions directed to cell $i+1$ (positive waves) and contributions directed to cell i (negative waves):

$$\tilde{\mathbf{J}}_{i+1/2}\delta\mathbf{U}_{i+1/2} = (\tilde{\mathbf{P}}\tilde{\mathbf{\Lambda}}^{-}\tilde{\mathbf{P}}^{-1})_{i+1/2}\delta\mathbf{U}_{i+1/2} + (\tilde{\mathbf{P}}\tilde{\mathbf{\Lambda}}^{+}\tilde{\mathbf{P}}^{-1})_{i+1/2}\delta\mathbf{U}_{i+1/2} \tag{43}$$

For the updating algorithm, as defined for a given cell i, only the contributions directed to cell i generated at the edges $i+1/2$ and $i-1/2$ are of interest. The contour integral of the numerical normal flux is equivalent to the sum of the in-going waves:

$$\tilde{\mathbf{J}}_{i-1/2}\delta\mathbf{U}_{i-1/2} + \tilde{\mathbf{J}}_{i+1/2}\delta\mathbf{U}_{i+1/2} \cong \sum_{m=1}^{N_\lambda}(\tilde{\lambda}^{+m}\alpha^m\tilde{\mathbf{e}}^m)_{i-1/2} + \sum_{m=1}^{N_\lambda}(\tilde{\lambda}^{-m}\alpha^m\tilde{\mathbf{e}}^m)_{i+1/2} \tag{44}$$

where $\lambda^- = \frac{1}{2}(\lambda - |\lambda|)$.

In order to enforce an exact equilibrium in steady state cases the normal source difference $\delta\mathbf{T}_{i+1/2}$ is also split in two kinds of waves:

$$\delta\mathbf{T}_{i+1/2} = (\delta\mathbf{T}\mathbf{n})^{+}_{i+1/2} + (\delta\mathbf{T}\mathbf{n})^{-}_{i+1/2} \tag{45}$$

with the same philosophy as before and where

$$\delta\mathbf{T}^{\pm}_{i+1/2} = \sum_{m=1}^{N_\lambda}(\beta\tilde{\mathbf{e}})^{\pm m}_{i+1/2} \tag{46}$$

with $\beta^{\pm m} = \frac{1}{2}(1 \pm \mathrm{sgn}(\tilde{\lambda}^m))\beta^m$.

The first order upwind scheme for the system gets the unified form:

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{\Delta x}\left(\sum_{m=1}^{N_\lambda}((\tilde{\lambda}^{+}\alpha - \beta^{+})\tilde{\mathbf{e}})^m_{i-1/2} + \sum_{m=1}^{N_\lambda}((\tilde{\lambda}^{-}\alpha - \beta^{-})\tilde{\mathbf{e}})^m_{i+1/2}\right) \tag{47}$$

This is a compact form with the focus on the waves generated at the cell edges, made of both the normal flux difference and the normal source term, and governed by the sign of the eigenvalues of the normal flux Jacobian.

As was done in the scalar case, the numerical scheme in (47) can be rewritten as:

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \left[\sum_{m=1}^{N\lambda}\nu^{m,+}_{i-/2}\delta\mathbf{U}_{i-/2} + \sum_{m=1}^{N\lambda}\nu^{m,-}_{i+/2}\delta\mathbf{U}_{i+/2}\right], \quad \nu^{m,\pm}_{i\mp1/2} = \lambda^{m*,\pm}_{i\mp1/2}\frac{\Delta t}{\Delta x} \tag{48}$$

with

$$\lambda^{m*,\pm}_{i\mp1/2} = \lambda^{m\pm}_{i\mp1/2}\,\theta^{m\pm}_{i\mp1/2} \quad \theta^{m\pm}_{i\mp1/2} = 1 - \left(\frac{\beta^-}{\alpha\tilde{\lambda}^-}\right)^{\pm}_{i\mp1/2} \tag{49}$$

where the ratio $\beta/(\alpha\tilde{\lambda})$ expresses the influence of the source terms over that of the flux differences.

In absence of source terms, the numerical scheme (48) is stable provided that

$$-1 \le \nu^{m,-}_{i+1/2} \le 0 \quad 0 \le \nu^{m.+}_{i-1/2} \le 1 \quad m = 1,\ldots,N\lambda \tag{50}$$

and the following condition over the conserved variables applies:

$$U_s^{\min} \le U_{s,i}^{n+1} \le U_s^{\max} \quad \mathbf{U} = (U_1,\ldots,U_s,\ldots,U_{nc})^T \tag{51}$$

where $U_S^{\max} = \max\{U^n_{S,i}, U^n_{S,i+1}, U^n_{S,i-1}\}$ and $U_S^{\min} = \min\{U^n_{S,i}, U^n_{S,i+1}, U^n_{S,i-1}\}$.

Monotonicity in the conserved variables in presence of source terms requires that for all m

$$\tilde{\lambda}^{m-,*}_{i+1/2} \leq 0 \quad \tilde{\lambda}^{m+,*}_{i-1/2} \geq 0 \quad m = 1, \ldots, N\lambda \tag{52}$$

which means that

$$\theta^{m,\pm}_{i\mp 1/2} \geq 0 \quad m = 1, \ldots, N\lambda \tag{53}$$

Under these assumptions the numerical stability when $\theta^{m,\pm}_{i\mp 1/2} \geq 0$ for all m, is provided by the intersection of the stability regions defined for each celerity $\tilde{\lambda}^{m,*}_{i+1/2}$

$$\begin{aligned} \Delta t &= CFL\, \Delta t_{\max}, \quad CFL \leq 1 \\ \Delta t_{\max} &= \min\{\Delta t_{i+1/2}\}_{k=0,N} \quad \Delta t_{i+1/2} = \frac{\Delta x}{\max\limits_{m}\{|\lambda^{m,*}_{i+1/2}|\}}, \quad \theta^{m,\pm}_{i+1/2} \geq 0 \quad m = 1, \ldots, N\lambda \end{aligned} \tag{54}$$

In the particular case of $\theta^{m,\pm}_{i+1/2} = 1$ for all m, (54) expresses the stability condition without source terms (CFL condition).

Equilibrium in steady state cases is ensured if the discretization of the source term has been constructed enforcing

$$(\tilde{\lambda}\alpha - \beta)^{m\pm}_{i\mp 1/2} = 0 \quad m = 1, \ldots, N\lambda \tag{55}$$

which is equivalent to $\lambda^{m\pm}_{i\mp 1/2} = 0$, leading to an unconditionally stable scheme in this particular case.

The scheme in (47) can also be expressed

$$\mathbf{U}^{n+1}_i = \mathbf{U}^n_i - \left[\sum_{m=1}^{N\lambda} \tilde{\lambda}^{m,+}_{i-/2}\delta\mathbf{D}_{i-/2} + \sum_{m=1}^{N\lambda} \tilde{\lambda}^{m,-}_{i+/2}\delta\mathbf{D}_{i+/2}\right]\frac{\Delta t}{\Delta x}, \quad \delta\mathbf{D}^m_{i\pm 1/2} = \theta^m_{i\pm 1/2}\delta\mathbf{U}^m_{i\pm 1/2} \tag{56}$$

which is useful to see that the stability region for (56) with source terms is enlarged if $0 \leq \theta^m_{i\pm 1/2} \leq 1$ for all m, as in this particular case $|\delta\mathbf{D}^m_k| \leq |\delta\mathbf{U}^m_k|$. Equilibrium is achieved when all $\delta\mathbf{D}^m_k = 0$, condition that can be automatically derived from (55).

On the other hand, as seen in the scalar case, when $\theta^m_{i\pm 1/2} < 0$ the source terms dominate over the flux differences. If (51) is still desired, the definition of γ has to be based on the specific necessities of the physical problem. A direct extension of (24) is not possible, as it is not feasible to define a $\gamma^m_{i+1/2}$ coefficient if $\theta^m_{i\pm 1/2} < 0$ for all m, since there is not a correspondence between the m-waves and the s-variables. In the special case of trying to preserve the sign of the solution in the s component, expressed as

$$U^{n+1}_{s,i} \leq 0 \text{ with } U^n_{s,i}, U^n_{s,i\pm 1} \leq 0 \quad \text{or} \quad U^{n+1}_{s,i} \geq 0 \text{ with } U^n_{s,i}, U^n_{s,i\pm 1} \geq 0 \tag{57}$$

the time step in the stability region must be computed following

$$\Delta t_{i+1/2} = \gamma\frac{\Delta x}{\max\limits_{m}\{|\lambda^{*,m}_{i+/2}|\}}, \quad \gamma = \frac{\min\{|U_{s,i}|, |U_{s,i+1}|, |\delta U_{s,i+1/2}|\}}{|\delta U_{s,i+1/2}|} \tag{58}$$

where $0 \leq \gamma \leq 1$. If the conserved variable is smooth or gradually varied, the coefficient γ is 1 and the time step in (58) reduces to (54).

In the particular case $\gamma = 0$ (58) predicts a zero time step. Actually, this means that no flux can cross the associated edge $i + 1/2$. When $\gamma = 0$ it is necessary to modify the eigenvalues as follows

$$\tilde{\lambda}^{m+,*}_{i+1/2} = \begin{cases} 0 & \textit{if } \tilde{\lambda}^{m+,*}_{i+1/2} < 0 \\ \tilde{\lambda}^{m+,*}_{i+1/2} & \textit{otherwise} \end{cases} \quad \tilde{\lambda}^{m-,*}_{i+1/2} = \begin{cases} 0 & \textit{if } \tilde{\lambda}^{m-,*}_{i+1/2} > 0 \\ \tilde{\lambda}^{m-,*}_{i+1/2} & \textit{otherwise} \end{cases} \tag{59}$$

When $\gamma << 1$ condition (58) forces to strong restrictions in the size of the time step and therefore unacceptable computational costs. This reduction can be avoided by means of a conservative strategy based on the redistribution of updating fluxes, involving a local time step $\Delta t_{Us,i}$, that replaces the γ coefficient in (58). The local time step $\Delta t_{Us,i}$ is defined in order to preserve the sign of the variable $U_{s,i}$. Expressing for each variable (50) as

$$U_{s,i}^{n+1} = U_{s,i}^{n} + \Delta t(\Psi_{i,i-1/2}^{Us} + \Psi_{i,i+1/2}^{Us}), \quad \Psi_{i,i\pm1/2}^{Us} = \frac{1}{\Delta x}\sum_{m=1}^{3}((\tilde{\lambda}^{\mp}\alpha - \beta^{\mp})e_1)_{i,i\pm1/2}^{m} \tag{60}$$

In the case of requiring the positivity of the variable $U_{s,i}$ $\Delta t_{Us,i}$ is computed as:

$$\Delta t_{Us,i} = -\frac{U_{s,i}^{n}}{\Psi_{i,i-1/2,k}^{Us} + \Psi_{i,i+1/2,k}^{Us}}, \quad \Psi_{i,i\pm1/2}^{Us} < 0 \tag{61}$$

It is worth noting that the quantity defined in (61) is always positive due to the condition on the contributions in the denominator. In the case of requiring the positivity of more variables more $\Delta t_{Us,i}$'s local time steps are defined. Then a redistribution factor $r_{t,i}$ is computed as:

$$r_{t,i} = \frac{\Delta t_v}{\Delta t_{\max}} \leq 1, \quad \Delta t_v = \min\{\Delta t_{\max}, \Delta t_{Us,i}\} \tag{62}$$

where $\Delta t_{\max}$ is computed as in (79). When in the cell i the ratio $r_{t,i} < 1$, we propose to redefine the updating fluxes-at those edges where $\Psi_{i,i\pm1/2}^{Us} \leq 0$ for one U_s according to:

$$\Psi_{i,i\pm1/2}^{Us} = \omega_{Us1}, \quad \Psi_{i\pm1,i\pm1/2}^{Us} = \Psi_{i\pm1,i\pm1/2}^{Us} + \omega_{Us2} \tag{63}$$

where, to preserve conservation, the quantities ω_{h1}, ω_{h2}, ω_{z1}, ω_{z2} are

$$\omega_{Us1} = \Psi_{i,i\pm1/2}^{Us} r_{t,i} \quad \omega_{Us2} = \Psi_{i,i\pm1/2}^{Us}(1 - r_{t,i}) \tag{64}$$

By means of this technique the variables are updated by computing the time step as in the homogeneous case. In some cases negative values can be obtained, as the truly time step is various orders of magnitude smaller. In that case it is enough to set $\Delta t_{\max} = \frac{1}{2}\Delta t_{\max}$ and to repeat the process until no negative values appear. This reduction has an irrelevant computational cost, generating always a solution.

4 APPLICATION TO THE 1D SHALLOW WATER EQUATIONS

4.1 *1D mathematical model of the shallow water equations*

In this work the schemes are used to solve the following system of equations:

$$\mathbf{U} = (h,\, q_x)^T \quad \mathbf{F} = \left(q,\, \frac{q^2}{h} + \frac{gh^2}{2}\right)^T \quad \mathbf{S} = \left(0, -gh\frac{\partial z}{\partial x} - \frac{\tau}{\rho_w}\right)^T \tag{65}$$

where h is the water depth, g is the acceleration of the gravity, $q = uh$ the unit discharge with u the averaged velocity. The bed shear-stress τ can be modelled according to the Chézy equation for open-channel hydraulics (Chow, 1959).

4.2 *Application of the explicit upwind numerical scheme*

For the Saint-Venant equations (Vázquez-Cendón, 1999), an approximate flux Jacobian matrix $\tilde{\mathbf{J}}$ has to be defined. The approximate matrix in our case is

$$\tilde{\mathbf{J}} = \frac{\delta \mathbf{F}}{\delta \mathbf{W}} = \begin{pmatrix} 0 & 1 \\ g\tilde{h} - \tilde{u}^2 & 2\tilde{u} \end{pmatrix} \tag{66}$$

The eigenvalues and eigenvectors of (66) are:

$$\begin{aligned} \tilde{\lambda}^1_{i+1/2} &= (u + \tilde{c})_{i,k} \quad \tilde{\lambda}^2_{i+1/2} = (u - \tilde{c})_{i,k} \\ \tilde{\mathbf{e}}^1_{i,k} &= \begin{pmatrix} 1 \\ \tilde{u} + \tilde{c} \end{pmatrix}_{i,k}, \quad \tilde{\mathbf{e}}^2_{i,k} = \begin{pmatrix} 1 \\ \tilde{u} - \tilde{c} \end{pmatrix}_{i,k} \end{aligned} \tag{67}$$

where

$$\tilde{u}_{i+1/2} = \frac{u_i\sqrt{h_i} + u_{i+1}\sqrt{h_{i+1}}}{\sqrt{h_i} + \sqrt{h_{i+1}}}, \quad \tilde{c}_{i+1/2} = \sqrt{g\frac{h_i + h_{i+1}}{2}} \tag{68}$$

From the difference in vector $\mathbf{U}$ across the grid edge the coefficients $\alpha^m_{i,+1/2}$ are:

$$\alpha^{1,2}_{i+1/2} = \frac{\delta h_{i+1/2}}{2} \pm \frac{1}{2\tilde{c}_{i+1/2}}(\delta q - \tilde{u}\delta h)_{i+1/2} \tag{69}$$

The non conservative term $\mathbf{S}$ is included defining $\delta \mathbf{T}_{i+1/2}$, with

$$\delta \mathbf{T}_{i+1/2} = \begin{pmatrix} 0 \\ -g\tilde{h}\,\delta z - \dfrac{\tilde{\tau}}{\rho_w}\Delta x \end{pmatrix}_{i+1/2} \tag{70}$$

Here it is worth remarking that if the matrix $\mathbf{S}$ were constant the equality $\mathbf{S}(\mathbf{U}) = \partial_x \mathbf{T}$ would hold. Obviously, this is not our case, but the linearization of the non conservative term can be considered locally for numerical approximation. Now, following (41) the coefficients β^m are defined as:

$$\beta^{1,2} = \mp\frac{1}{2}\left(\tilde{c}\,\delta z + \frac{\tilde{\tau}\Delta x}{\tilde{c}\rho_w}\right)_{i+1/2} \tag{71}$$

4.3 *Conservation properties and equilibrium at steady state*

The unified discretization of the source terms is successfully constructed when it ensures an exact balance in first order (Bermudez and Vazquez-Cendon 1998). For that reason, the discretization of the fluxes and source terms proposed in §3.1 is here analysed. Steady state is expressed as:

$$\tilde{\mathbf{J}}_{i+1/2}\delta \mathbf{U}_{i+1/2} = \delta \tilde{\mathbf{T}}_{i+1/2} \tag{72}$$

According to the form of the matrices involved, the first equation in (72) gives

$$\delta(hu)_{i+1/2} = \delta(q)_{i+1/2} = 0 \tag{73}$$

which means that the normal discharge is constant at the edge. The second line in (72) yields

$$\delta h_{i+1/2}(\tilde{c}^2 - \tilde{u}^2)_{i+1/2} + \delta q_{i+1/2}(2\tilde{u})_{i+1/2} = -(g\tilde{h}\delta z)_{i+1/2} - \left(\frac{\tilde{\tau}\Delta x}{\rho_w}\right)_{i+1/2} \tag{74}$$

Considering (74) and expressing the friction term in function of the energy slope $S_{f_{i+1/2}}$,

$$\left(\frac{\tilde{\tau}}{\rho_w}\right)_{i+1/2} = (g\tilde{h}S_f)_{i+1/2} \tag{75}$$

the expression for the water depth profile for steady flow appears:

$$\delta h_{i+1/2}(1 - Fr^2_{i+1/2}) + \delta z_{i+1/2} = -\Delta x S_{i+1/2} \tag{76}$$

where $Fr_{i+1/2}$ is the Roe average Froude number, $Fr_{i+1/2} = \tilde{u}_{i+1/2}/\tilde{c}_{i+1/2}$. Equation (74) can be derived directly enforcing

$$\delta \mathbf{D}^m_{i+1/2} = 0 \tag{77}$$

or $(\tilde{\lambda}^- \alpha - \beta^-)^m_{i+1/2} = 0$ for $m = 1$ and $m = 2$. Therefore the discretization ensures equilibrium at steady state correctly for first order, and the C-Property is satisfied not only in the case of quiescence flow over discontinuous bed topography, but also in the case of steady fluid in motion.

4.4 *Interpretation of the θ coefficient*

In the homogenous case, without source terms, $\theta^{\pm}_{i+1/2} = 1$, the original stability region is automatically recovered. The case $\theta^{\pm}_{i+1/2} = 0$, corresponds to the equilibrium case (C-Property is satisfied) and the numerical scheme becomes unconditionally stable. On the other hand, when $\theta^{\pm}_{i+1/2} < 0$ the positivity requirements over the variables are satisfied if the stability region is defined using:

$$\gamma = \frac{\min\{h_i, h_{i+1}, |\delta h_{i+1/2}|\}}{|\delta h_{i+1/2}|}, \quad \theta^{\pm}_{i+1/2} < 0 \tag{78}$$

Condition (78) is of special relevance when dealing with wetting and drying fronts. In the special case, $h_{i+1} = 0$ or $h_i = 0$, γ is nil and it is necessary to modify the eigenvalues as in (59) otherwise the positivity condition over the depth is not ensured. As no flux can cross the associated edge $i + 1/2$ the following condition must be applied to the solution at the cells sharing edge $i + 1/2$:

$$u^{n+1}_i = u^{n+1}_{i+1} = 0 \tag{79}$$

To avoid strong restrictions in the magnitude of the time step when $\gamma << 1$ the redistribution factor is computed as:

$$r_{t,i} = \frac{\Delta t_v}{\Delta t_{\max}} \leq 1, \quad \Delta t_v = \min\{\Delta t_{\max}, \Delta t_{h,i}\} \tag{80}$$

where $\Delta t_{h,i}$ is computed following (61).

Also, to preserve conservation it is necessary to impose (79) in the edges where $\Psi^h_{i\pm1/2} \leq 0$ and $r_{t,i} \leq 1$. The redefinition of the fluxes according to (63) is done at the cell edges where $\Psi^h_{i,i\pm1/2} \leq 0$.

Near wetting/drying fronts, characterized by small values of water depth, the bed friction term may dominate over any other term, leading to numerical instabilities. When $0 < \gamma < 1$ numerical instabilities are avoided by requiring that friction alone is not able to change the sign of the discharge, so the following conditions are enforced over the unit discharge function, *hu*

$$(hu)^{n+1}_i \geq 0 \quad (hu)^n_i, (hu)^n_{i\pm1} \geq 0 \quad \text{or} \quad (hu)^{n+1}_i \leq 0 \quad (hu)^n_i, (hu)^n_{i\pm1} \leq 0 \tag{81}$$

These conditions must be enforced to determine the maximum allowable time step, otherwise the numerical scheme leads to numerical instabilities. Let us assume without lost of generality the one-dimensional case, where the updated value can be expressed as:

$$(hu)^{n+1}_i = (hu)^n_i - (g\tilde{h}S_f \Delta x)^n_k \frac{\Delta t}{\Delta x} \tag{82}$$

that can be rewritten as

$$(hu)_i^{n+1} = (hu)_i^n \left[1 - \frac{(g\tilde{h}S_f)_k^n}{(hu)_i^n} \Delta t \right] \tag{83}$$

The second term on the right hand side of (91) must be positive to ensure (89). Hence, in general, the time step Δt_k, taking into account also condition (86), is limited by

$$\Delta t_{i+1/2} = \min \left\{ \left| \frac{\tilde{u}_{i+1/2}}{gS_{f,i+1/2}} \right|, \frac{\Delta x}{\max\limits_m \{|\tilde{\lambda}^m_{i+1/2}|\}} \right\}_{i+1/2} \tag{84}$$

to prevent instabilities.

5 APPLICATION TO THE 1D SHALLOW WATER EQUATIONS WITH SOLUTE TRANSPORT

5.1 *1D mathematical model of the shallow water equations with solute transport*

In this work the schemes are used to solve the following system of equations:

$$\mathbf{U} = (h,\, q_x, h\phi)^T, \quad \mathbf{F} = \left(q, \frac{q^2}{h} + \frac{gh^2}{2}, q\phi \right)^T \quad \mathbf{S} = \left(0, -gh\frac{\partial z}{\partial x} - \frac{\tau}{\rho_w}, 0 \right)^T \tag{85}$$

where ϕ is the depth averaged solute concentration.

5.2 *Application of the explicit upwind numerical scheme*

The eigenvalues and eigenvectors are in this case:

$$\begin{aligned} \tilde{\lambda}^1_{i+1/2} &= (\tilde{u} + \tilde{c})_{i+1/2} \quad \tilde{\lambda}^2_{i+1/2} = (\tilde{u} - \tilde{c})_{i+1/2} \quad \tilde{\lambda}^3_{i+1/2} = \tilde{u}_{i+1/2} \\ \tilde{\mathbf{e}}^1_{i,k} &= \begin{pmatrix} 1 \\ \tilde{u} + \tilde{c} \end{pmatrix}_{i,k}, \quad \tilde{\mathbf{e}}^2_{i,k} = \begin{pmatrix} 1 \\ \tilde{u} - \tilde{c} \end{pmatrix}_{i,k} \quad \tilde{\mathbf{e}}^3_{i,k} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{i,k} \end{aligned} \tag{86}$$

where

$$\tilde{\phi}_{i+1/2} = \frac{\tilde{\phi}_i \sqrt{h_i} + \tilde{\phi}_{i+1}\sqrt{h_{i+1}}}{\sqrt{h_i} + \sqrt{h_{i+1}}} \tag{87}$$

The coefficients $\alpha^m_{i,+1/2}$ are:

$$\alpha^{1,2}_{i+1/2} = \frac{\delta h_{i+1/2}}{2} \pm \frac{1}{2\tilde{c}_{i+1/2}} (\delta q - \tilde{u}\delta h)_{i+1/2} \quad \alpha^3_{i+1/2} = \delta(h\phi)_{i+1/2} - \tilde{\phi}_{i+1/2}(\delta h)_{i+1/2} \tag{88}$$

Now, the coefficients β^m are

$$\beta^{1,2} = \mp \frac{1}{2} \left(\tilde{c}\, \delta z + \frac{\tilde{\tau}\Delta x}{\tilde{c}\rho_w} \right)_{i+1/2}, \quad \beta^3 = 0 \tag{89}$$

5.3 *Interpretation of the θ coefficient*

When moving to the coupled set of water flow and solute transport equations, γ is defined with the same purpose of preserving the sign:

$$\gamma = \min\{\gamma_h, \gamma_{h\phi}\}$$
$$\gamma_h = \frac{\min\{h_i, h_{i+1}, |\delta h_{i+1/2}|\}}{|\delta h_{i+1/2}|}, \quad \gamma_{h\phi} = \frac{\min\{(h\phi)_i, (h\phi)_{i+1}, |\delta(h\phi)_{i+1/2}|\}}{|\delta(h\phi)_{i+1/2}|} \quad \theta^{\pm}_{i+1/2} < 0 \quad (90)$$

In presence of solute fronts, that is clean/mixed water boundaries with continuous water level surface, this leads to unrealistic results in the solute advance. If this fact is not considered, negative values of solute mass and concentration can be obtained and the necessity to tune and alter the scheme results becomes necessary. This can be avoided by using again a conservative redistribution of the updating contributions (Murillo et al. 2006) that ensures adequate bounding properties over the solute concentration.

Now the redistribution factor is computed as:

$$r_{t,i} = \frac{\Delta t_v}{\Delta t_{\max}} \leq 1, \quad \Delta t_v = \min\{\Delta t_{\max}, \Delta t_{h,i}, \Delta t_{h\phi,i}\} \quad (91)$$

and we propose to redefine the updating fluxes to cells at those edges where $\Psi^h_{i,i\pm1/2} \leq 0$ according to (63) over the water depth and over the solute mass.

6 APPLICATION TO THE 1D BED LOAD SEDIMENT TRANSPORT

6.1 *1D mathematical model for sediment transport*

The mathematical model used in this work is based in the differential form of the equations described by Hudson involving the friction term. In this model the solid phase has a velocity equal to the liquid one and presents a uniform distribution over the flow depth. No exchanges through the bottom interface are considered. The system of equations is

$$\mathbf{U} = \begin{pmatrix} h \\ uh \\ z \end{pmatrix}, \mathbf{F} = \begin{pmatrix} hu \\ (u^2 h + \frac{1}{2} g h^2) \\ \xi q_l \end{pmatrix}, \mathbf{S} = \begin{pmatrix} 0 \\ -gh\dfrac{\partial z}{\partial x} - \dfrac{\tau}{\rho_w} \\ 0 \end{pmatrix} \quad (92)$$

Where h is the water depth, u is the depth averaged velocity, z is the bottom elevation, $\xi = 1/(1-\varepsilon)$ where ε is the porosity, g is gravity, ρ_w is the density of the water, τ is the bed shear-stress, and q_l is the total volumetric sediment rate. The sediment transport flux is defined using the Grass formulation (Grass, 2000)

$$q_l = A\, u(u^2 + v^2)^{1/2(m-1)} \quad (93)$$

where A is a dimensional constant usually determined by experimental data and $1 \leq m \leq 4$ (Hudson, 2000) In this work m is set constant and equal to 3. It is important to note that to avoid the singularity in the Jacobian matrix of the flux, the flux $\mathbf{F}$ and the slope term $\mathbf{S}_1$ are modified as follows (Hudson, 2000)

$$\mathbf{F} = \begin{pmatrix} hu \\ u^2 h + \frac{1}{2} g h^2 + ghz \\ q_l \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} 0 \\ gz\dfrac{\partial h}{\partial x} - \dfrac{\tau}{\rho_w} \\ 0 \end{pmatrix} \quad (94)$$

leading to the following Jacobian matrix

$$\mathbf{J} = \frac{\partial \mathbf{F}}{\partial \mathbf{W}} = \begin{pmatrix} 0 & 1 & 0 \\ g(h+z) - u^2 & 2u & gh \\ -ud & d & 0 \end{pmatrix} \tag{95}$$

with $d_1 = 3A\xi u^2 h^{-1}$.

6.2 *Application of the explicit upwind numerical scheme*

Hudson defined an approximate Jacobian matrix where the averaged eigenvalues where computed solving the following third order polynomial,

$$P(\tilde{\lambda}) = \tilde{\lambda}^3 - 2\tilde{u}\tilde{\lambda}^2 + \tilde{u}^2 - [g(\tilde{h} + \tilde{z} + \tilde{h}\tilde{d})]\tilde{\lambda} + g\,\tilde{h}\,\tilde{u}\,\tilde{d} = 0 \tag{96}$$

where

$$\tilde{d}_{i+1/2} = c_b \beta \frac{\sqrt{h_{i+1}} + \sqrt{h_i}}{\sqrt{h_{i+1}} h_i + \sqrt{h_i} h_{i+1}} (u_{i+1}^2 + u_i^2 + u_i u_{i+1}) \tag{97}$$

leading to three real and distinct eigenvalues. Therefore the problem is strictly hyperbolic. Two eigenvalues have the same sign of the particle velocity $\tilde{u}$, and one has the opposite. We name it in the ascending order: with $\tilde{u} > 0$, $\tilde{\lambda}^1 < 0$, $\tilde{\lambda}^2 \geq 0$, $\tilde{\lambda}^3 > 0$ and $\tilde{\lambda}^3 > \tilde{\lambda}^2$, while $\tilde{\lambda}^2 = 0$ only if $\tilde{u} = 0$, the condition of fluid at rest. The right eigenvector associated to each λ^m are

$$\tilde{\mathbf{e}}_{i+1/2}^m = \begin{pmatrix} 1 \\ \tilde{\lambda}^m \\ (-g(\tilde{h} + \tilde{z}) + (\tilde{\lambda}^m - \tilde{u})^2)(g\tilde{h})^{-1} \end{pmatrix}_{i+1/2} \qquad m = 1, 2, 3 \tag{98}$$

where the coefficients $\alpha_{i+1/2}^m$ are

$$\alpha_{i+1/2}^m = \left[\frac{(\tilde{\lambda}^a \tilde{\lambda}^b + g(\tilde{h} + \tilde{z}) - \tilde{u}^2)\delta h + (2\tilde{u} - \tilde{\lambda}^a - \tilde{\lambda}^b)\delta(hu) + g\tilde{h}\,\delta z}{(\tilde{\lambda}^k - \tilde{\lambda}^a)(\tilde{\lambda}^k - \tilde{\lambda}^b)} \right]_{i+1/2} \qquad m = 1, 2, 3 \tag{99}$$

with $a \neq b \neq k$. The different β^m coefficients are computed as

$$\beta_{i+1/2}^m = \left(\frac{(2\tilde{u} - \tilde{\lambda}^a - \tilde{\lambda}^b)(g\tilde{z}\delta h - \tilde{\tau}\Delta x/\rho_w)}{(\tilde{\lambda}^m - \tilde{\lambda}^a)(\tilde{\lambda}^m - \tilde{\lambda}^b)} \right)_{i+1/2} \qquad m = 1, 2, 3 \tag{100}$$

In the case that $\tilde{u}_{i+1/2}$ becomes nil, $u_{i+1} = u_i = 0$, $\tilde{\lambda}^2$ also becomes nil and $\tilde{\mathbf{e}}^2 = \mathbf{0}$, so the linearization in(36) fails. Also, in the case that h_{i+1} or h_i are nil the average quantity $\tilde{d}_{i+1/2}$ become useless. In such cases we consider that no sediment transport occurs and the system of equations reduces to the shallow water equations.

The positivity requirement over the water depth is satisfied defining γ as in (78). To avoid strong restrictions in the magnitude of the time step when $\gamma << 1$, the redistribution factor is computed as in (80). The redefinition of the fluxes according to (63) is done at the cell edges where $\Psi_{i,i\pm1/2}^h \leq 0$ over the water depth and bottom elevation. Also the same restrictions imposed by the friction term must be observed.

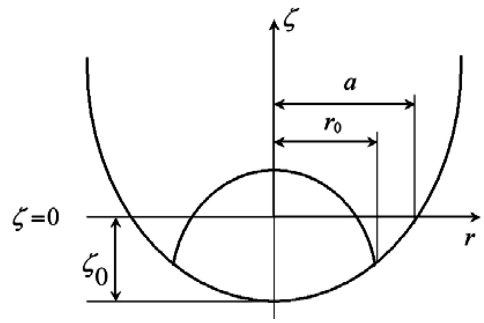


Figure 2. Initial free surface and water depth profile for the parabolic basin test.

7 APPLICATIONS

7.1 *Long wave resonance in a parabolic basin*

The analytical solution of a long wave, driven by gravity and resonating in a frictionless, dry circular parabolic basin was presented by Thacker (Thacker, 1981) for the shallow water equations, where the free surface displacement is given by

$$\zeta(r,t) = \zeta_o((1-A^2)^{1/2}B^{-1} - 1 - r^2a^{-2}\{(1-A^2)B^{-2} - 1\}) \tag{101}$$

and the bottom elevation is given as

$$z(r,t) = -\zeta_0(1 - r^2a^{-2}) \tag{102}$$

where $A = (a^4 - r_o^4)(a^4 + r_o^4)^{-1}$, $B = 1 - A\cos\omega t$, $\omega = a^{-1}\sqrt{8g\zeta_o}$, ζ_o is the center point water depth, r is the distance from the center point, a is the radial distance from the center point to the zero elevation on the shoreline and r_o is the distance from the center point to the point where the water depth is initially nil. Those values are represented in Figure 2. Numerical values are $\zeta_o = 20.0$ m, $r_o = 1200$ m, $a = 1500$ m. The domain is divided in using intervals with $\Delta x = 25$ m. The water surface movement will be simulated together with an initial solute concentration:

$$\phi(r,t_o) = \phi_o \exp(-r(2r_o)^{-1}) \tag{103}$$

There is no analytical solution for the solute concentration evolution in time but, if no diffusion is assumed, the solution for each T oscillation period will be:

$$\phi(r, t = TK) = \phi_o \exp(-r(2r_o)^{-1}), \quad K = 1, \ldots, \infty \tag{104}$$

The numerical experiment is performed using $\phi_o = 1$. This test case illustrates the concepts described in this work, as it includes wetting/drying fronts and the generation of dry regions from wet areas. In particular, the advance of the wetting/drying front is produced in the first half period, during the wave expansion, while during the wave contraction both wetting/drying fronts and drying process are present.

Figure 3 shows the water depths given by the exact solution and by the proposed method at times, $1/4T$, $2/4T$, $3/4T$, T, $3/2T$, $2T$, $5/2T$, $3T$, $7T/2$ and $4T$.. The simulated results prove in every case in good agreement with the analytical solution, including water depth and inundated area, even in the fourth oscillation. Figure 4 shows the comparison between the exact and numerical solution for the solute concentration distribution at times T, $2T$, $3T$ and $4T$. Better accuracy in the final solution can be achieved reducing the size of the cells generating a better representation of the bottom elevation.

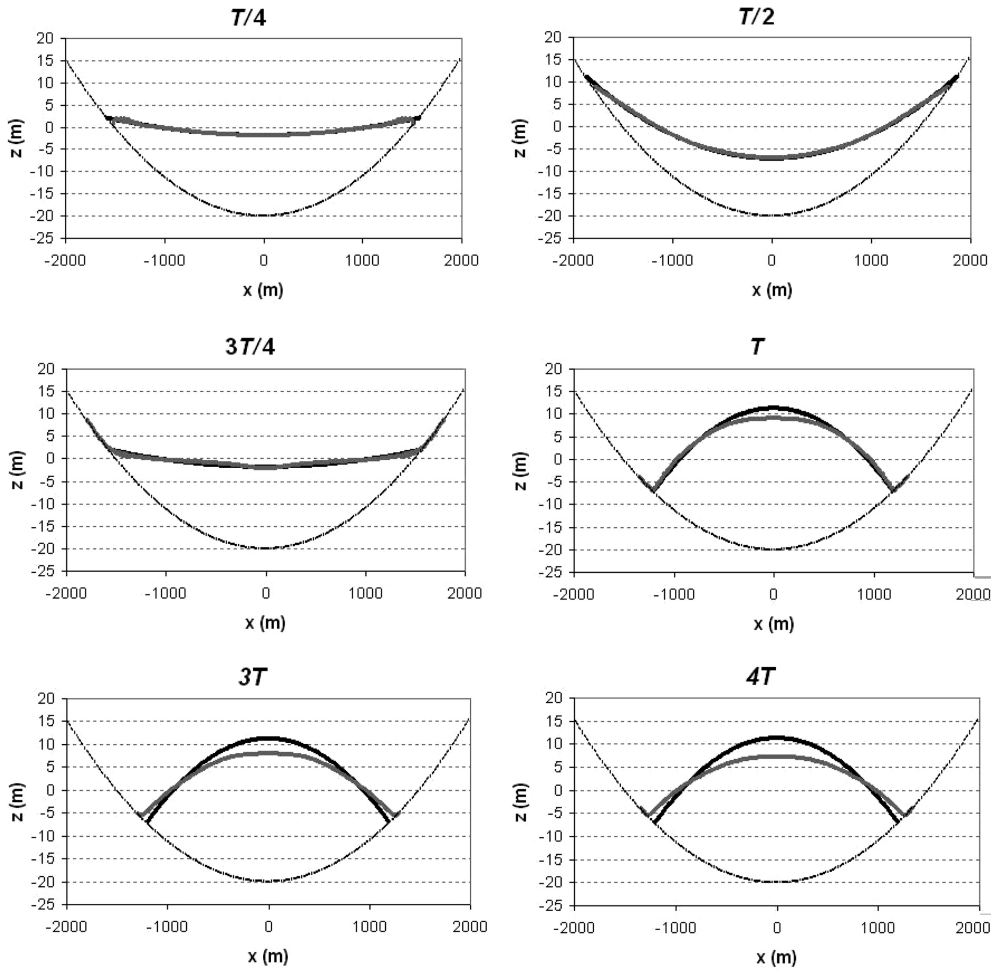


Figure 3. Water elevation surface (in meters): exact and simulated for times 1/4T, 2/4T, 3/4T, T, 3T and 4T.

7.2 *Bed evolution to equilibrium*

In the following test case the bed evolution to equilibrium in a channel departing from a situation of still water is analysed. A domain 100 meters long is divided in cells with $\Delta x = 1$ m, imposing initially $u = 0$, $z = 0$ m and $h = 2$ m. From the beginning of the simulation a constant inlet discharge is imposed upstream, $hu = 10\ \mathrm{m}^2/\mathrm{s}$, while a constant water level surface $h + z = 2$ m is imposed downstream, assuming $\xi = 0.5$ and $A = 0.05$. The shear stress is defined using the Manning equation:

$$\tau = \rho_w g n^2 u|u|h^{-1/3} \tag{105}$$

with $n = 0.04$.

Figure 5 shows the final bed and water level surface equilibrium slopes. In case of use a unified discretization for the friction term no remarkable differences appear in the bed and water surface elevations, but for the pointwise simulation the discharge varies along the domain, as Figure 6 shows.

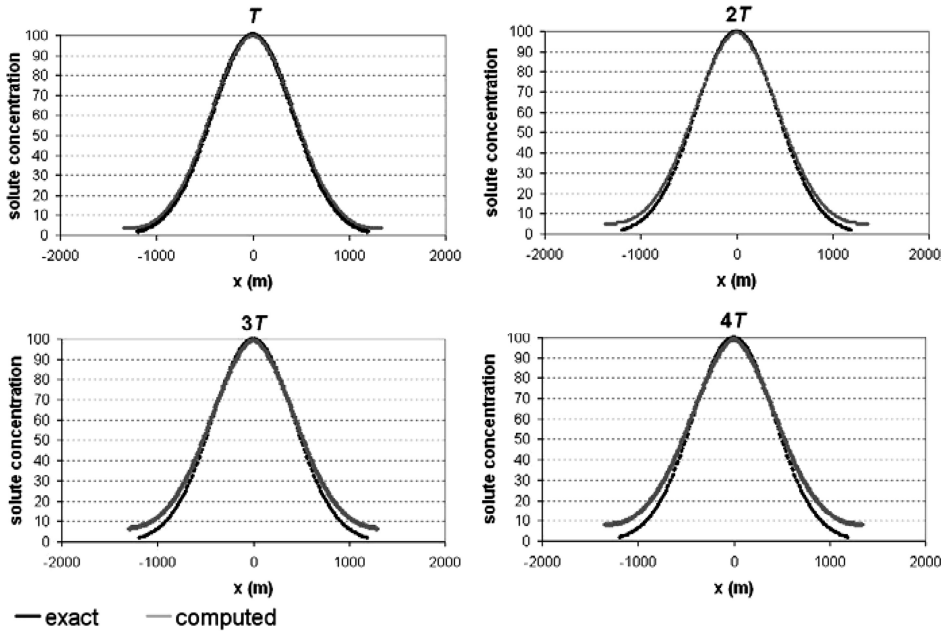


Figure 4. Solute distribution results and exact solution at times T, 2T, 3T and 4T.

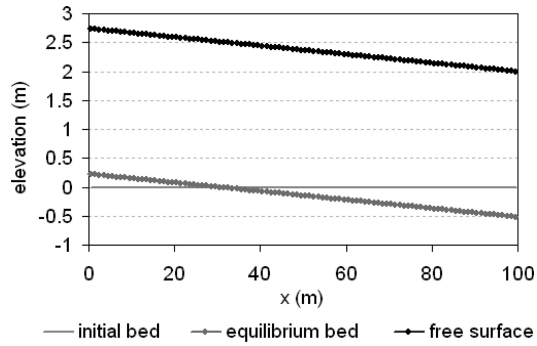


Figure 5. Final bed and water surface levels in equilibrium.

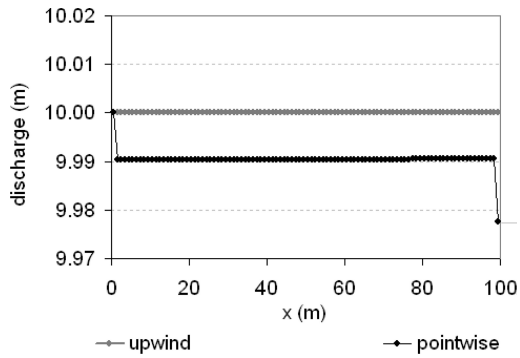


Figure 6. Equilibrium discharge using a unified and a pointwise formulation for the friction term.

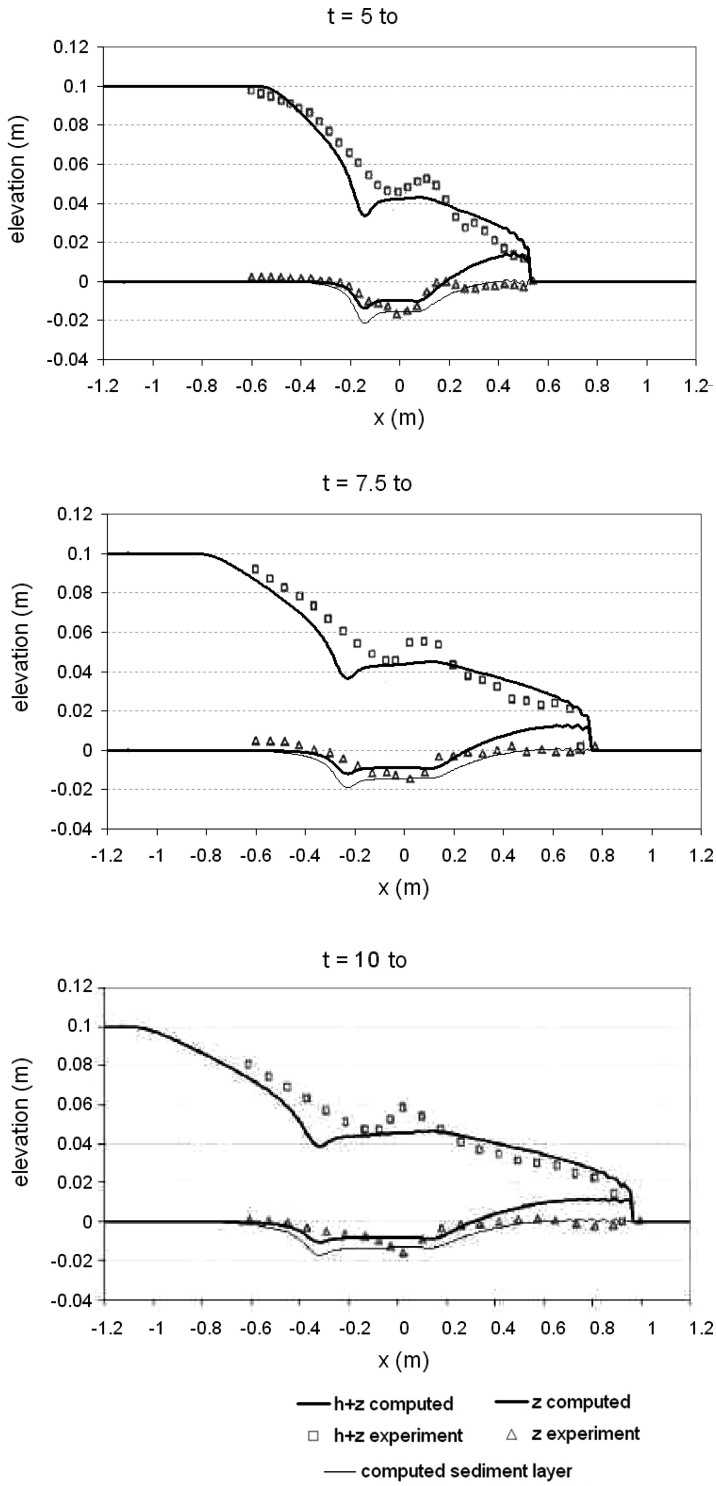


Figure 7. Experimental data and the computational solution for the Louvain test case, at times $t = 5t_o$, $7.5t_o$, $10t_o$.

7.3 *Dam break over movable bed*

The model proposed by Hudson is here applied to erodible bed dam-break problems. A body of water at rest of constant depth h_o, is released over a flat erodible bed channel extending on both sides. Two different experiments described by Capart and Fraccarollo, (Capart and Fraccarollo, 2002) are used to check the performance of the numerical scheme. The first was performed in the Université Catholique de Louvain (Louvain-la-Neuve) and the second in the National Taiwan University (Taipei) with coarse granular material of different densities. In the Louvain experiment an intermediate density for the material was chosen ($s = \rho_s/\rho_w = 1.540$) whilst in the Taipei experiment a much lighter material was chosen ($s = 1.048$), in both cases lighter than natural sediment ($s \approx 2.65$). The initial water depth in both cases is $h_o = 0.01$ m. It is possible to define the sediment load as in (Rosatti and Fraccarollo, 2006, Summer et al., 1996)

$$q_l = c_b \psi u^3 \tag{106}$$

where c_b is the volumetric concentration of sediments in the bed ($c_b = 1 - \varepsilon$). In both cases $c_b = 0.5$, and the shear stress is obtained using an expression similar to the Chézy equation:

$$\tau = \rho_w s C_f u|u| \tag{107}$$

Following (Fraccarollo and Capart 2002) ψ is defined as

$$\psi = sC_f(gc_b(1-s)\tan\varphi)^{-1} \tag{108}$$

using in both cases $C_f = 0.014$, $\tan\varphi = 0.5$. In the Louvain case ψ is 0.016 and in the Taipei case ψ is 0.125. With this formulation it is possible to estimate the depth of the sediment layer in movement,

$$h_s = \psi u^2 \tag{109}$$

Figure 7 compares the experimental data and the computational solution for the Louvain test case, at times $t = 5t_o$, $7.5t_o$, $10t_o$, being t_o the geomorphologic time (Capart and Fraccarollo, 2002) approximately 0.101 s. The main differences are attributed to the fact that in the mathematical model mass and momentum exchanges through the bottom interface are not considered.

In the Taipei test case the exchanges through the bottom interface can not be neglected, and the model predicts unrealistic solutions as Figure 8 shows for time $t = 3t_o$.

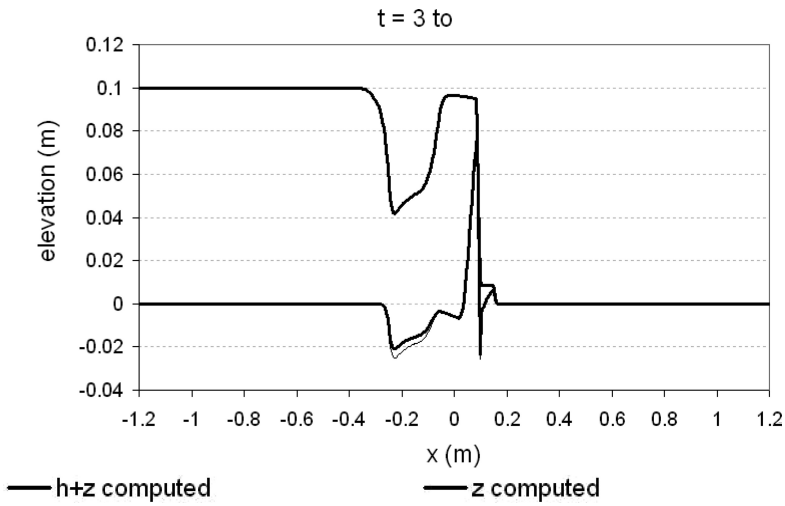


Figure 8. Experimental data and the computational solution for the Louvain test case, at time $t = 3t_o$.

8 CONCLUSIONS

In this work, we have illustrated the results of the analysis of the stability region for different systems of conservation laws. The local linearization of the source terms in hyperbolic systems and the unified formulation are the key to extend the C-property from the equilibrium of water at rest to equilibrium in flows over fixed and mobile bed. Also, robust and accurate numerical techniques have been defined to satisfy the positivity of the variables when demanded, leading to a perfect conservation in complex cases.

REFERENCES

Bermúdez A. & Vázquez M.E. 1994. Upwind methods for hyperbolic conservation laws with source terms, *Computers and Fluids*, 8: 1049–107.

Chow V.T. 1959. Open channel flow. MacGraw-Hill Book Co. Inc.

Burguete J. & García-Navarro P. 2001. Efficient construction of high-resolution TVD conservative shemes for equations with source terms: application to shallow water flows. *International Journal for Numerical Methods in Fluids*, 37, pp. 209–248.

Burguete J., García-Navarro P. & Murillo J. 2007 Analysis of the friction term in one-dimensional shallow water model: application to open channel and river flow. *Journal of Hydraulic Engineering,* (in press).

Brufau P., García-Navarro P., Ghilardy P., Natale L. & Savi F. 2000. 1D modelling of debris flow, *J. Hydrol. Res.* 38(6), 435–446.

Črnjarić-Žic N., Vuković S., & Sopta N. 2004. Extension of ENO and WENO schemes to one-dimensional transport equations. *Comput. Fluids* 33, 31–56.

Fraccarollo L. & Capart H. 2002. Riemann wave description of erosional dam-break flows, J. Fluid Mech, 461, 183–228.

Grass A.J. 1997 Sediment transport by waves and currents. *SERC London Cent. Mar. Technol Report No: FL29*.

Hudson J. & Sweby P.K. 2003. Formulations for Numerically Approximating Hyperbolic Systems Governing Sediment Transport. Journal of Scientific Computing, 19, 225–252, 2003.

Hudson J. & Sweby P.K. 2005. A high-resolution scheme for the equations governing 2D bed-load sediment transport, International Journal for Numerical Methods in Fluids 47, 1085–1091

Hudson J. 2002. Numerical techniques for morphodynamic modelling. *Ph.D. Thesis*, University of Reading.

Murillo J., Burguete J., Brufau P. & García-Navarro P. 2005, Coupling between shallow water and solute flow equations: analysis and management of source terms in 2D, *International Journal for Numerical Methods in Fluids*, 49:267–299.

Murillo J., García-Navarro P., Burguete J. & Brufau P. 2007. The influence of source terms on stability, accuracy and conservation in two-dimensional shallow flow using triangular finite volumes. *International Journal for Numerical Methods in Fluids*, 54, 543–590.

Rosatti G. & Fraccarollo L. 2006. A well-balanced approach for flow over movibile-bed with high sediment transport. Journal of Computational Physics, 220, 312–338.

Roe P.L. 1986. Discrete models for the numerical analysis of time dependent multidimensional gas dynamics. *J. Comp. Physics*, 63:458–476.

Summer B.M., Kozakievicz A., Fredsoe J. & Deigaard R. 1996. Velocity and concentration profiles in sheet-flow layer of movable bed. Journal of Hydraulic Eng. 122, 549–558.

Thacker W.C. 1981. Some exact solutions to the non linear shallow water equations. *Journal of Fluid Mechanics*. 107, 499–508.

Vázquez-Cendón M.E. 1999. Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry, *J. of Comp. Physics*, 148: 497–526.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Data-model integration – An approach to assimilation of sea level anomaly data into an oceanographic model

Vladan Babovic
Department of Civil Engineering, National University of Singapore, Singapore

ABSTRACT: Currents in Malacca and Singapore Strait are due to tides with significant contributions by residual currents, mainly related to monsoon cycles. Remote sensing data are expected to add vital information to assess physical parameters in the real life system. The paper describes the utilization of satellite altimetry data generated by TOPEX/POSEIDON (T/P) and Jason-1 altimeter data of the South China Sea and the Andaman Sea bordering the Indian Ocean. This paper is prelude to improving seasonal transport patterns in Southeast Asian waters using remote altimeter data in operational (forecasting) modelling. Both the regional and local scales are addressed.

1 INTRODUCTION

The South China Sea and Malacca Strait domain area considered in this paper covers approximately 3500 km by 3500 km. The marine area consists of a number of basins, each with characteristic depth, length and width scales, often separated by high sills. While depths in the South China Sea can be of the order of 5000 m, the depth in the Java Sea is almost everywhere below 100 m. Tides vary strongly over the area, from fully diurnal to fully semi-diurnal, with many areas being a mixture of the two. The main tidal forcing originates in the Pacific Ocean, with secondary forcing from the Indian Ocean through various inlets in the Island chain of the Greater and Lesser Sunda Islands. The area is also characterised by monsoons. Oceanographically speaking, there is a large scale balance between the tidal forcing from the Pacific, the monsoon and topography/geometry.

1.1 *Hydrodynamics on a regional scale*

Good knowledge of the driving hydrodynamic processes on a regional scale is a prerequisite to properly assess the processes on fine spatial-temporal resolutions. With this aim WL|Delft Hydraulics developed hydrodynamic model of the Malacca Strait (Figure 1). The dimensions and geometrical variations give rise to a complicated tidal as well as residual current response in the sub-basins which is not seen elsewhere in the world.

1.2 *Transport on a local scale*

The tide- and wind-induced residual currents are some of the main hydrodynamic carriers for the transport of dissolved and particulate matter. The requirements for increasing accuracy of such models make it necessary to extend the fine-scale domain of coastal models into a scale where the large scale spatial-temporal variation over the open boundaries have to be taken into account.

Prevailing monsoon wind forcing and the spatially varying Coriolis force induces a mean sea level gradient on a monthly and seasonal time scales. These residual currents are rather sensitive for temporal variations in the mean sea level gradient at the open model boundaries. Although the processes that trigger these phenomena are relatively well-understood, their quantification

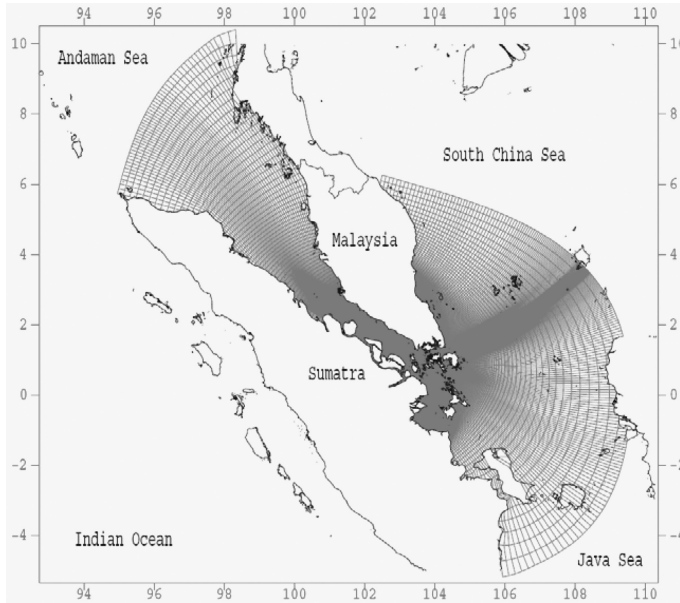


Figure 1. The extent of computational domain.

in real-life applications requires characterisation of driving forces in order to meet the accuracy demands.

Within the present effort the synoptic Sea Surface Anomalies (SSAs) extracted from the thirteen year records of altimeter data (RADS) are used to assess the effect of large scale geostrophic circulation effects. The Sea Surface Anomaly (SSA) is both spatial and temporal variation in water levels with respect to the sea bed which remains after removal of all strictly harmonic variations. The main objective is to establish an efficient manner for forecasting of non-tidal currents and their assimilation in hydrodynamic model.

2 TOPEX/POSEIDON SATELLITE ALTIMETRY

The T/P project is the result of a collaboration between the US space agency NASA and French space agency CNES. An important contribution of the T/P program is that some of the difficulties raised in the traditional way of setting up a tide resolving model can be made easier by means of satellite altimetry. Also, satellite altimetry enables us to accurately estimate sea sufrace annomaly on high spatial-temporal resolutions. Presently, the outdated T/P satellite has been replaced by Jason-1.

2.1 *Satellite altimetry: An orbiting tide gauge*

The principle of satellite altimetry is that of an orbiting nadir-looking radar system. By means of such a system the response of the ocean is measured from a short hypothetical EM pulse that bounces off the sea surface. The received information on board the satellite consists of radar waveform samples. These are basically histograms describing the distribution of power against time of the reflected signal. All information retrieved from the satellite altimeter measurements are derived from the radar waveform samples or auxiliary information, either directly observed by instruments on the satellite or obtained from geophysical models.

2.2 *Altimeter footprint*

The altimeter footprint is an area of 5 to 10 km in diameter depending on the state of the sea level. The ranges observed by the altimeter are a statistical mean in that area thereby filtering out wind wave effects from the sea level data. Wind waves with spatial wavelengths under the footprint diameter are by definition invisible to the T/P altimeter. A second filter consists of an algorithm used in the post-processing of altimeter data. It is based upon the sea state bias correction, which removes the asymmetry in the radar reflection from signals returning from either the crests or troughs of wind waves in the radar footprint.

2.3 *Orbit characteristics*

Since the start of satellite altimetry for scientific applications with GEOS-3 in 1975 the trajectories were usually chosen between 750 and 1330 km resulting in orbital periods between 6000 and 6700 seconds. All altimeter orbits are circularised in the best possible way to minimise height variations while high inclinations are chosen to maximise the latitude range seen by the radar. A unique orbit design criterion for satellite altimeter systems is that trajectories with repeating ground tracks are often preferred for oceanographic applications. For T/P and Jason-1 it implies that the ground track overflown by the altimeter will start to repeat itself after 127 orbital revolutions which take place in a period of 9.9156 (wall-clock) days which corresponds to 10 (nodal rate corrected) days. This 9.9156 day window is normally referred to as a cycle. The altimeter database storage structure is fully described in the RADS project report [9]. Any altimeter system that senses the ocean will under-sample signals that occurs within a time frame less than a cycle. At best, use can be made of intersecting ground tracks, so-called cross-over points, or assumptions can be made on the spatial or temporal smoothness of the phenomenon that we expect to be observed by the altimeter.

2.4 *Altimeter data*

The pass files contain geophysical data records (GDRs) for all available altimeter systems. Each pass files contains $\sim$2000 records depending on the land/ocean/ice distribution along the track in question. Based on data from RADS, it is possible to calculate fully corrected sea surface anomaly (SSA). However, due to orbiting period of approximately 10 days over the exactly same location, it is not possible to update SSA information on a continuous basis.

3 SSA-RESOLVING MODEL

The main difference between the tide-resolving simulations and the SSA simulations is that the latter depend on meteorological forcing by wind and possibly by atmospheric pressure whereas, traditionally, tidal simulations are not forced by atmospheric flows although diurnal and annual cycles in wind stress vector can occur. In order to obtain the SSA, the elimination of the water level harmonics is essential. For this we chose to eliminate all harmonics detected in the altimeter observations rather than using tidal model because the difference between the latter model estimates and the harmonics in the altimeter observations may result with spurious oscillations in the SSA.

In addition to SSA information ECMWF wind data at spatial resolution of $0.33^0 \times 0.33^0$ and at temporal resolution of 6 hours were also available. It has been stated previously that SSA is essentially induced by long-term wind setup, such as monsoon.

Figure 3 is an illustration of spatial cross-correlation between wind intensity and SSA. The fact that in most of the domain correlation coefficients take values between 0.7 and 0.75, clearly demonstrates the interdependence.

At the same time, large spatial extent of high correlation values poses a problem of different kind. If we were to construct a model which fully exploits the correlation structure, the model would have to be of very high dimensionality. Highly dimensional models, in turn, are notorious of its ability to overfit and require extraordinarily large amounts of data to address this issue.

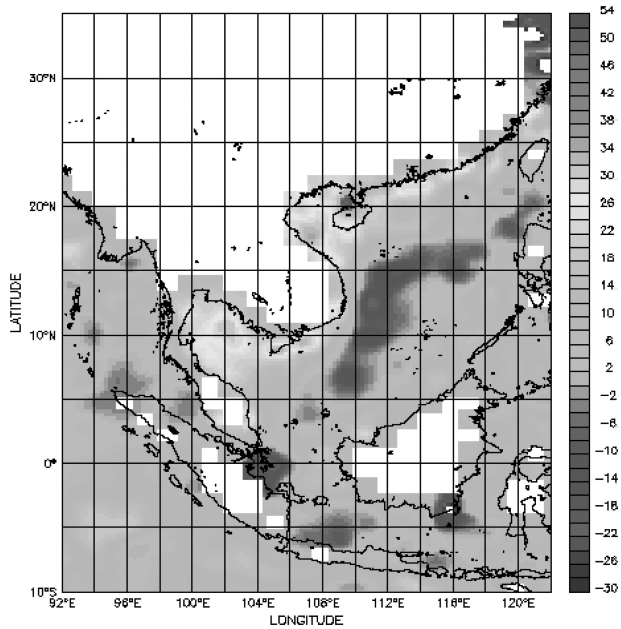


Figure 2. Sea Surface Anomaly as observed on December 14, 2005.

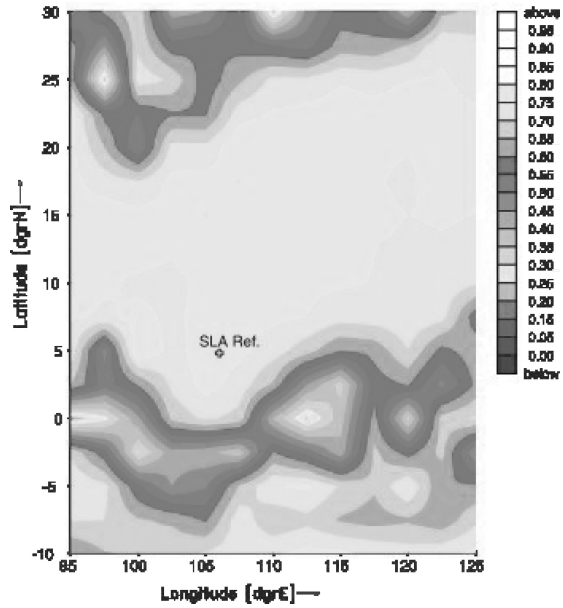


Figure 3. Cross-correlation coefficients between ECMWF wind and SSA.

4 FORECAST AND ASSIMILATION OF SEA SURFACE ANOMALY

The desire to predict the future and understand the past governs the search for laws that explain the behaviour of observed phenomena. If the underlying deterministic equations are known, in principle they can be solved to forecast the outcome based on knowledge of the initial conditions and evolution of forcing terms. In hydraulic modelling, for example, governing laws are Navier-Stokes equations,

whereas the forcing term is the evolution of the status of atmosphere (atmospheric pressures and resulting wind fields). Initial conditions describe the sea status (current speeds and directions, water levels) in entire computational domain at the beginning of computation. Once the initial and forcing terms are *precisely* specified, it should be possible to *precisely* calculate the evolution of the status of the sea, from its specified initial conditions and as a consequence of applied forcing.

However, even under these, almost ideal circumstances, the model results are not precise. Every model is indeed only a *model* of reality; it employs a number of simplifying assumptions which inevitably produces inaccuracies. In a numerical modelling one discretises a domain, and therefore is not able to resolve numerous sub-grid phenomena. Also, the errors in the model parametrisation greatly contribute to errors of numerical models. Finally, it is impossible to *precisely* define initial conditions and forcing terms in the *entire* computational domain. All these imprecisions and uncertainties can accumulate to produce fairly poor model results despite perfect knowledge of the governing laws [1].

4.1 *How can we make more accurate models?*

Various schemes may be utilised to make models more accurate. When observations of the modelled phenomena are available and some knowledge of the aforementioned errors exists, data assimilation methods can be used in order to improve the model solution. Data assimilation is a methodology which can optimise the extraction of reliable information from observations and combine it with, or assimilate it in, numerical models. A number of different data assimilation procedures can be adopted. These are designed to either improve description of initial conditions at the time of forecast or provide correction of model predictions during a forecast period [2].

The errors associated with the final model outputs make the model to be far from the reality despite our perfect knowledge of governing laws. There are different data assimilation procedures that can be applied to improve accuracy of model outputs. These are designed to either improve description of initial conditions (boundary conditions) or to provide correction of model predictions. The data assimilation procedures can be classified according to the variables modified during the updating process. In WMO (1992, see also Refsgard, 1997) four different methodologies have been defined:

1. *Updating of input parameters* – This is the classical method justified by the fact that input uncertainties may be dominant by the dominant error in operational forecasting.
2. *Updating of state variables* – Adjustment of state variables can be done in different ways. The theoretically most comprehensive methodology is based on Kalman filtering (Gelb, 1974). Kalman filtering is the optimal updating procedure for linear systems, but can with some modifications also provide an approximate solution for non-linear systems.
3. *Updating of model parameters* – continuous adaptation of model parameters is a matter of continuous debate. The prevailing view seems to be that for hydrodynamic models of non-trivial complexity recalibration of model parameters at every time step has no real advantages, as the operation of any hydrodynamic system can not significantly change over the short interval of time.
4. *Updating of output variables (error prediction)* – The deviations between the observed variables and the simulation mode forecast are model errors. Possibility of forecasting these errors and superimposing them to the simulation model forecasts, usually gives more accurate performance. This method is most often referred to as error prediction and is the method investigated in this study.

Recently, neural networks (NNs) have been applied successfully in various fields, as a modelling technique for nonlinear systems. A major characteristic of NNs is that they are capable of approximating nonlinear functions to any arbitrary degree of accuracy (Hornik, 1989; Cybenko, 1989). As in many other processes, the nonlinearity, complexity, and lack of understanding of wastewater related processes favour the potential use of neural networks. However, a typical data-driven (stochastic) model (which can be either linear or a nonlinear function) suffers from a drawback

in that it is synthesised only on the available data, without detailed knowledge of the underlying processes (black box approach). It may be inadequate and inaccurate for prediction and extrapolation when the data are sparse or noisy. Consequently, the alternative of using a hybrid model that integrates both deterministic (prior) model and stochastic model for modelling wastewater processes appears promising, especially when the more accurate model outputs (predictions) are required. Philosophically, this approach is based on the assumption that the source of complex behaviour is not only intrinsic but also extrinsic due to the stochastic influences, also referred to as noise or exogenous schocks, i.e., outside perturbations of the system.

In the context of the present effort, a data-driven model is developed to in error prediction sense. The forecast will be assimilated into hydrodynamic model to provide correct initial conditions and account for anomalies in sea levels.

5 DATA-DRIVEN FORECAST OF SEA SURFACE ANOMALY

For obtaining short-term forecasts, statistical models could be applied. Simplest models take the form of a multiple linear regression model, as initially developed by Yule [13].

$$x_t = \boldsymbol{a}\boldsymbol{x}_{t-1} + \boldsymbol{b}\boldsymbol{e}_t \tag{1}$$

where $\mathbf{x}_t = (x_t, x_{t-1}, x_{t-2}, \ldots, x_{t-(d-1)})$, $\mathbf{e}_t = (e_t, e_{t-1}, e_{t-2}, \ldots, e_{t-(d-1)})$, $\mathbf{b} = (b_1, b_2, b_3, \ldots, b_d)$ and $\mathbf{a} = (a_1, a_2, a_3, \ldots, a_d)$.

Such lag vectors are also called tapped delay lines. Yule's original idea for forecasting was that future predictions can be generated by using the immediately preceding values. There is a deep connection between time-lagged vectors and the underlying dynamics which was first proposed by Packard [1] and Takens [12] who published the first formal demonstration of such a connection, and later this was strengthened by Sauer et al [11].

Takens' Time-delay Embedding Theorem [12] can be written as follows: Given a dynamical system with a d-dimensional solution space and an evolving solution $\mathbf{x}(t)$, let y be some observation $y(\mathbf{x}(t))$. Let us also define the lag vector (with dimension d and time delay τ) $\mathbf{y}_t \equiv (y_t, y_{t-\tau}, y_{t-2\tau}, y_{t-3\tau}, \ldots, y_{t-(d-1)\tau})$. Then, under very general conditions, the space of vectors $\mathbf{y}_t$ generated by the dynamics contains all of the information of the space of solution vectors $\mathbf{x}_t$. The mapping between them is smooth and invertible. This property is referred to as diffeomorphism and this kind of mapping is referred to as an embedding. Thus, the study of the time series y_t is also the study of the solutions of the underlying dynamical system $\mathbf{x}(t)$ via a particular coordinate system given by the observable y. Takens' theorem implies that there exists, in principle, a non-linear autoregression of the general form:

$$y(t) = g[y(t-1), y(t-2), y(t-3), \ldots, y(t-T)] \tag{2}$$

which models the series exactly (assuming that no noise is present). The problem addressed in this article is the one of finding this non-linear auto-regression and for that purpose we use Radial Basis Function Neural Networks. Takens' theorem implies that we should be able to use the history of the SSA evolution to accurately forecast future SSAs.

5.1 *Stochastic data-driven modelling*

The linear *AutoRegressive* (stochastic data-driven) model with *eXogenous* (ARX) inputs applied here, can be given as:

$$o_{k+1} = w_1 x_{1k} + w_2 x_{2k} + w_3 x_{3k} + w_4 \tag{3}$$

where, $x_{1_k} = o_k$, o_k stands for error prediction of the deterministic model (in (5) and (6) o_k is shown as $\hat{y}$*stochastic*), x_1 stands for deterministic model error, $\overrightarrow{x}_2$ *and* $\overrightarrow{x}_3$ stand for known exogenous inputs ($\overrightarrow{x}_2$ is observed SLA data and $\overrightarrow{x}_3$ is output from the deterministic model), w_i contains the unknown model parameters ($i = 1, \ldots 3$, 3 is the dimension of the input space), and k is the time delay. The input vector is augmented with the bias input $x_4 = 1$.

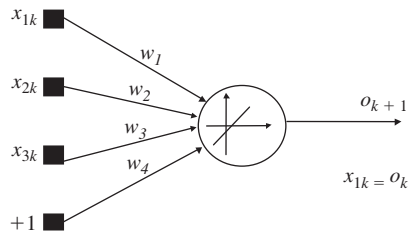


Figure 4. Linear ARX model.

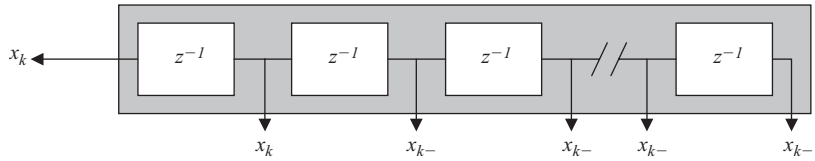


Figure 5. Delay space embedding (or tapped delay-line memory model). Each z^{-1} operator introduces a one time step delay and n stands for memory depth.

The model given in (1), can be shown graphically in Fig 4.

Generally, equation (1) the n-dimensional input vector (multiple features data set) represents mapping $\Re^n \rightarrow \Re^1$, i.e., it is a hyperplane in a $n+1$ dimensional space. Such a form of traditional ARX models is based on a delay space embedding (tapped delay-line memory). Hence, a memory is a buffer containing the n most recent inputs (Fig 5).

The nonlinear extension (NARX) of the same ARX model used here is a neural network that uses radial basis functions as the activation functions in the hidden layer (unlike the *multilayer perceptron*, which uses sigmoidal function,) and is referred to as *radial basis function neural network* (RBF NN). It is obtained when several RBF neurons in hidden layer are used in parallel and are connected to a linear output neuron, which represents mapping $\Re^n \rightarrow \Re^1$ (see Figure 6). The structure of our RBF NN model consists of one input layer, one hidden layer and one output layer. As reported in Kecman [7] there is no 'sound' evidence that the NN structures with more hidden layers generate more accurate results than structures with one hidden layer. Thus, all experiments performed here use a single hidden layer structure.

In the NARX model, the basic form of autoregressive function is retained. As can be seen in Figure 6, without hidden layer nodes RBF NNs are equivalent to the linear ARX models.

An introduction of a hidden layer containing nonlinear RBFs enables RBF NN to model any nonlinear dependency to any desired degree of accuracy. This is a rationale for application of RBF NNs in this paper. In mathematical formulation the RBF network can be written as:

$$o = F(\mathbf{x}) = \sum_{j=1}^{J} w_j \varphi_j(\|\mathbf{x} - \mathbf{c}_j\|) \tag{4}$$

where $\| \mathbf{x} - \mathbf{c}_j \|$ stands for the distance between data point $\mathbf{x}$ and RBF center $\mathbf{c}_j$.

Therefore, the RBF NN's output is a weighted sum of the hidden layer's activation functions. The most commonly used basis functions φ are Gaussian RB functions. There are three types of parameters that can be optimised (learned) in RBF neural network model with Gaussian basis functions:

1. *RBF Centers (hidden layer parameters),*
2. *Widths or standard deviations of RBFs (hidden layer parameters)*
3. *Output layer weights*

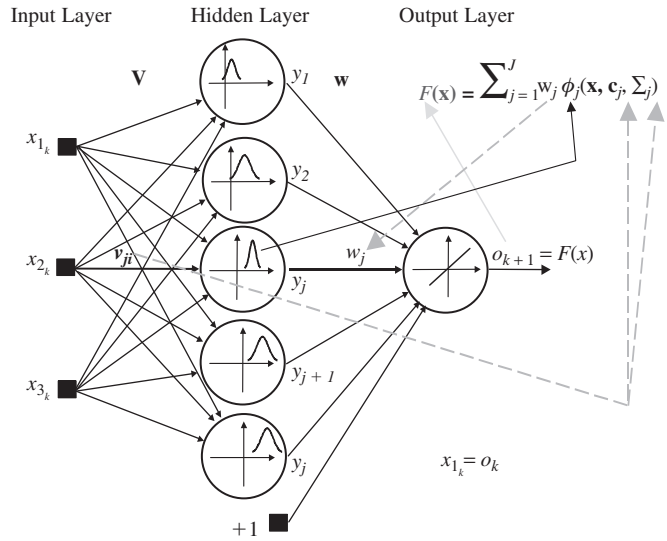


Figure 6. Schematisation of radial basis function neural network.

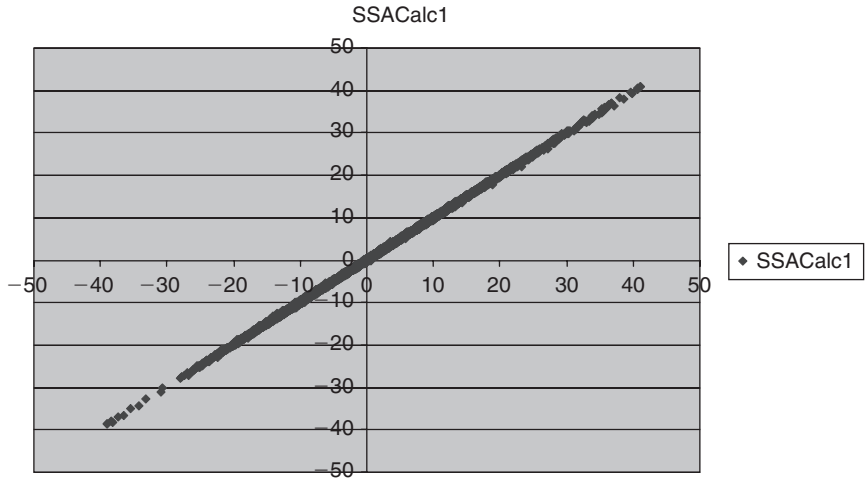


Figure 7. Scatter plot for sea level anomaly for lead time of one day.

For training of RBF NN models different strategies exist. The strategy applied here is to determine the hidden layer weights first and subsequently the output layer weights are estimated by using the pseudoinversion of the so-called design matrix **G** (see more in Kecman, 2001).

Results of application of RBF NN are most encouraging as exemplified in Figure 7 and Table 1.

6 CONCLUSION

Means for data collection and distribution have never been as advanced as they are today. While advances in data storage and retrieval continue at a breakneck pace, the same cannot be asserted about advances in information and knowledge extraction from data. Without such developments,

Table 1. Summary of statistical measures of accuracy for varying lead times

Performance	dSSA1	dSSA2	dSSA3	dSSA4	dSSA5
MSE	0.025	0.177	1.346	3.403	6.288
NMSE	0.024	0.042	0.086	0.147	0.202
MAE	0.120	0.294	0.860	1.388	1.901
Min Abs Error	0.000	0.000	0.000	0.000	0.001
Max Abs Error	1.050	14.488	14.439	13.400	13.599
r	0.988	0.979	0.956	0.925	0.894

however, we risk missing most of what data have to offer. However, as this paper demonstrated, forecasting on the basis of the data alone is not the complete story, at least not in scientific and engineering domains. Although a datacentric approach is at the focal point in many research areas, the current paper advocates a slightly different approach. Clearly, there is an enormous amount of knowledge and understanding of physical processes that should not merely be discarded. Forecast skill based on deterministic models should not be ignored. We strongly believe that the most appropriate way forward is to combine the best of the two approaches – theory-driven, understanding rich with data-driven modeling processes. The present paper presented an approach where a theorydriven deterministic model was coupled with a data-driven technique to produce a hybrid model with very good forecast skills. The resulting hybrid model performs much better than each of its constituents taken exclusively. The benefits of coupling were expressed not only in terms of accuracy, but also in terms of extension of the forecast horizon.

REFERENCES

Babovic, V., R. Cañizares, H.R. Jensen, and A. Klinting: 2001, Artificial neural networks as a routine for updating of numerical models. *ASCE Journal of Hydraulic Engineering*, **127**, 181–193.

Babovic, V., and Keijzer, M., 1999., Forecasting of river discharges in the presence of chaos and noise, in Marsalek, J., (ed), Coping with Floods, Kluwer.

Babovic, V. and M. Keijzer, 2000, Genetic programming as a model induction engine. *Journal of Hydroinformatics*, **2**, 35–60.

Casdagli, M.C., and Weigend, A.S., 1994., Exploring the continuum between deterministic and stochastic modelling. Time Series Prediction: Forecasting the Future and Understanding the Past, SFI Studies in the Sciences of Complexity, Proc. Vol. XV, in Weigend, A. S., and Gerschenfeld, N.A., (eds.) Addison-Wesley.

Farmer, J.D., and Sidorowich, J.J., 1987, Predicting chaotic time series, *Physical Review Letters*, Vol. **59**, No. 8, pp.62–65.

Koza, J.R.: 1992, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, MA, USA.

Packard, N.H., Crutchfeld, J.P., Farmer, J.D., and Shaw, R.S., 1980., Geometry from a time series, *Phys. Rev. Lett.*, Vol. **45**(**9**), pp. 712–716.

Ray, R.D. 1999 *A Global Ocean Tide Model from TOPEX/POSEIDON Altimetry: GOT99.2*, NASA M-1999-209478, NASA/GSFC, Greenbelt Md 20771. 1999.

Schrama, E., 1997 *Satellite Altimetry, Ocean Dynamics and the Marine Geoid*, Lecture notes in Earth Sciences, Vol. 65, pp 405–454, Springer Verlag.

Schrama E., R. Scharroo & M. Naeije, 2000 *Radar Altimeter Database System (RADS): Towards a generic multi-purpose altimeter database system*. Netherlands Remote Sensing Board (BCRS) report.

Sauer, T., Yorke, J., A., and Casdagli, M., 1991., Embedology, J. Stat. Phys., Vol. 65 (3/4), pp. 579–616.

Takens F., 1981, Detecting strange attractors in turbulence. *Dynamical Systems and Turbulence,* in Rand, D.A. and Young, L.-S. (eds), Lecture notes in mathematics, Vol. 898, pp. 336–381, Springer-Verlag.

Yule, G.U., 1927., Philosophical Transactions of Royal Society London, A **227**, 267.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Data-driven knowledge discovery: Four roads to vegetation-induced roughness formulae

Vladan Babovic
Department of Civil Engineering, National University of Singapore, Singapore

ABSTRACT: The paper describes the process of induction of equations for the description of vegetation-induced roughness from several angles. Firstly, it describes two approaches for obtaining theoretically well-founded analytical expressions for vegetation resistance. The first of the two is based on simplified assumptions for the vertical flow profile through and over vegetation, whereas the second is based on an analytical solution to the momentum balance for flow through and over vegetation. In addition to analytical expressions the paper also outlines a numerical 1-DV k-ε turbulence model which includes several important features related to influence plants exhibit on the flow. Last but not least, the paper presents a novel way of applying genetic programming to the results of the 1-DV model, in order to obtain an expression for roughness based on synthetic data. The resulting expressions are evaluated and compared with an independent data set of flume experiments.

1 INTRODUCTION

Proper modelling of the flow resistance of wetlands and vegetated floodplains is of great practical importance. Many research initiatives have been undertaken in order to improve on the description of the relationship between flow resistance and the presence and spatial distribution of vegetation. Both analytical and experimental studies of vegetation-related resistance to flow and the equivalent resistance coefficients have shown that the resistance coefficients are water depth dependent. Consequently, the traditional approach of using a single resistance coefficient fails to correctly describe the physics of the phenomenon. One way of improving upon this description is updating the equivalent resistance coefficient based on the computed water depth. In order to do so, a relation between vegetation characteristics, bed resistance, water depth and equivalent resistance coefficient is needed.

In this paper, four different methods to obtain such a relation are elaborated. The first is based on a simple analytical description of the division in discharge through and over submerged vegetation. The second is based on an analytical solution to the momentum balance of flow through and over vegetation. The third is based on a numerical 1-DV model and finally, the fourth is based on data mining.

Data mining is concerned with extracting useful information from data. However, mining the data *alone* is not the entire story. Scientific theories encourage the acquisition of new data and this data in turn leads to the generation of new theories. Traditionally, the emphasis is on a theory, which demands that appropriate data be obtained through observation or experiment. In such an approach, the discovery process is what we may refer to as *theory-driven*. Especially when a theory is expressed in mathematical form, theory-driven discovery may make extensive use of strong methods associated with mathematics or with the subject matter of the theory itself. The converse view takes a body of data as its starting point and searches for a set of generalisations, or a theory, to describe the data parsimoniously or even to explain it. Usually such a theory takes the form of a precise mathematical statement of the relations existing among the data. This is the *data-driven* discovery process.

We believe that the most appropriate way for scientific applications of data mining is to combine the better of the two approaches: theory-driven, understanding-rich with data-driven discovery process. Clearly, there is an enormous amount of knowledge and understanding of physical processes that should be incorporated in the discovery process. The work described here is part of a research effort aiming at providing new hypotheses built from data. The ultimate objective is to build models which can be interpreted and further manipulated by the domain experts. Once a model is interpreted, it can be used with confidence. It is only in this way that one can take full advantage of knowledge discovery and advance our understanding of physical processes, of, in this case, vegetation resistance.

2 BACKGROUND

The effect of vegetation on flow is generally expressed as an effect on the hydraulic roughness. Early measurements (18th century) of flow velocities in channels revealed that the depth-averaged flow velocity [m/s] was a function of the water level slope i[m/m] and the hydraulic radius R[m]. In 1776 Antoine de Chézy published a simple equation that includes an engineering factor C, the Chézy value, which was at first thought to be a constant (Vernon-Harcourt, 1896). The well-known Chézy formula is:

$$\bar{u} = C\sqrt{Ri} \tag{1}$$

In this equation C is a parameter that expresses the hydraulic roughness of the bed and banks of a channel [$m^{1/2}$ s^{-1}]. Further investigations, by Nikuradse (1930), revealed that the roughness of the bed affects the roughness length z_0 [m] in the logarithmic velocity profile for a fully rough bed derived by, among others, Prandtl and Von Kármán:

$$u(z) = \frac{u_*}{\kappa} \ln\left(\frac{z}{z_0}\right) \tag{2}$$

Nikuradse showed that for hydraulically rough walls, the roughness length of the logarithmic velocity profile can be expressed as $k_N/30$, where k_N is the Nikuradse equivalent roughness (Nikuradse, 1930). Calculating the depth-averaged velocity, and applying the Chézy formula and the Nikuradse roughness height yields the White-Colebrook formula for the Chézy value:

$$C = 18 \log\left(\frac{12R}{k_N}\right) \tag{3}$$

With an increasing roughness height the value for C decreases. Various alternative expressions for flow resistance exist, for example those of Strickler or Manning, which can all be mutually converted. Essentially, in using a single roughness equation such as the White-Colebrook formula, vegetation is treated as large bed structures with a logarithmic flow profile above them. In reality, however, there is flow over and through submerged vegetation, and the vertical flow profile deviates from the logarithmic one. A typical time-averaged velocity profile for submerged vegetation is shown in Figure 1. Four distinct zones can be identified in the time-averaged vertical velocity profile for flow through and above submerged vegetation:

1. In the first zone, near the bed, the velocity is highly influenced by the bed, and its vertical profile joins the logarithmic boundary layer profile.
2. In the second zone, which corresponds to the zone inside the vegetation sufficiently away from the bed and from the top of the vegetation, the velocity is uniform.
3. In the third zone, near the top of the vegetation, there is a transitional profile between the uniform velocity inside the vegetation and the logarithmic profile above it. The profile in this zone can be approximated by an exponential function.
4. Finally, the fourth zone corresponds to the zone above the vegetation, where a logarithmic profile is observed, which has a zero-plane displacement below the top of the vegetation layer.

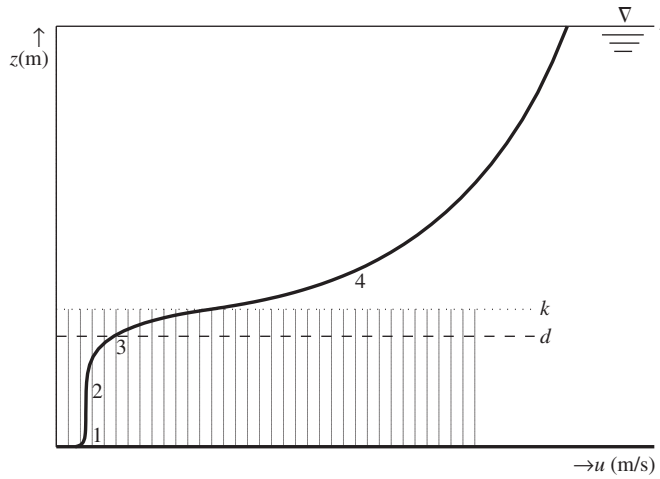


Figure 1. Four zones in the vertical profile for horizontal velocity, $u(z)$, through and over vegetation, $h =$ water depth (m), $k =$ vegetation height (m), $d =$ zero-plane displacement (m).

The White-Colebrook formula fails here and another type of resistance formula should be sought for. A considerable amount of research has been carried out on the effects of vegetation on the hydraulic resistance, extending the basic ideas of Nikuradse (1930). Early work includes Einstein & Banks (1950), Kouwen *et al*. (1969), Kouwen & Unny (1973), Klaassen & Van der Zwaard (1974) and Petryck & Bosmajian III (1975). In a study by Dawson & Charlton (1988), a literature search has been carried out on the hydraulic resistance of vegetation, resulting in over 360 publications. Since then, many more publications have followed. A limited overview of recent research includes studies on the improvement of flow resistance formulae (Darby, 1999; Hasegawa *et al*., 1999; Meijer & Van Velzen, 1999; Stephan & Gutknecht, 2002; Järvelä, 2002; Mason *et al*., 2003; James *et al*., 2004 and Järvelä, 2004), on analytical approaches for the vertical profile of horizontal velocity (Klopstra *et al*., 1997; Carollo *et al*., 2002; Katul *et al*., 2002), on biomechanics and streamlining of vegetation (Fahti-Maghadan & Kouwen, 1997) and on turbulence characterisation for submerged rods and vegetation (Shimizu & Tsujimoto, 1994; Ikeda & Kanazawa, 1996; Nezu & Naot, 1999; Nepf & Vivoni, 2000; Ikeda *et al*., 2001; Fisher-Antze *et al*., 2001; López & García, 2001; Ghisalberti & Nepf, 2002; Righetti & Armanini, 2002; Wilson *et al*., 2003 and Ghisalberti & Nepf, 2004). However, no acceptable formulation for roughness induced by submerged vegetation, valid for a wide range of vegetation properties and water depths has been found as of yet. This is the main inspiration for the present work.

3 THEORETICAL, ANALYTICAL FORMULATIONS OF RESISTANCE DUE TO VEGETATION

3.1 *Case of non-submerged vegetation*

Non-submerged flow conditions can be successfully treated analytically. The balance of horizontal momentum in uniform steady flow conditions through non-submerged vegetation expressed as cylinders dictates that total shear stress is equal to the sum of the bed shear stress and the equivalent shear stress due to vegetation drag:

$$\tau_t = \tau_b + \tau_v \tag{4}$$

where τ_t denotes the total fluid shear stress:

$$\tau_t = \rho_0 ghi \tag{5}$$

τ_b denotes the bed shear stress:

$$\tau_b = \rho_0 g \frac{u^2}{C_b^2} \tag{6}$$

and τ_v is the vegetation resistance force per unit horizontal area:

$$\tau_v = \tfrac{1}{2}\rho_0 C_D m D h u^2 \tag{7}$$

The vegetation resistance force is thus modelled as the drag force on a random or staggered array of rigid cylinders with uniform properties. The uniform flow velocity through non-submerged vegetation follows from the momentum balance and is given by:

$$u_{cb} = \sqrt{\frac{hi}{\frac{1}{C_b^2} + \frac{C_D m D h}{2g}}} \tag{8}$$

The discharge per unit width through the vegetation is given by:

$$q = h u_{cb} \tag{9}$$

From the discharge through the vegetation the representative Chézy value for non-submerged vegetation is calculated as:

$$C = \frac{q}{h\sqrt{hi}} \tag{10}$$

Therefore, the representative Chézy value for non-submerged vegetation becomes:

$$C_k = \sqrt{\frac{1}{\frac{1}{C_b^2} + \frac{C_D m D h}{2g}}} \tag{11}$$

When the bed resistance is negligible with respect to the drag force of the vegetation, the resistance coefficient can be reduced to:

$$C_k = \sqrt{\frac{2g}{C_D m D h}} \tag{12}$$

3.2 *Case of submerged vegetation*

Two different methods to derive theoretical resistance formulae for submerged vegetation were applied by Baptist (2005), i.e. the "method of effective water depth" and a method utilising an analytical derivation of the momentum balance for flow through vegetation.

3.2.1 *Submerged vegetation – Method of effective water depth*

The "method of effective water depth" is inspired by earlier approaches by Hong (1995), Campana (1999) and Stone & Shen (2002). In this method, the four zones in the velocity profile of flow through and above vegetation are represented in a simplified way and reduced to two flow zones:

1. A uniform flow velocity, u_c, inside the vegetation, and
2. A logarithmic flow profile, u_u, above the vegetation, with zero-plane displacement at height k and with an additional slip velocity of size u_c.

This simplified velocity profile is presented in Figure 2. The flow velocity in the vegetated section follows from the momentum balance for flow through vegetation. The flow velocity in the

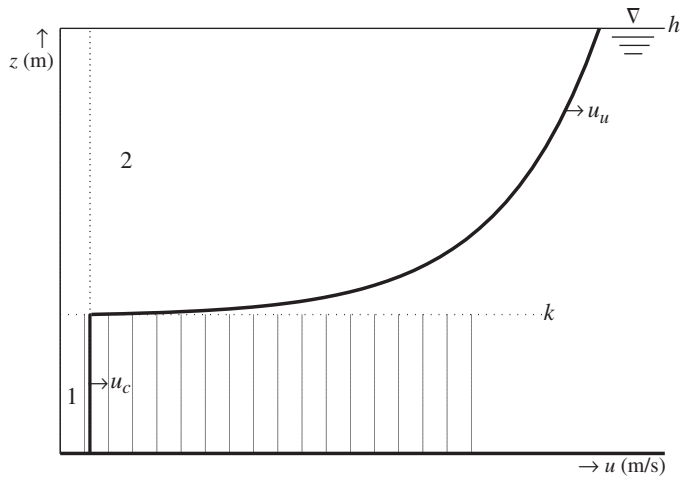


Figure 2. Representation of the vertical velocity profile in two zones for the method of effective water depth, h = water depth (m), k = vegetation height (m), u_c = uniform flow velocity profile (m/s), u_u = logarithmic flow velocity profile (m/s).

uniform part of the profile is, therefore, given by the formula for flow through fully submerged vegetation, which resembles Eq. (8):

$$u_c = \sqrt{\frac{hi}{\frac{1}{C_b^2} + \frac{C_D m D k}{2g}}} \tag{13}$$

The logarithmic velocity profile above the vegetation (u_u) is given by:

$$u_u(z) = \frac{u_*}{\kappa} \ln\left(\frac{z-k}{z_0}\right) + u_c \tag{14}$$

Note that the uniform flow velocity through the vegetation, u_c, is added to the logarithmic velocity as a constant 'slip velocity' to match both profiles at $z = k$.

The height-averaged velocity above the vegetation then becomes:

$$\bar{u_u} = \frac{1}{h-k}\int_k^h u_u(z)dz = \frac{u_*}{\kappa} \ln\left(\frac{h-k}{z_0} - 1\right) + u_c = \frac{u_*}{\kappa} \ln\left(\frac{h-k}{ez_0}\right) + u_c \tag{15}$$

where e is the base of the natural logarithm.

From the discharge relationship, Eq (10), the representative Chézy value follows as:

$$C_r = \sqrt{\frac{1}{\frac{1}{C_b^2} + \frac{C_D m D k}{2g}}} + \frac{(h-k)^{\frac{3}{2}} \frac{\sqrt{g}}{\kappa} \ln\left(\frac{h-k}{ez_0}\right)}{h^{\frac{3}{2}}} \tag{16}$$

Note that the first term on the right-hand side equals the representative roughness of non-submerged vegetation for $h = k$, Eq. (11). Note further that part of the second term on the right-hand side can be approximated by the White-Colebrook formula:

$$\frac{\sqrt{g}}{\kappa} \ln\left(\frac{h-k}{ez_0}\right) \equiv 18 \log\left(\frac{12\,(h-k)}{k_N}\right) \tag{17}$$

where k_N, the equivalent Nikuradse roughness height of the top of the vegetation layer [m], is taken equal to $30z_0$ (Nikuradse, 1930). This approach presents one unknown variable: z_0 being the roughness height of the top of the vegetation layer.

3.2.2 *Submerged vegetation – Analytical method*

In the analytical method the velocity profile inside the vegetation layer is obtained by analytically solving the momentum equation:

$$\frac{\partial \tau_{xz}}{\partial z} - \rho_0 g \frac{\partial h}{\partial x} - \frac{1}{2}\rho_0 C_D m D u^2(z) = 0 \tag{18}$$

After quite some manipulations presentation of which exceeds the scope of the present paper, the formula for the equivalent Chézy resistance coefficient becomes:

$$C_r = \frac{1}{h^{3/2}} \begin{cases} L\left[2\left(u_{vk} - \sqrt{a_v + u_{v0}^2}\right) + u_{v0} \ln\left(\frac{(u_{vk}-u_{v0})\left(\sqrt{a_v+u_{v0}^2}+u_{v0}\right)}{(u_{vk}+u_{v0})\left(\sqrt{a_v+u_{v0}^2}-u_{v0}\right)}\right)\right] \\ + \frac{\sqrt{g(h-k)}}{\kappa(h-k)}\left[(h-d)\ln\left(\frac{h-d}{z_0}\right) - (k-d)\ln\left(\frac{k-d}{z_0}\right) - (h-k)\right] \end{cases} \tag{19}$$

The formula for the vegetation-related resistance coefficient for flow through non-submerged vegetation can be applied for known vegetation characteristics: diameter D, density m, bulk drag coefficient C_D and the water depth h. The bulk drag coefficient for flow through vegetation is a parameter that is difficult to determine, and many researchers have been working on it, for instance Li & Shen (1973) and Neary (2003). In this study we take the theoretical viewpoint that vegetation can be modelled as rigid cylinders and we simply apply a drag coefficient of 1.0, disregarding the placement of the cylinders or the flow Reynolds number.

4 1 DV TURBULENCE MODEL

As argued in section 3, one way of obtaining a detailed description of resistance of flow through and above vegetation, is to perform detailed numerical simulations based on a 1-DV k-ε turbulence model that has been developed by Uittenbogaard (2003). This model is a simplification of the full 3-D Navier-Stokes equations in order to account for horizontal flow conditions only. The model describes vegetation as rigid cylinders, similar to the previous approaches. At the same time and in order to include the effects of vegetation into the k-ε turbulence closure, the following modifications have been included: (i) the decrease of the available cross-section for the vertical exchange of momentum, turbulence kinetic energy and turbulent dissipation, (ii) the drag force exerted by the plants in the horizontal direction, (iii) an additional turbulence production term due to vegetation, and (iv) an additional turbulence dissipation term due to vegetation. The 1-DV model assumes uniform flow in horizontal direction. The momentum equation reads:

$$\rho_0 \frac{\partial u(z)}{\partial t} + \frac{\partial p}{\partial x} = \frac{\rho_0}{1 - A_p(z)} \frac{\partial}{\partial z}\left((1 - A_p(z))\,(\nu + \nu_T(z))\,\frac{\partial u(z)}{\partial z}\right) - \frac{F(z)}{1 - A_p(z)} \tag{20}$$

where F is the drag force of the vegetation per unit volume [Nm^{-3}]:

$$F(z) = \tfrac{1}{2}\rho_0 C_D(z) D(z) m(z) u(z) |u(z)| \tag{21}$$

The pressure gradient is constant along the water depth, according to the hydrostatic pressure assumption, and is either provided as input to the model, or numerically adjusted to satisfy a given

depth-averaged bulk velocity U. The bulk velocity is defined by the volume flux of water divided by the channel's wetted cross-section and relates to the pore velocity by:

$$U = \frac{1}{h}\int_0^h (1 - A_p)u(z)\,\mathrm{d}z \tag{22}$$

The continuity equation is given by:

$$\frac{\partial u}{\partial x} = 0 \tag{23}$$

The k-ε turbulence model provides a closure for the eddy viscosity, relating it to the turbulent kinetic energy, k_T[m^{2}s^{-2}] and its dissipation rate, ε [m^{2}s^{-3}]:

$$\nu_T = c_\mu \frac{k_T^2}{\varepsilon} \tag{24}$$

where c_μ is a constant (with the standard value of 0.09). The k_T-equation in the k-ε turbulence model is modified to take into account the effect of vegetation on the additional production of turbulence and on the vertical diffusion of turbulent kinetic energy:

$$\frac{\partial k_T}{\partial t} = \frac{1}{1 - A_p}\frac{\partial}{\partial z}\left(\left(1 - A_p\right)\left(\nu + \nu_T/\sigma_k\right)\frac{\partial k_T}{\partial z}\right) + T + P_k - B_k - \varepsilon \tag{25}$$

The first term in the right-hand side of Eq. (25) is the vertical diffusion of turbulent kinetic energy by its own mixing action, corrected for the specific area, (1-A_p). A closure coefficient of $\sigma_k = 1.0$ was applied. The second term, T, denotes the additional turbulence generated by the vegetation. Considering fully turbulent flow, all the work done by the fluid against the plants drag force is converted into turbulent kinetic energy, making the expression for T:

$$T\,(z) = F\,(z)\,u\,(z) \tag{26}$$

For transient or laminar flow, part or all of this power would be transferred into heat by work against viscous forces and correction terms depending on Reynolds number would be needed. The third term in the right hand side, P_k, is the turbulence production in shear flows:

$$P_k = \nu_T\left(\frac{\partial u}{\partial z}\right)^2 \tag{27}$$

The fourth term, B_k, is the buoyancy term which represents the conversion of turbulent kinetic energy into gravitational energy according to:

$$B_k = -\frac{\nu_T}{\sigma_k}\frac{g}{\rho_0}\frac{\partial \rho}{\partial z} \tag{28}$$

Finally, the last term, ε, corresponds to the dissipation rate of the turbulent kinetic energy. This is modelled by the ε-equation:

$$\frac{\partial \varepsilon}{\partial t} = \frac{1}{1 - A_p}\frac{\partial}{\partial z}\left(\left(1 - A_p\right)\left(\nu + \nu_T/\sigma_\varepsilon\right)\frac{\partial \varepsilon}{\partial z}\right) + P_\varepsilon - B_\varepsilon - \varepsilon_\varepsilon + c_{2\varepsilon}\frac{T}{\tau_{eff}} \tag{29}$$

The first term in the right hand side of Eq. (29) represents the vertical diffusion of ε by the turbulent eddies. For the closure coefficient, $\sigma_\varepsilon = 1.3$ is applied. The next 3 terms P_ε, B_ε and ε_ε,

correspond to the production, buoyancy and dissipation of ε, respectively, and are related to the production, buoyancy and dissipation of k_T by the expressions:

$$P_\varepsilon = c_{1\varepsilon} \frac{\varepsilon}{k_T} P_k \tag{30}$$

$$B_\varepsilon = c_{1\varepsilon} \left(1 - c_{3\varepsilon}\right) \frac{\varepsilon}{k_T} B_k \tag{31}$$

$$\varepsilon_\varepsilon = c_{2\varepsilon} \frac{\varepsilon^2}{k_T} \tag{32}$$

where $c_{1\varepsilon} = 1.44$, $c_{2\varepsilon} = 1.92$ and $c_{3\varepsilon} = 0$ or 1 (depending on stratification). Universal values for closure coefficients, as derived by Launder and Spalding (1974), have been applied.

The important part is in the last term in the right-hand side, which corresponds to the dissipation rate of turbulence produced by vegetation. This dissipation rate depends on the effective turbulence dissipation time scale (τ_{eff}) and it is affected by the closure coefficient $c_{2\varepsilon}$. To obtain τ_{eff}, Uittenbogaard (2003) related it to the different length scales that control turbulence inside vegetation. First of all, at sufficient distance from the bed as well as from the top of the vegetation, the length scale of internally generated turbulence is smaller than the available fluid space inside the vegetation, and therefore the relevant time scale of this small scale turbulence corresponds to the intrinsic turbulence time scale:

$$\tau_{int} = \frac{k_T}{\varepsilon} \tag{33}$$

This time scale is adopted as an effective time scale by Shimizu & Tsujimoto (1994) and López & García (2001). It is here where the turbulence model of Uittenbogaard (2003) differs from that of previous authors. Uittenbogaard includes the penetration of shear turbulence from above the vegetation into the top layer of the vegetation. Above the vegetation a shear layer is formed by the vertical exchange of horizontal momentum with the retarded flow inside the vegetation. The large eddies that are advected from above the vegetation have to be squeezed into smaller-scale eddies of the length scale of the available pore space inside the vegetation. In this way, the relevant time scale for the dissipation is determined by the geometrical properties of the vegetation, according to:

$$\tau_{geom} = \left(\frac{L_p^2}{c_\mu^2 T} \right)^{\frac{1}{3}} \tag{34}$$

where L_p is the available length scale for eddies between the vegetation:

$$L_p\left(z\right) = c_\ell \left(\frac{1 - A_p\left(z\right)}{m\left(z\right)} \right)^{\frac{1}{2}} \tag{35}$$

in which c_ℓ is a coefficient that affects the geometrical length scale. For cylinders, it is assumed that $c_\ell = 1$, but it is noted that for real vegetation with twigs and foliage c_ℓ might be smaller (Baptist, 2005). The 1-DV model computes both time scales τ_{int} and τ_{geom} over the vertical and evaluates the effective time scale by a MAX-operator:

$$\tau_{eff}^{-1} = \max\left(\tau_{int}^{-1}, \tau_{geom}^{-1}\right) \tag{36}$$

Finally, the representative roughness is obtained from the 1-DV model by calculating the energy slope, i:

$$i = \frac{\partial p / \partial x}{\rho_0 g} \tag{37}$$

and subsequently applying Chézy's relationship. The validity of the 1-DV model has been demonstrated by comparing the outcome with measurements of vertical profiles for flow and turbulence by Meijer & Van Velzen (1999), Nepf & Vivoni (2000) and López & García (2001), and will be demonstrated in a later section of this paper.

5 EQUATION BUILDING WITH GENETIC PROGRAMMING

When refining a model of a physical process, a scientist focuses on the agreement of theoretically predicted and experimentally observed behaviour. If these agree in some accepted sense, then the model is 'correct' within that context. In the process of making sense of experimental data it is generally desirable to express the relation between the variables in a symbolic form: an equation. In this work we consider the problem inverse to verification of theoretical models: how can we obtain the governing equations directly from measurements? To do this, we will apply genetic programming.

5.1 *Dimensionally-aware genetic programming*

Genetic programming (GP) is a technique that can be used to find the symbolic form of an equation, including a set of coefficients. One of the advantages of genetic programming over other methods for regression is the symbolic nature of the solutions that are produced. This is especially pronounced in empirical modelling of unknown phenomena where an underlying theoretical model is not known. For a detailed description of genetic programming from a water resources perspective the interested reader is referred to Babovic & Abbott (1997) and Babovic & Keijzer (2000). Inspired by Koza's (1992) pioneering work on GP and in order to improve performance of his standard approach, an augmented version of GP has been proposed – dimensionally aware GP (Keijzer & Babovic, 1999) – which is arguably more useful in the process of scientific discovery. Dimensionally aware genetic programming (Keijzer & Babovic, 1999) differs from the standard approach in that the raw observations are used together with their units of measurement. The system of units of measurement can be viewed as a typing scheme and as such can be used in some form of typed genetic programming. The dimensionally aware approach proposes what can be called a *weakly typed* or *implicit casting* approach. Here dimensional correctness is not enforced, but promoted. An extra objective for selection, goodness-of-dimension, is introduced that is used next to a goodness-of-fit objective. These two objectives are then used in a multi-objective optimization routine using the concepts of dominance and Pareto optimality. Goodness-of-dimension is measured by calculating how many constants with appropriate units should be introduced to render an equation dimensionally correct. The result of a single run of such unit typed genetic programming is a number of equations – a so-called Pareto front of non-dominated solutions – that balance dimensional correctness (goodness-of-dimension) with goodness-of-fit.

5.2 *Determination of a vegetation-related resistance formula for submerged vegetation (h>k) using genetic programming*

Dimensionally-aware genetic programming was applied to a set of 990 results of the 1-DV model for submerged vegetation. These 990 cases were chosen to represent a wide variety of vegetation properties and water depths for cylinders that represent pioneer species, (stiff) grasses, herbaceous vegetation, reed and bushes. The input variables are presented in Table 1 below.

For purposes of dimensional consistence a slightly adapted Chézy coefficient was used:

$$C' = \frac{C}{\sqrt{g}} \tag{38}$$

By virtue of using such a dimensionless coefficient, time-related units are avoided and the resistance coefficient became solely a function of the geometry of the system. GP was employed

Table 1. Inputs to the 1-DV k-ε model.

Input	Dimension	Description
D	L	Diameter of the stems.
m	L^{-2}	Number of stems per square meter.
k	L	Vegetation height.
C_D	–	Drag coefficient of a single stem.
C_b	$L^{0.5}/T$	Bed Chézy resistance coefficient.
h	L	Water depth.

in a multi-objective sense, so that the following three objectives were simultaneously optimised: (i) root mean square error (RMSE): measure of overall accuracy of the formula, (ii) coefficient of determination (CoD): measure of the goodness of the shape of the formula and (iii) dimensional error: measure of dimensional consistency of the formulae.

The GP induction system was run many times with different parameters using various subsets of the 990 cases. Subsequently the front of non-dominated solutions for all the runs was examined to find a suitable formula. Genetic programming results provided a dimensionally consistent formula that had both smallest RMSE and the highest CoD:

$$\frac{C_r}{\sqrt{g}} = \sqrt{\frac{2}{c_D m D k}} + \ln\left\{\left(\frac{h}{k}\right)^2\right\} \tag{39}$$

This formula can be rearranged to:

$$C_r = \sqrt{\frac{2g}{c_D m D k}} + 2\sqrt{g}\ln\left(\frac{h}{k}\right) \tag{40}$$

Clearly, the relationship of the resistance coefficient C_r and the water depth h consists of a h-independent term plus a h-dependent logarithmic term. For $h = k$, the logarithmic term reduces to zero, making the resistance equivalent to the resistance for flow through non-submerged vegetation, for maximally flooded conditions, $h = k$, see Eq. (12). Having found this, it would be better to have an expression compatible with the more accurate expression for resistance for non-submerged conditions that includes the bed resistance, as well, Eq (11) for fully submerged conditions:

$$C_k = \sqrt{\frac{1}{\frac{1}{C_b^2} + \frac{C_D m D k}{2g}}} \tag{41}$$

Finally, replacing the constant value of 2 with the more theoretically founded Von Kármán's constant $1/\kappa \approx 2.5$, the final expression was determined to be:

$$C_r = \sqrt{\frac{1}{\frac{1}{C_b^2} + \frac{C_D m D k}{2g}}} + \frac{\sqrt{g}}{\kappa}\ln\left(\frac{h}{k}\right) \tag{42}$$

Although fractionally more complicated than the original formulation, expression (42) was found to be in good agreement with the data, especially in regions with higher Chézy values (Figure 3d). Furthermore, it is theoretically well founded, combining the resistance for the flow inside the vegetation with the observed logarithmic profile above the vegetation.

Subsequent research on the formulation revealed an agreement with the work by Kouwen *et al*. (1969), where a general formula for resistance induced by vegetation was proposed as:

$$C_r = C_1 + \frac{\sqrt{g}}{\kappa} \ln\left(\frac{h}{k}\right) \tag{43}$$

Kouwen *et al*. proposed several relationships for C_1, but no definitive conclusions were drawn. Equation (42) gives an exact formulation for C_1 and as such presents a step forward in the modelling of resistance. Note that the final formulation is a combination of a computer induced expression that fits the data well, and theoretically based modifications to fit the theory. Further analysis reveals that the equation is equivalent to the equation produced by the method of effective water depth, Eq. (16) if, (i) the depth balance is ignored (*i.e*., the factors k and *(h-k)* are considered equal to h), and (ii) the roughness length z_0 is set to k/e where e is the base of the natural logarithm. This leads to the assumption that the resistance is mainly governed by a logarithmic flow profile over the submerged depth, with a boundary condition at the level of zero intercept, which is at the top of the vegetation. Note further that Eq. (42) is obtained by integrating the differential equation:

$$\frac{\partial C_r}{\partial h} = \frac{\sqrt{g}}{\kappa h} \tag{44}$$

with the boundary condition of $C = C_{ref}$ at $h = h_{ref}$. In this case the logical boundary condition is $h_{ref} = k$.

6 COMPARISON OF THE FORMULATIONS

There are several ways to evaluate the formulations. First and foremost is the ability to model the data under study. This is evaluated by comparing the results of the analytical formulations with those of the 1-DV numerical model. Evaluation parameters are the Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_{Oi} - x_{Mi})^2} \tag{45}$$

where N is the number of observations, and x_{Oi} and x_{Mi} are the observed and modelled values, respectively, and the Coefficient of Determination (CoD), which in this linear regression case equals R^2, where R is the correlation coefficient. Results are presented in Table 2. It can be seen that the expressions based on the genetic programming results are in better agreement with the synthetic dataset than the manually induced formulations. There is a small level of systematic error, but the GP-formulae model the variance admirably. Even though the modified GP-formula has a higher RMSE than the original GP-formula, the scatter plot in Figure 3 reveals that the improved formulation removes the large residuals associated with high Chézy values, and thus is more reliable over the entire domain than the original formulation.

Up to this point, both training and comparisons were performed on synthetic data, generated by the 1-DV model. To ultimately test our approaches, 177 experimental runs based on laboratory flume experiments were collected from 10 independent studies. These data were not used in the equation building process, but kept aside to validate the equations induced.

These studies include rigid and flexible, artificial and natural vegetation types from Kouwen *et al*. (1969), Ree & Crow (1977), Murota *et al*., (1984), Tsujimoto & Kitamura (1990), Tsujimoto *et al*., (1993), Ikeda & Kanazawa (1996), López & García (1997), Meijer (1998a), Meijer (1998b), López & García (2001) and Järvelä (2003). The data contained all input information, except the bed roughness C_b, which was assumed to be negligible in the experiments for smooth flume beds,

Table 2. Root Mean Squared Errors (RMSE) and Coefficients of Determination (CoD) for the four formulations of the resistance coefficients, including the 1-DV model which was used for generation of training data, compared with the 1-DV model results.

Equation	RMSE 1-DV Data [$m^{1/2}s^{-1}$]	CoD 1-DV Data[-]
Method of effective water depth, Eq. (16)	1.3048	0.97485
Analytical solution method, Eq. (19)	1.4705	0.93724
Original GP-formula, Eq. (39)	0.9746	0.97065
Modified GP-formula, Eq. (42)	1.2061	0.97579
1-DV numerical model	0	1.00

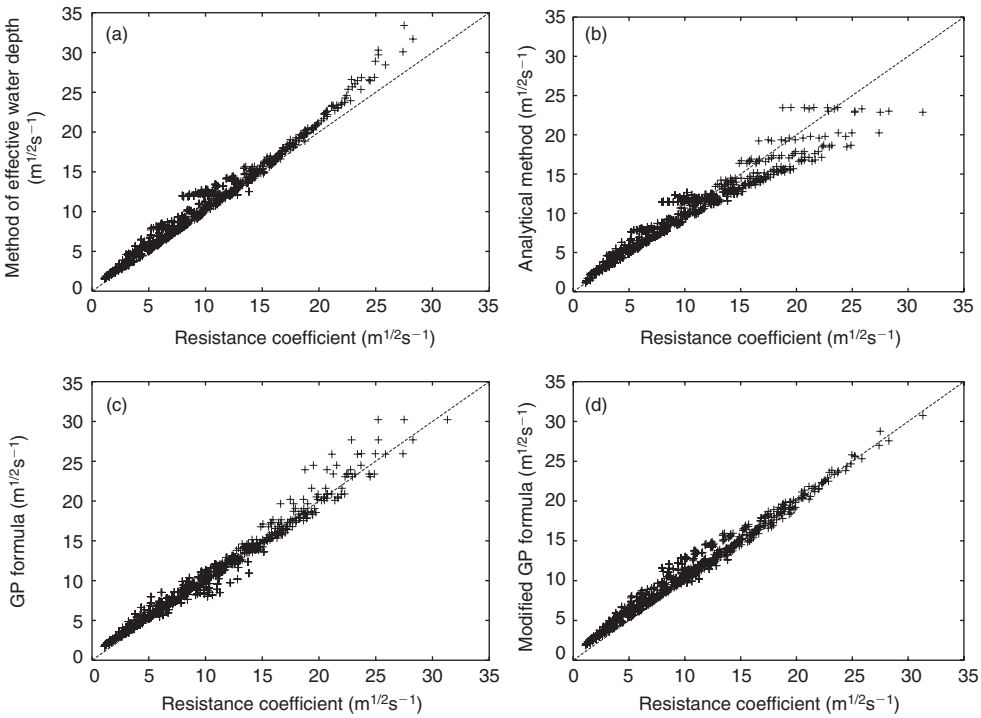


Figure 3. Scatter plots for the four equations on 1-DV data. (a) upper left: Method of effective water depth, Eq. (16), (b) upper right: Analytical solution method, Eq. (19), (c) lower left: Original GP-formula, Eq. (39) and (d) lower right: Modified GP-formula, Eq. (42).

whereupon it was set to $60\,m^{1/2}s^{-1}$. In cases where the drag coefficient was not defined it was assumed to be 1.0. For flexible vegetation experiments, the deflected height was applied in the formulations. An overview of the data of the 177 experiments is given in Appendix A of Baptist (2005). It is noted that although the range in parameter values in these experiments is quite large, in comparison with natural vegetation types, flume experimental data are not covering the full range of existing types.

For this particular dataset it was possible to also test the performance of the 1-DV model itself. Results for the comparison can be found in Table 3 and Figure 4, and it can be seen that the genetic programming induced equations give a highly competitive agreement with the data. What is particular startling is that their performance is even competitive with the 1-DV model the equations are based upon.

Table 3. Root Mean Squared Errors (RMSE) and Coefficients of Determination (CoD) for the four formulations of the resistance coefficients, including the 1-DV model which was used for generation of training data, compared with flume experiments.

Equation	RMSE flume experiments [$m^{1/2}s^{-1}$]	CoD flume experiments[-]
Method of effective water depth, Eq. (16)	2.7187	0.83594
Analytical solution method, Eq. (19)	2.2796	0.81325
Original GP-formula, Eq. (39)	2.1093	0.83737
Modified GP-formula, Eq. (42)	2.1826	0.87418
1-DV numerical model	1.8600	0.87300

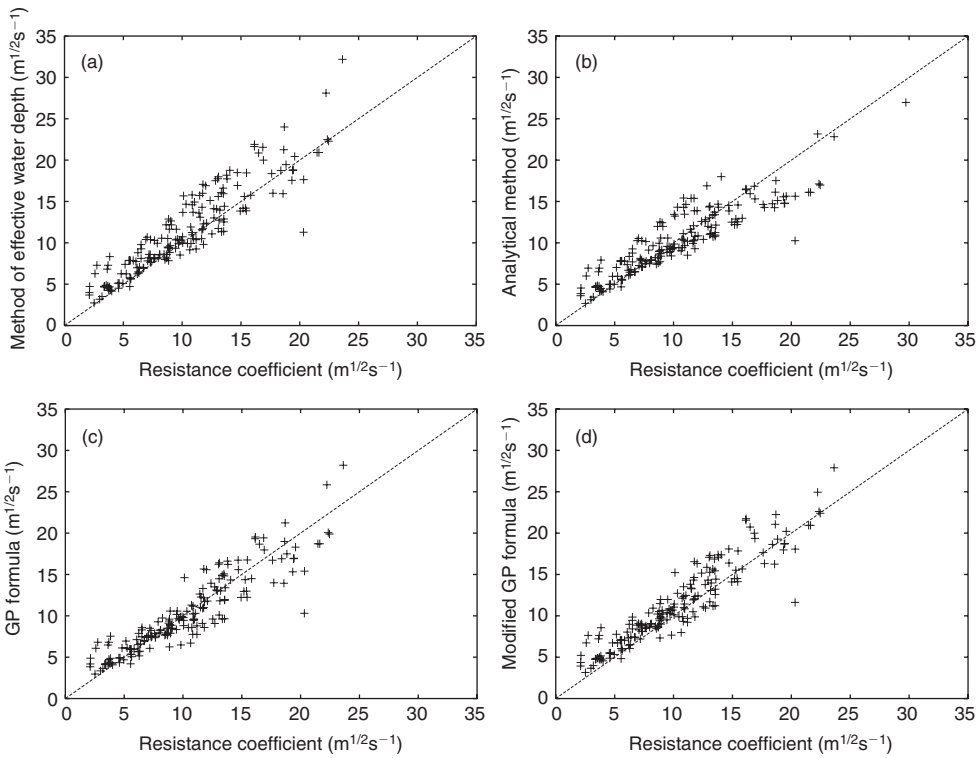


Figure 4. Scatter plots for the four equations for experimental flume data. (a) upper left: Method of effective water depth, Eq. (16), (b) upper right: Analytical solution method, Eq. (19), (c) lower left: Original GP-formula, Eq. (39) and (d) lower right: Modified GP-formula, Eq. (42).

Finally, Figure 5 presents a comparison between the improved genetic programming equation and the original 1-DV model, both applied to the validation set of flume experiments. No serious discrepancies between the dynamical model and the simple equation are observed.

From the perspective of simplicity of the equations, it might be enlightening to compare the apparatus of expressions (Eqs. (4) to (37)) that leads to the definition of Eqs. (16) and (19), with the conciseness and elegancy of the genetic programming induced Eq. (39) and its human-manipulated variant, Eq. (42). The genetic programming induced equations are based purely on the 1-DV model data, and yet ignore considerations about turbulence length scales and turbulence intensity. They

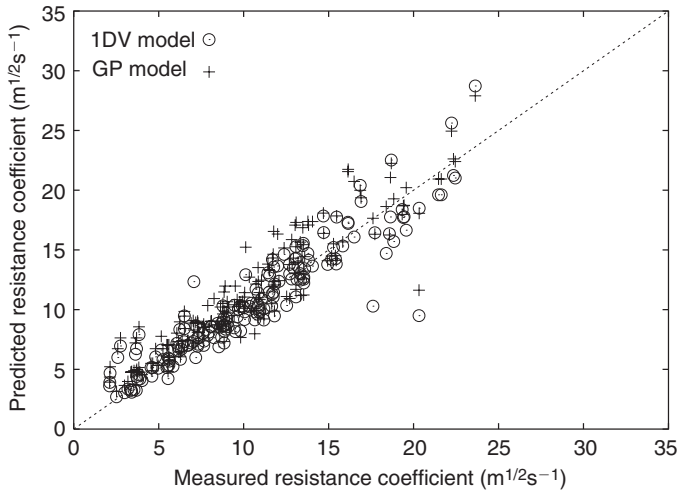


Figure 5. Direct comparison between the 1-DV model (circles) and the modified GP-formula, Eq. (42) (plusses), using experimental flume data (i.e. out of sample validation).

focus primarily on obtaining good agreement with the data, and in this case, the problematic interplay between the vegetation and the turbulence induced by the vegetation apparently are of secondary importance to the simple logarithm on the ratio of water depth over plant height.

7 DISCUSSION

This paper reports the comparison of four different methods for calculating flow resistance due to vegetation. In particular, a GP algorithm showed to induce an appropriate equation for vegetation resistance from the output of a 1-DV numerical model. In comparison to a variety of flume data, a modified version of this equation is shown to provide the best fit with minimised residuals over the whole data range. The induced equation is sufficiently simple to be applied to wide-area depth-averaged models, to calculate, for example, flow over vegetated floodplains.

The main limitation of the study is the assumption that vegetation can be represented as rigid cylinders. In vegetated floodplains, several vegetation types, such as herbs and grasses will bend due to the force of the flow. Bending of vegetation decreases the frontal area for drag, introduces lift forces and furthermore, swaying vegetation may introduce additional resistance. It has been found that for rigid vegetation the linear relation between force and the square of velocity holds, as depicted in this paper, but for flexible plants a linear increase of drag force with flow velocity is observed (Armanini *et al*., 2005). This implies that the equation for resistance found is less suitable for flexible vegetation, although the comparison with flume data, which included flexible vegetation as well, showed a reasonable fit.

Testing expressions for their ability to model a phenomenon such as resistance induced by vegetation requires experimental data. Particularly when using data driven methods, data are needed for steering the error minimisation process. For equations induced by scientists such data is needed to test the proposed equation for its capacity to model the phenomenon under study.

Even though the variables for this relationship should be measurable in wetlands and vegetated floodplains, this cannot yet be accomplished for resistance coefficients measured at a fine scale, at different water depths and with different types of vegetation. Obtaining data at such a fine scale in realistic circumstances is prohibitive in terms of effort and cost. Instead, we here chose to use a fine-scaled numerical 1-DV model, describing turbulence and flow properties to generate data.

Such a model employs available knowledge about characteristics of vegetation, turbulence induced by vegetation, and resistance caused by the drag forces on the plants. Determining values of the resistance coefficients from such a microscopic model is trivial, and can be understood as a noise free approximation of the phenomenon under study. However, a compact expression describing the phenomenon is not readily available from a numerical 1-DV model.

It could be argued that using synthetic data defeats the purpose of finding an equation. If a dynamic model exists, why not simply use that one instead of going the laborious route of finding a compact equation. The purpose of finding a compact equation lies in the type of modelling that it enables. Vegetation resistance is a typical three-dimensional problem due to the spatial heterogeneity of vegetation and the water depth dependency of submerged or non-submerged vegetation. A full dynamical model thus operates on a 3-D grid, which is computationally expensive. An analytical solution to the problem of resistance induced by vegetation, which includes water depth dependency, makes 2-dimensional, depth-averaged modelling possible, allowing for faster model computations and the possibility to apply the model to larger areas. In addition, an analytical expression can also be used in 1-DH computations, or even in spreadsheet models.

A difficulty of the approach, even in simple models, is that floodplains have inhomogeneous vegetation types with a complex structure (Baptist et al., 2005). This shifts the modelling problem from estimating an equivalent bed roughness value to estimating plant properties for density, diameter, height and drag coefficient. In principle it is possible, although not easy, to obtain the complete horizontal and vertical structure of grasses, herbaceous vegetation, bushes, floodplain forests and other vegetation types. The drag coefficient, however, is a property that cannot be measured directly in the field. In principle, the drag coefficient for smooth cylinders is known from experimental studies and theory and is dependent on the value for the Reynolds number of the flow and the spatial arrangement of the cylinders. In applications on vegetation roughness this coefficient is usually estimated or used as a calibration parameter, for example in an effort to account for rough surfaces or for the foliage. Given Eq. (42), and calibrated on the drag coefficient, it's main advantage over just calibrating a value for C_r directly, is that it gives a relationship with water depth.

8 CONCLUSIONS

Four formulations for water depth-related resistance induced by vegetation were studied and compared. Two of the equations studied here were created through analysis and a process of derivation by a scientist. One equation was derived by dimensionally aware genetic programming, and finally one expression was created by manually analysing and improving the genetic programming equation. It was found that the genetic programming equations were superior to the manually derived equations, both on their performance on synthetic training and laboratory testing data, and in the economy of detail that needs to be modelled. The manually improved expression was found to be in good agreement with previous expressions found in the literature, and performed competitive on experimental data with the detailed numerical model that was used to generate the training data.

This paper presented a case study in the use of genetic programming as a hypothesis generator for use in scientific discovery. Not only does it show that genetic programming is capable of producing equations that are comparable or perhaps better than their human derived competitors, it produces expressions that are amenable to further analysis and manual improvement. This is potentially a much more useful result, as it shows that the symbolic nature of genetic programming can be used to build up knowledge in a problem domain. In contrast with many machine learning algorithms, where the trained model is the end result of a problem statement, the genetic programming induced expressions can be used to start a new cycle of inquiry.

The elegant Eq. (42) that was induced with genetic programming is both theoretically and experimentally justified, and can be used to estimate the resistance coefficient of submerged vegetation.

REFERENCES

Armanini, A., Righetti, M. and Grisenti, P. (2005). "Direct measurement of vegetation resistance in prototype scale", *Journal of Hydraulic Research*, 43(5): 481–487.

Babovic, V. and Abbott, M.B. (1997). "The evolution of equation from hydraulic data, Part I: Theory and Part II: Application", *Journal of Hydraulic Research*, 35(3):397–430.

Babovic, V. and Keijzer, M. (2000). "Genetic programming as a model induction engine." *Journal of Hydroinformatics*, 2(1):35–60.

Baptist, M.J. (2005). "Modelling floodplain biogeomorphology," Ph.D. thesis, ISBN 90-407-2582-9, 213 pp., Delft University of Technology, Faculty of Civil Engineering and Geosciences, Section Hydraulic Engineering, Delft.

Baptist, M.J., Van den Bosch, L.V., Dijkstra, J.T. and Kapinga, S. (2005). Modelling the effects of vegetation on flow and morphology in rivers. *Archiv für Hydrobiologie Supplement,* 155/1–4, *Large Rivers*, 15(1–4): 339–357.

Campana, T.R. (1999). "Hydraulic resistance of submerged floodplain vegetation," M.Sc. thesis H.E.043, IHE-Delft, Delft.

Carollo, F.G., Ferro, V. and Termini, D. (2002). "Flow Velocity Measurements in Vegetated Channels." *Journal of Hydraulic Engineering*, 128(7), 664–673.

Dawson F.H. and Charlton, F.G. (1988). "Bibliography on the hydraulic resistance of vegetated watercourses." Technical report, Freshwater Biological Association, Occasional Publication No. 25, ISNN 0308–6739, 25 pp.

Darby, S.E. (1999). "Effect of Riparian Vegetation on Flow Resistance and Flood Potential." *Journal of Hydraulic Engineering*, 125(5), 443–445.

Einstein, H.A. and Banks, R.B. (1950). "Fluid resistance of composite roughness." *Transactions of the Amer. Geoph. Union*, 31(4), 603–610.

Fathi-Maghadam, M. and Kouwen, N. (1997). "Nonrigid, Nonsubmerged, Vegetative Roughness on Floodplains." *Journal of Hydraulic Engineering*, 123(1), 51–57.

Fischer-Antze, T., Stoesser, T., Bates, P. and Olsen, N.R.B. (2001). "3D numerical modelling of open-channel flow with submerged vegetation." *Journal of Hydraulic Research*, 39(3), 303–310.

Garratt, J. (1992). *The atmospheric boundary layer*, Cambridge University Press, New York, NY., 316 pp.

Ghisalberti, M. and Nepf, H.M. (2002). "Mixing layers and coherent structures in vegetated aquatic flows." *Journal of Geophysical Research*, 107(C2), 3-1 to 3-11.

Ghisalberti, M. and Nepf, H.M. (2004). "The limited growth of vegetated shear layers." *Water Resources Research*, 40(doi:10.1029/2003WR002776, 2004).

Hasegawa, K., Asai, S., Kanetaka, S. and Baba, H. (1999). "Flow properties of a deep open experimental channel with a dense vegetation bank." *Journal of Hydroscience and Hydraulic Engineering*, 17(2), 59–70.

Hoffmann, M. (2004). "Application of a simple space-time averaged porous media model to flow in densely vegetated channels." *J. of Porous Media*, 7(3), 183–191.

Hong, H.T.M. (1995). Hydraulic Resistance of Flexible Roughness. M.Sc thesis H.H.237, IHE Delft.

Ikeda, S. and Kanazawa, M. (1996). "Three-dimensional organized vortices above flexible water plants." *Journal of Hydraulic Engineering*, 122(11), 634–640.

Ikeda, S., Yamada, T. and Toda, Y. (2001). "Numerical study on turbulent flow and honami in and above flexible plant canopy." *International Journal of Heat and Fluid Flow*, 22, 252–258.

Jackson, P.S. (1981). "On the displacement height in the logarithmic velocity profile." *Journal of fluid mechanics*, 111, 15–25.

James, C.S., Birkhead, A.L., Jordanova, A.A. and O'Sullivan, J.J. (2004). "Flow resistance of emergent vegetation." *Journal of Hydraulic Research*, 42(4), 390–398.

Järvelä, J. (2002). "Flow resistance of flexible and stiff vegetation: a flume study with natural plants." *Journal of Hydrology*, 269(1–2), 44–54.

Järvelä, J. (2003). "Influence of vegetation on flow structure in floodplains and wetlands." *River, Coastal and Estuarine Morphodynamics 2003*, Barcelona, Spain, 845–856.

Järvelä, J. (2004). "Determination of flow resistance caused by non-submerged woody vegetation." *Int. J. River Basin Management*, 2(1), 1–10.

Katul, G., Wiberg, P., Albertson, J. and Hornberger, G. (2002). "A mixing layer theory for flow resistance in shallow streams." *Water Resources Research*, 38(11), 1250, doi:10.1029/2001WR000817, 2002.

Keijzer, M. and Babovic, V. (1999). "Dimensionally aware genetic programming." Banzhaf, W., Daida, J., Eiben, A.E., Garzon, M.H., Honavar, V., Jakiela, M. and Smith, R.E. (Eds), *Proceedings of the Genetic*

and Evolutionary Computation Conference, volume 2, pages 1069–1076, Orlando, Florida, USA, 13–17 July 1999. Morgan Kaufmann.

Klaassen, G.J. and Van der Zwaard, J.J. (1974). "Roughness coefficients of vegetated floodplains." *Journal of Hydraulic Research*, 12(1), 43–63.

Klopstra, D., Barneveld, H.J., Van Noortwijk, J.M. and Van Velzen, E.H. (1997). "Analytical model for hydraulic roughness of submerged vegetation." *27th International IAHR Conference*, San Fransico, 775–780.

Kouwen, N., Unny, T.E. and Hill, H.M. (1969). "Flow retardance in vegetated channels." *Journal of the Irrigation and Drainage Division*, 95(IR2), 329–342.

Kouwen, N. and Unny, T.E. (1973). "Flexible roughness in open channels." *Journal of the Hydraulics Division*, 99(HY5), 713–728.

Koza, J.R., (1992) *Genetic Programming: On the Programming of the Computes by Means of Natural Selection*, MIT Press.

Launder, B. and Spalding, D. (1974). "The numerical computation of turbulent flows." *Computer Methods in Applied Mechanics and Engineering* 3, 269–289.

Li, R.M. and Shen, H.W. (1973). "Effect of Tall Vegetations on Flow and Sediment." *Journal of the Hydraulics Division*, 99(HY6), 1085–1103.

López, F. and García, M.H. (1997). "Open-channel Flow Through Simulated Vegetation: Turbulence Modelling and Sediment Transport." *Technical Report WRP-CP-10*, Hydrosystems Laboratory, University of Illinois, Urbana.

López, F. and García, M.H. (2001). "Mean Flow and Turbulence Structure of Open-Channel Flow through Non-Emergent Vegetation." *Journal of Hydraulic Engineering*, 127(5), 392–402.

Mason, D.C., Cobby, D.M., Horritt, M.S. and Bates, P.D. (2003). "Floodplain friction parameterization in two-dimensional river flood models using vegetation heights derived from airborne scanning laser altimetry." *Hydrol. Process.*, 17, 1711–1732.

Meijer, D.M. (1998a). "Flume studies of submerged vegetation." *PR121.10*, HKV, Lelystad (in Dutch).

Meijer, D.M. (1998b). "Flume studies of submerged reed." *PR177*, HKV, Lelystad (in Dutch).

Meijer D.G. and Van Velzen, E.H. (1999). "Prototype-scale flume experiments on hydraulic roughness of submerged vegetation." *28th International IAHR Conference*, Graz.

Murota, A., Fukuhara, T. and Sato, M. (1984). "Turbulence structure in vegetated open channel flows." *Journal of Hydroscience and Hydraulic Engineering*, 2(1), 47–61.

Neary, V.S. (2003). "Numerical Solution of Fully Developed Flow with Vegetative Resistance." *Journal of Engineering Mechanics*, 129(5), 558–563.

Nepf, H.M. and Vivoni, E.R. (2000). "Flow structure in depth-limited, vegetated flow." *Journal of Geophysical Research*, 105(C12), 28,547–28,557.

Nezu, I. and Naot, D. (1999) "Partly Vegetated Open Channels; New Experimental Evidence." *28th International IAHR Conference*, Graz.

Nikuradse, J. (1930). "Turbulente Strömung in nichtkreisförmigen Rohren." *Ing.-Arch.*, 1(306).

Petryck, S. and Bosmajian III, G. (1975). "Analysis of flow through vegetation." *Journal of the Hydraulics Division*, 101(HY7), 871–884.

Ree, W.O. and Crow, F.R. (1977). "Friction factors for vegetated waterways of small slope." *ARS-S-151*, Agricultural Research Service, U.S. Department of Agriculture, Stillwater, Oklahoma, U.S.A.

Righetti, M. and Armanini, A. (2002). "Flow resistance in open channel flows with sparsely distributed bushes." *Journal of Hydrology*, 269(1–2), 55–64.

Shimizu, Y. and Tsujimoto, T. (1994). "Numerical analysis of turbulent open-channel flow over vegetation layer using a k-ε turbulence model." *J. Hydrosci. Hydr. Eng.*, 11(2), 57–67.

Stephan, U. and Gutknecht, D. (2002). "Hydraulic resistance of submerged flexible vegetation." *Journal of Hydrology*, 269(1–2), 27–43.

Stone, B.M. and Shen, H.T. (2002). "Hydraulic Resistance of Flow in Channels with Cylindrical Roughness." *Journal of Hydraulic Engineering*, 128(5), 500–506.

Taylor, R.P., Coleman, H.W. and Hodge, B.K. (1985). "Prediction of turbulent rough-wall skin friction using a discrete element approach." *Journal of Fluids Engineering*, 107, 251–257.

Thom, A.S. (1971). "Momentum absorption by vegetation." *Q.J.R. Meteorol. Soc.*, 97, 414–428.

Tsujimoto, T. and Kitamura, T. (1990). "Velocity profile of flow in vegetated-bed channels." *KHL progressive report 1*, Hydraulic Laboratory, Kanazawa University.

Tsujimoto, T., Okada, T., and Kontani, K. (1993). "Turbulent structure of open-channel flow over flexible vegetation." *KHL progressive report 4*, Hydraulic Laboratory, Kanazawa University.

Uittenbogaard, R. (2003). "Modelling turbulence in vegetated aquatic flows." *International workshop on RIParian FORest vegetated channels: hydraulic, morphological and ecological aspects*, 20–22 February 2003, Trento, Italy.

Van Velzen, E.H., Jesse, P., Cornelissen, P. and Coops, H. (2003). "Hydraulic resistance of vegetation in floodplains, Part 2 Background document version 1-2003." Ministry of Transport, Public Works and Water Management, Institute for Inland Water Management and Waste Water Treatment, report 2003.029. Lelystad, the Netherlands (in Dutch).

Vernon-Harcourt, L.F. (1896). *Rivers and Canals, Vol. 1 Rivers*, the Clarendon Press, Oxford, 312pp.

Wilson, C.A.M.E., Stoesser, T., Bates, P.D. and Batemann Pinzen, A. (2003). "Open Channel Flow through Different Forms of Submerged Flexible Vegetation." *Journal of Hydraulic Engineering*, 129(11), 847–853.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Modeling subsurface water flow and solute transport with HYDRUS and related numerical software packages

J. Šimůnek
Department of Environmental Sciences, University of California Riverside, Riverside, USA

M.Th. van Genuchten
US Salinity Laboratory, Riverside, USA

M. Šejna
PC-Progress, Ltd., Prague, Czech Republic

ABSTRACT: HYDRUS and related software packages represent a system of numerical models for simulating variably-saturated water, heat and solute movement in the subsurface. The software packages can simulate flow and transport in one-dimensional (e.g., HYDRUS-1D, HP1, SOILCO2, and UNSATCHEM), two-dimensional (SWMS_2D, CHAIN_2D, or HYDRUS-2D), and three-dimensional (SWMS_3D or HYDRUS (2D/3D)) transport domains. While water flow in all of the models is simulated in a relatively similar manner, different solute transport modules exist in the different codes. Some models can simulate only transport of solutes independent of one another, whereas other models include biogeochemical modules simulating complex interactions between the various solutes. For example, HYDRUS-1D and UNSATCHEM models can also consider the transport of major ions and their mutual reactions, such as dissolution-precipitation, cation exchange, or aqueous complexation. The wetland module in HYDRUS (2D/3D) furthermore considers biochemical processes associated with the flow of wastewater through constructed or natural wetlands. Finally, HP1, a model resulting from coupling HYDRUS-1D and PHREEQC, accounts for a broad range of instantaneous and kinetic chemical and biological reactions, including aqueous complexation, precipitation-dissolution, cation exchange, surface complexation, and redox reactions. In this manuscript we review the various models and provide references to many of their applications.

1 INTRODUCTION

We describe in this paper a number of numerical models for simulating water flow and movement of heat and various solutes in the subsurface. The models were developed during the past 15 years or more in close collaboration between the University of California, Riverside, the U.S. Salinity Laboratory in Riverside, PC-Progress in Prague, Czech Republic, and SCK•CEN, Mol, Belgium. This manuscript serves as a companion paper to an article that will appear in the upcoming "Vadose Zone Modeling" Special Section of Vadose Zone Journal (Šimůnek et al., in press). While that article qualitatively describes the various conceptual processes forming the basis of the different programs, here we focus on the implemented mathematical models. We provide the governing equations for the various water flow, heat transport, and solute transport modules, and discuss into which HYDRUS programs they were implemented. We additionally discuss the numerical techniques used to discretize the governing flow and transport equations, and the importance of

Table 1. Synopsis of the HYDRUS suite of windows-based software packages.

Model	Ver	Dim	Brief description	Reference
HYDRUS-1D	3.0	1D	Water flow, heat and solute transport, carbon dioxide transport, transport of major ions, major ion chemistry, transport of colloids and bacteria.	Šimůnek et al. (2005)
HP1	1.0	1D	Water flow, transport of heat and multiple components, mixed equilibrium/kinetic biogeochemical reactions.	Jacques & Šimůnek (2005)
HYDRUS-2D	2.0	2D	Water flow, heat and solute transport (discontinued, replaced with HYDRUS (2D/3D)).	Šimůnek et al. (1999)
HYDRUS (2D/3D)	1.0	2/3D	Water flow, heat and solute transport, transport of colloids and bacteria wetland module.	Šimůnek et al. (2006b), Šejna & Šimůnek (2007)

Ver – version, OS – operating system, Dim – dimensions, DOS – Disc Operating System, 1D – one-dimensional, 2D – two-dimensional and axisymetrical three-dimensional, 3D – three-dimensional.

the developed graphical user interfaces to ensure wide acceptance of the software packages. We further discuss selected applications of the HYDRUS programs to demonstrate different features of the codes. Our primary focus is on the programs listed in Table 1, i.e., HYDRUS-1D, HP1, HYDRUS-2D, and HYDRUS (2D/3D).

2 WATER FLOW

2.1 *Uniform flow*

Uniform variably-saturated water flow in all of these models is described using the Richards equation:

$$\frac{\partial \theta(h)}{\partial t} = \frac{\partial}{\partial x_i}\left[K(h)\left(K_{ij}^A \frac{\partial h}{\partial x_j} + K_{iz}^A\right)\right] - S(h) \tag{1}$$

where θ = volumetric water content [L^3L^{-3}]; h = pressure head [L]; K = unsaturated hydraulic conductivity [LT^{-1}]; K_{ij}^A = components of a dimensionless anisotropy tensor $\mathbf{K}^A$ (which reduces to the unit matrix when the medium is isotropic); S = general sink/source term [$L^3L^{-3}T^{-1}$] accounting for root water uptake; t = time [T]; and x_i = spatial coordinate [L].

Because of strongly nonlinear makeup of the Richard equation, only a relatively few simplified analytical solutions can be derived. Most practical applications of (1) require a numerical solution, which can be obtained using a variety of methods such as finite differences or finite elements. Equation (1) is generally referred to as the mixed form of the Richards equation since it contains two dependent variables (i.e., the water content and the pressure head). Solutions of (1) require knowledge of two soil hydraulic functions describing the soil water retention and hydraulic conductivity properties of the porous material. The soil water retention curve (or also called the soil water characteristic curve, the capillary pressure-saturation relationship, or the pF curve) describes the relationship between the water content and the pressure head. The HYDRUS packages permit the use of analytical functions suggested by Brooks & Corey (1964), van Genuchten (1980), Kosugi

(1996), and Durner (1994):

$$S_e(h) = \frac{\theta - \theta_r}{\theta_s - \theta_r} = \begin{cases} \left(\dfrac{h_e}{h}\right)^{\lambda} \\ \dfrac{1}{[1 + |\alpha h|^n]^m} \\ \dfrac{1}{2} erfc\left\{\dfrac{\ln(h/h_m)}{\sqrt{2}\sigma}\right\} \\ \displaystyle\sum_i \dfrac{w_i}{[1 + |\alpha_i h|^{n_i}]^{m_i}} \end{cases} \tag{2}$$

respectively, where S_e = effective saturation [-]; θ_r and θ_s = residual and saturated water contents [L^3L^{-3}], respectively; h_e = air-entry value [L]; λ = pore-size distribution index that characterizes the width of the pore-size distribution (or the steepness of the retention function); α [L^{-1}], n [-] and $m = 1/1/n$ [-] = empirical shape parameters; h_m = median pressure head [L] for which $S_e(h_m) = 0.5$ and corresponding to the median pore radius; σ = standard deviation of the log-transformed soil pore radius [-] characterizing the width of the pore-size distribution; and w_i = weighting factors for the overlapping regions.

Equation (1) contains the hydraulic conductivity function which characterizes the ability of a soil to transmit water, and as such is inversely related to the resistance to water flow. Similarly as for the soil water retention curve, HYDRUS allows analytical models for the hydraulic conductivity function as suggested by Brooks & Corey (1964), van Genuchten (1980), Kosugi (1996), and Durner (1994):

$$K(S_e) = K_s \begin{cases} S_e^{2/\lambda + l + 1} \\ S_e^l [1 - (1 - S_e^{1/m})^m]^2 \\ S_e^l \left\{\dfrac{1}{2} erfc\left[\dfrac{\ln(h/h_m)}{\sqrt{2}\sigma} + \sigma\right]\right\}^2 \\ \dfrac{(w_1 S_{e_1} + w_2 S_{e_2})^l (w_1 \alpha_1 [1 - (1 - S_{e_1}^{1/m_1})^{m_1}] + w_2 \alpha_2 [1 - (1 - S_{e_2}^{1/m_2})^{m_2}])^2}{(w_1 \alpha_1 + w_2 \alpha_2)^2} \end{cases} \tag{3}$$

respectively, where K_s = saturated hydraulic conductivity [LT^{-1}]; and l = pore-connectivity parameter [-].

2.2 *Nonequilibrium flow*

The Richards equation (1) predicts a uniform flow process in the vadose zone, although possibly modified macroscopically by allowing the use of spatially variable soil hydraulic properties (e.g., as dictated by different soil horizons, or varying laterally). Unfortunately, the vadose zone can be extremely heterogeneous at a range of scales, from the microscopic (e.g., pore scale) to the macroscopic (e.g., field scale). Some of these heterogeneities can lead to a preferential flow process that is very difficult to capture macroscopically with the standard Richards equation. One obvious example of preferential flow is the rapid movement of water and dissolved solutes through macropores (e.g., between soil aggregates, or caused by earthworm pathways or decayed root channels) or within rock fractures, with much of the water bypassing (short-circuiting) the soil or rock matrix. However, many other causes of preferential flow exist, such as flow instabilities caused by soil textural changes or water repellency (Hendrickx & Flury 2001; Šimůnek et al. 2003), as well as lateral funneling of water due to inclined or other textural boundaries (e.g., Kung 1990).

Preferential flow in macroporous soils and fractured rocks can be described using a variety of dual-porosity or dual-permeability models (Gerke & van Genuchten 1993; Šimůnek et al. 2003).

Dual-porosity and dual-permeability models both assume that the porous medium consists of two interacting pore regions, one associated with the inter-aggregate, macropore, or fracture system, and one comprising the micropores (or intra-aggregate pores) inside soil aggregates or the rock matrix. Dual-porosity models assume that water in the matrix is stagnant, while dual-permeability models allow also for water flow within the soil or rock matrix.

2.2.1 *Dual-porosity model*

The dual-porosity models implemented in the HYDRUS-1D and HYDRUS (2D/3D) packages assume that water flow is restricted to the macropores (or the inter-aggregate pores and fractures), and that water in the matrix (the intra-aggregate pores or the rock matrix) does not move at all. This conceptualization partitions the liquid phase into mobile (flowing, inter-aggregate), θ_{mo}, and immobile (stagnant, intra-aggregate), θ_{im}, regions [L^3L^{-3}]:

$$\theta = \theta_{mo} + \theta_{im} \tag{4}$$

The dual-porosity formulation for water flow is then based on a mixed formulation of the Richards equation (1) to describe water flow in the macropores (the preferential flow pathways) augmented with a mass balance equation to describe the moisture dynamics in the matrix as follows (Šimůnek et al. 2003):

$$\begin{aligned} \frac{\partial \theta_{mo}(h_{mo})}{\partial t} &= \frac{\partial}{\partial x_i}\left[K(h_{mo})\left(K_{ij}^A \frac{\partial h_{mo}}{\partial x_j} + K_{iz}^A\right)\right] - S_{mo}(h_{mo}) - \Gamma_w \\ \frac{\partial \theta_{im}(h_{im})}{\partial t} &= - S_{im}(h_{im}) + \Gamma_w \end{aligned} \tag{5}$$

where h_{im} and h_{mo} = pressure heads for both regions [T^{-1}]; S_{im} and S_{mo} = sink terms for both regions [T^{-1}]; and Γ_w = transfer rate for water from the inter- to the intra-aggregate pores [T^{-1}].

Examples of applications of dual-porosity models to a range of laboratory and field data involving transient flow and solute transport are given by Šimůnek et al. (2001), Castiglione et al. (2003), Zhang et al. (2004), Köhne et al. (2004a, 2005), Kodešová et al. (2005) and Haws et al. (2005).

2.2.2 *Dual-permeability model*

Research versions of HYDRUS-1D and the two-dimensional computational module of HYDRUS (2D/3D) implement the dual-permeability approach suggested by Gerke & van Genuchten (1993). This approach applies Richards equations to each of two pore regions. The flow equations for the macropore or fracture (subscript f) and matrix (subscript m) pore systems in this approach are given by:

$$\begin{aligned} \frac{\partial \theta_f(h_f)}{\partial t} &= \frac{\partial}{\partial x_i}\left[K_f(h_f)\left(K_{ij}^A \frac{\partial h_f}{\partial x_j} + K_{iz}^A\right)\right] - S_f(h_f) - \frac{\Gamma_w}{w} \\ \frac{\partial \theta_m(h_m)}{\partial t} &= \frac{\partial}{\partial x_i}\left[K_m(h_m)\left(K_{ij}^A \frac{\partial h_m}{\partial x_j} + K_{iz}^A\right)\right] - S_m(h_m) + \frac{\Gamma_w}{1-w} \end{aligned} \tag{6}$$

respectively, where w = ratio of the volumes of the macropore or fracture domain and the total soil system [-]. Note that the water contents θ_f and θ_m in (1) have different meanings than in (5) where they represented water contents of the total pore space (i.e., $\theta = \theta_{mo} + \theta_{im}$), while in (6) they refer to water contents of the two separate (fracture or matrix) pore domains such that $\theta = w\theta_f + (1-w)\theta_m$.

Examples of applications of dual-permeability models to a range of laboratory and field data involving transient flow and solute transport are given by Pot et al. (2005), Kodešová et al. (2005), and Köhne et al. (2006).

2.3 *Water flow and vapor transport*

The Richards equation (1) considers only water flow in the liquid phase and ignores the effects of the vapor phase on the overall water mass balance. While this assumption is justified for the majority of applications, a number of problems exist in which the effect of vapor flow can not be neglected. Vapor movement is often an important part of the total water flux when the soil moisture becomes relatively low. Scanlon et al. (2003) showed that water fluxes in deep vadose zone profiles of the arid and semiarid regions of the western U.S. are often dominated by thermal vapor fluxes. Nonisothermal liquid and vapor flow in an unpublished research version of HYDRUS is described using the equation (e.g., Saito et al. 2006):

$$\frac{\partial\theta(h)}{\partial t} = \frac{\partial}{\partial x_i}\left[(K + K_{vh})\left(K_{ij}^A \frac{\partial h}{\partial x_j} + K_{iz}^A\right) + (K_{LT} + K_{vT})K_{ij}^A \frac{\partial T}{\partial x_j}\right] - S(h) \tag{7}$$

where $\theta =$ total volumetric water content, being the sum ($\theta = \theta_l + \theta_v$) of the volumetric liquid water content, θ_l, and the volumetric water vapor content, θ_v (both expressed as an equivalent water content), $T =$ temperature [K]; $K =$ isothermal hydraulic conductivity for the liquid phase [LT^{-1}]; $K_{LT} =$ thermal hydraulic conductivity for the liquid phase [$L^2K^{-1}T^{-1}$]; $K_{vh} =$ isothermal vapor hydraulic conductivity [LT^{-1}]; $K_{vT} =$ thermal vapor hydraulic conductivity [$L^2K^{-1}T^{-1}$]. Overall water flow in (7) is given as the sum of isothermal liquid flow, isothermal vapor flow, gravitational liquid flow, thermal liquid flow, and thermal vapor flow. Since several terms of (7) are a function of temperature, this equation should be solved simultaneously with the heat transport equation (11) to properly account for temporal and spatial changes in soil temperature.

Examples of HYDRUS-1D applications that consider the simultaneous movement of water in the liquid and vapor phases are given by Scanlon et al. (2003) and Saito et al. (2006). Scanlon et al. (2003) evaluated response of several semiarid (High Plains, Texas) and arid (Chihuahuan Desert, Texas; Amargosa Desert, Nevada) sites to paleoclimatic forcing using water-potential and Cl profiles, and by modeling nonisothermal liquid and vapor flow and Cl transport. They used HYDRUS-1D to show that upward nonisothermal liquid and vapor flow for at least 1 to 2 kyr in the High Plains and for 12 to 16 kyr in the Chihuahuan and Amargosa desert sites is required to reproduce measured upward water potentials and Cl profiles. Saito et al. (2006) additionally combined a numerical solution of the water flow (7) and heat transport (11) equations with water and energy balances at the soil surface. Their analyses of the distributions of the liquid and vapor fluxes versus depth showed that the soil water dynamics close to the soil surface is strongly associated with the soil temperature regime.

2.4 *Water flow with freezing and thawing*

Although considering both liquid and vapor flow, equation (7) still neglects one phenomenon that is important in cold climatic regions, i.e., freezing and thawing. To include freezing and thawing processes, (7) needs an additional term that accounts for frozen water (e.g., Hansson et al. 2004) as follows:

$$\begin{aligned}\frac{\partial\theta(h)}{\partial t} + \frac{\rho_i}{\rho_w}\frac{\partial\theta_i(T)}{\partial t} = \\ = \frac{\partial}{\partial x_i}\left[(K + K_{vh})\left(K_{ij}^A \frac{\partial h}{\partial x_j} + K_{iz}^A\right) + (K_{LT} + K_{vT})K_{ij}^A \frac{\partial T}{\partial x_j}\right] - S(h)\end{aligned} \tag{8}$$

where $\theta_i =$ volumetric ice content [L^3L^{-3}]; $\rho_i =$ density of ice [ML^{-3}] (931 kg m^{-3}); and $\rho_w =$ density of liquid water [ML^{-3}] ($\sim$1000 kg m^{-3}). This equation is even more tightly coupled with the heat transport equation (12) through the second term on the left side, since the change in the volumetric ice content depends on temperature and can be expressed using the generalized Clapeyron equation.

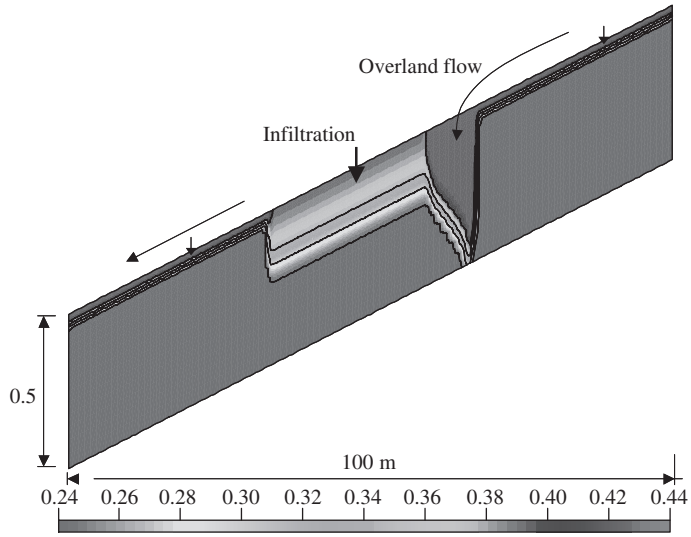


Figure 1. Calculated contours of the water content in a hillside transect following rainfall and overland flow (Šimůnek 2003).

2.5 *Overland flow*

Traditionally, surface and subsurface flow processes have been considered mostly separately, with some of the processes being severely simplified when coupled surface and subsurface flow is simulated. For example, existing surface irrigation models typically still use empirical infiltration functions such as the Philip, Kostiakov, modified Kostiakov, and Branch infiltration equations, rather than rigorously simulating subsurface water flow and solute transport using the Richards equation for variably-saturated flow, and advection-dispersion type equations for solute transport. To overcome this deficiency, the following overland flow equation has been incorporated in HYDRUS-2D for use in combination with the atmospheric boundary condition at the surface boundary (Šimůnek 2003):

$$\begin{aligned} &\frac{\partial h}{\partial t} + \frac{\partial Q}{\partial x} = q(x,t) \\ &Q = \alpha h^m \qquad \text{where} \qquad \alpha = \frac{S^{1/2}}{n} \qquad \text{and} \qquad m = 5/3 \end{aligned} \tag{9}$$

where $h =$ unit storage of water (or mean depth for smooth surfaces) [L]; $Q =$ discharge per unit width [L^2T^{-1}], $x =$ distance coordinate over the soil surface [L]; $q(x,t) =$ rate of local input, or lateral inflow (i.e., local precipitation minus local infiltration) [LT^{-1}]; α [$L^{2-m}T^{-1}$] and m [$-$] = parameters related to slope, surface roughness, and the flow conditions (laminar or turbulent flow), usually evaluated using Manning's hydraulic resistance law; $S =$ slope [-]; and $n =$ Manning's roughness coefficient for overland flow [$-$].

Figure 1 shows a simple application of the overland flow module in HYDRUS-2D to simulate flow in a hillside transect in which the soil material in the middle third of the transect (between 33 and 66 m) has a 100 times higher saturated hydraulic conductivity than the remaining soil (25 versus 0.25 m d^{-1}). Because of this, the middle section can accommodate the infiltration of water from both the rainfall itself and from runoff from the upgradient part of the hillslope. The soil transect had a slope of 0.01, while the roughness coefficient n was assumed to be equal to 0.01.

2.6 *HYDRUS package for MODFLOW*

Although computer power has increased tremendously during the last few decades, large scale three-dimensional applications evaluating water flow in the vadose zone are often still prohibitively expensive in terms of computational resources. To overcome this problem, Seo et al. (2006, 2007) developed a computationally efficient one-dimensional unsaturated flow HYDRUS package and linked it to the three-dimensional modular finite-difference ground water model MODFLOW-2000 (Harbaugh et al. 2000). The HYDRUS unsaturated flow package used HYDRUS-1D to simulate one-dimensional vertical variably-saturated flow. MODFLOW zone arrays were used to define the cells to which the HYDRUS package was applied. MODFLOW used the time-averaged flux from the bottom of the unsaturated zone as recharge, and calculated a water table depth which was then used as a pressure head bottom boundary for HYDRUS. Twarakavi et al. (in press) provided a comparison of the HYDRUS package to other MODFLOW packages that evaluate processes in the vadose zone and presented a field application demonstrating the functionality of the package.

3 HEAT TRANSPORT

3.1 *Heat transport without vapor transport*

All HYDRUS models consider heat transport due to conduction and convection with flowing water. Neglecting the effects of water vapor diffusion, the heat transport can be described as:

$$C(\theta)\frac{\partial T}{\partial t} = \frac{\partial}{\partial x_i}\left(\lambda_{ij}(\theta)\frac{\partial T}{\partial x_j}\right) - C_w q_i \frac{\partial T}{\partial x_i} \tag{10}$$

where $\lambda_{ij}(\theta)$ = apparent thermal conductivity of the soil [$MLT^{-3}K^{-1}$] (e.g. $Wm^{-1}K^{-1}$); $C(\theta)$ and C_w = volumetric heat capacities [$ML^{-1}T^{-2}K^{-1}$] (e.g. $Jm^{-3}K^{-1}$) of the porous medium and the liquid phase, respectively; and q_i = fluid flux density [LT^{-1}]. The HYDRUS models at each time step first numerically solve the Richards equation (1) and then the heat transport equation (10) using calculated values of the water content and the fluid flux.

3.2 *Heat transport with vapor transport*

When the effects of water vapor diffusion can not be neglected, the heat transport must be expanded to the form (e.g., Saito et al. 2006):

$$C(\theta)\frac{\partial T}{\partial t} + L_0\frac{\partial \theta_v}{\partial t} = \frac{\partial}{\partial x_i}\left(\lambda_{ij}(\theta)\frac{\partial T}{\partial x_j}\right) - C_w q_i \frac{\partial T}{\partial x_i} - C_v \frac{\partial q_{vi} T}{\partial x_i} - L_0\frac{\partial q_{vi}}{\partial x_i} \tag{11}$$

where L_0 = volumetric latent heat of vaporization of liquid water [$ML^{-1}T^{-2}$] (e.g., Jm^{-3}); and q_{vi} = vapor flux density [LT^{-1}]. In equation (11), the total heat flux density is defined as the sum of the conduction of sensible heat as described by Fourier's law (the first term on the right side), sensible heat by convection of liquid water (the second term) and water vapor (the third term), and of latent heat by vapor flow (the forth term) (de Vries 1958).

As an example, Figure 2 shows calculated water content, total flux, temperature and concentration profiles for a 10-cm long soil sample with zero water fluxes at both top and bottom boundaries, and with a temperature gradient along the sample. Increasing temperatures (Fig. 2c) from the top to the bottom of the sample cause vapor flow (Fig. 2b) from the warmer bottom end of the sample toward the colder end. Water evaporates at the warmer end, flows upward as vapor and condensates at the colder end. Water contents correspondingly decrease at the warmer end and increase at the colder bottom (Fig. 2a). As a consequence of changing water contents, a pressure head gradient develops in the sample, leading to water flow in a direction opposite to vapor flow. A steady-state is eventually reached when the upward vapor flow fully balances the downward liquid flow (Fig. 2b).

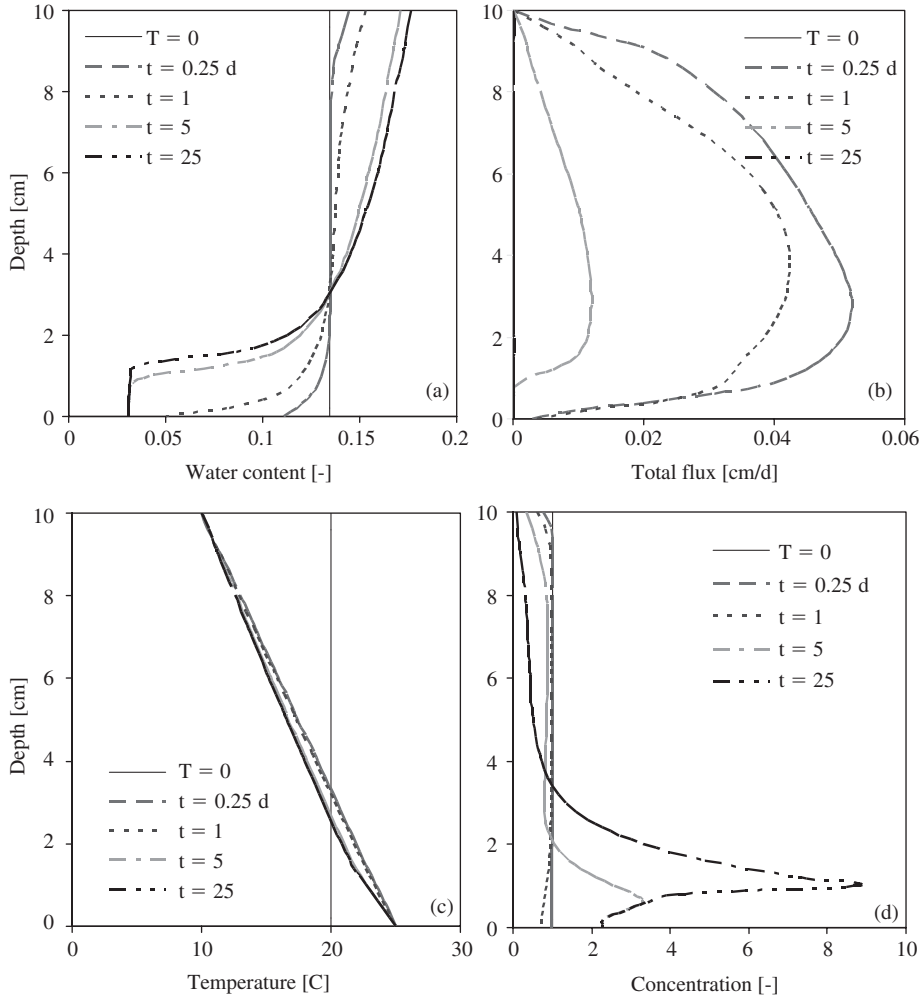


Figure 2. Water content (a), total flux (b), temperature (c), and solute concentration (d) distributions in a 10-cm long vertical soil sample with zero water fluxes across the top and bottom boundaries, and with temperature increasing from top to bottom.

Since water evaporates at the bottom of the sample and condensates at the top, solute becomes more concentrated near the bottom and more diluted near the top (Fig. 2d). Also, the concentration profile should eventually reach steady-state, although at a much larger time, when the advective downward solute flux balances the upward diffusive flux.

3.3 *Heat transport with freezing and thawing*

Similar to water flow, considering freezing and thawing in the heat transport equation requires an additional term to account for the energy stored in the frozen water (e.g., Hansson et al. 2004):

$$C(\theta)\frac{\partial T}{\partial t} - L_f \rho_i \frac{\partial \theta_i}{\partial t} + L_0 \frac{\partial \theta_v}{\partial t} = \\ = \frac{\partial}{\partial x_i}\left(\lambda_{ij}(\theta)\frac{\partial T}{\partial x_j}\right) - C_w q_i \frac{\partial T}{\partial x_i} - C_v \frac{\partial q_{vi} T}{\partial x_i} - L_0 \frac{\partial q_{vi}}{\partial x_i} \tag{12}$$

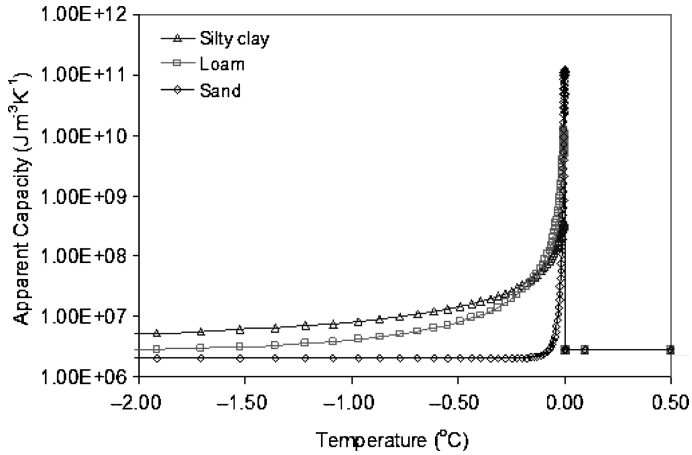


Figure 3. Apparent volumetric heat capacity ($\mathrm{Jm^{-3}K^{-1}}$) for three soil textural classes (sand, loam, and silty clay).

where L_f = latent heat of freezing [$\mathrm{L^2T^{-2}}$] ($\sim$3.34e + 5 J kg^{-1}). The first term on the left-hand side represents changes in the energy content, whereas the second and third terms represent changes in the latent heat of the frozen and vapor phases, respectively. Terms on the right-hand side represent respectively, soil heat flow by conduction, convection of sensible heat with flowing water, transfer of sensible heat by diffusion of water vapor, and transfer of latent heat by diffusion of water vapor. Hansson et al. (2004) discussed the extreme nonlinearity of (12) caused by the dependency of the apparent volumetric heat capacity ($=C - L_f \rho_i d\theta_i/dT$) on temperature. While the apparent heat capacity increases by about two orders of magnitude for silty clay when the freezing point is reached, the increase for sand is almost five orders of magnitude (Fig. 3). For sands the increase in the apparent heat capacity due to freezing becomes negligible below about −0.05°C, which corresponds to a pressure head of −6.2 m. At that temperature almost all soil water is frozen, except for the residual water content. For fine-textured soils the increase in the apparent heat capacity extends to much lower temperatures, which reflects the fact that for these soils a significant amount of water remains unfrozen at slightly subzero temperatures.

Hansson et al. (2004) used HYDRUS-1D to evaluate data from laboratory column freezing experiments. The experiments involved 20-cm long soil columns with an internal diameter of 8 cm that were exposed at the top to a circulating fluid with a temperature of −6°C. Water and soil in the columns froze from the top down during the experiment, with the freezing process inducing significant water redistribution within the soil. Values of the total water content calculated with HYDRUS-1D compared well with measured values.

Figure 4 shows temperatures, pressure heads, liquid water contents, and concentrations at different depths in a loamy soil profile subject to fluctuating surface temperatures. Surface temperatures followed the sinusoidal curve with a mean value of −1°C and a daily amplitude of 5°C. Changes in temperature can be observed down to a depth of about 10 cm (Fig. 4a). The Clapeyron equation predicts that a temperature decrease of 1°C below zero corresponds to a pressure head decrease of about −120 m. This shows that negative temperatures can cause extremely negative pressure heads in the soil profile (Fig. 4b). These pressure heads have their corresponding water contents (Fig. 4c), which follow immediately from the soil water retention curve. The remaining water within the soil profile stays in a frozen state. From Figure 4 it is obvious that the process of freezing is very similar to drying, with frozen water remaining in the larger pores and liquid water in the smaller pores. Concentrations fluctuate in a similar fashion as the other variables since solutes are excluded from frozen water and remain in the liquid phase (Fig. 4d).

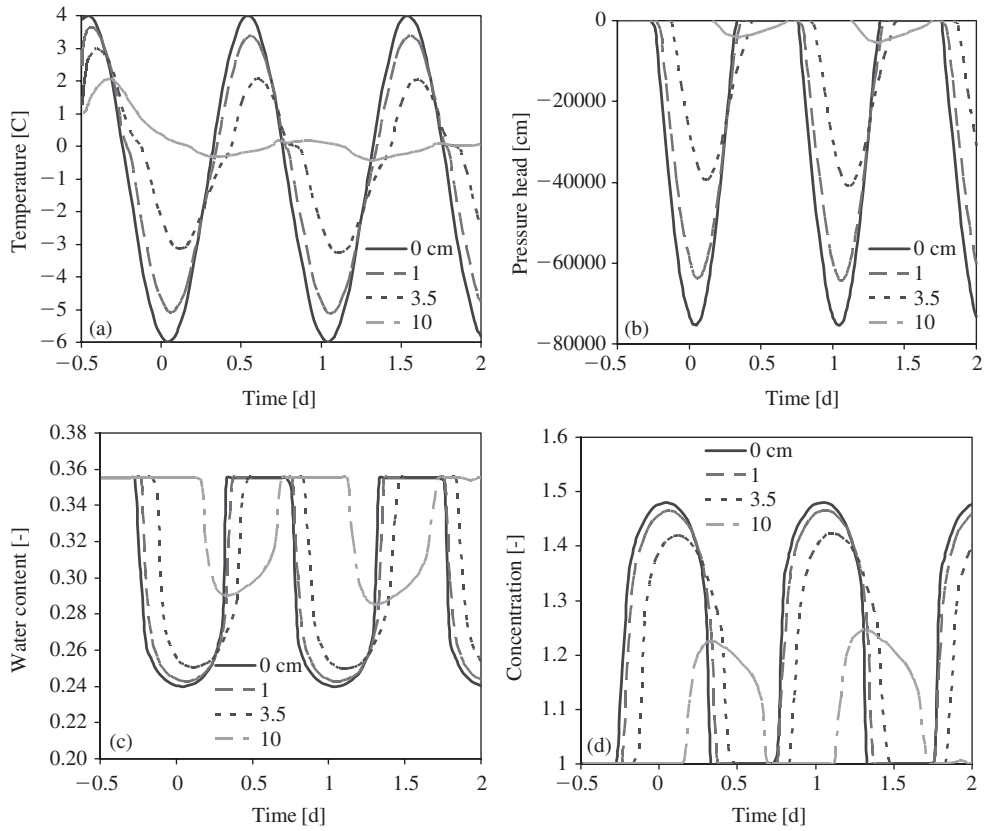


Figure 4. Temperatures (a), pressure heads (b), water contents (c), and concentrations (d) at different depths in a loamy soil profile subject to fluctuating surface temperatures.

4 SOLUTE TRANSPORT

While water flow in the various HYDRUS software packages is simulated in a relatively similar manner (i.e., using the Richards equation to describe uniform flow, or using dual-porosity or dual-permeability models for nonequilibrium flow), different solute transport modules exist in the different codes. Some modules simulate single-component transport, while others contain biogeochemical modules for simulating potentially very complex interactions between the various solutes. In this section we provide brief descriptions of the various solute transport modules.

4.1 *Single component solute transport*

Most HYDRUS predecessors (e.g., SWMS_2D and SWMS_3D) allowed solutes to exist only in the liquid and solid phases. The newer windows-based HYDRUS models (i.e., HYDRUS-1D, HYDRUS-2D and HYDRUS (2D/3D)) assume that solutes can exist in all three phases (liquid, solid, and gaseous) and that the decay and production processes can be different in each phase. While in the HYDRUS predecessors solutes could be transported only by advection and dispersion in the liquid phase, the more recent versions assume that they can be transported also by diffusion in the gas phase. In the various solute transport formulations below we will use the following general conservation equation for solutes in a variably saturated rigid porous medium

$$\frac{\partial M}{\partial t} = -\frac{\partial J_{Si}}{\partial x_i} - \phi \tag{13}$$

where M = mass of solute [ML^{-3}]; J_{Si} = solute flux density [$ML^{-2}T^{-1}$]; and ϕ = coefficient representing various sinks and sources [$ML^{-3}T^{-1}$].

4.1.1 *Equilibrium transport*

The SWMS_2D and SWMS_3D models (Šimůnek et al. 1994, 1995) assume that solutes reside only in the liquid and solid phases. The various terms of (13) are then given by:

$$\begin{aligned} M &= \theta c + \rho s \\ J_{Si} &= -\theta D_{ij}^{w} \frac{\partial c}{\partial x_j} + q_i c \\ \phi &= -\mu_w \theta c - \mu_s \rho s + \gamma_w \theta + \gamma_s \rho - S c_r \end{aligned} \tag{14}$$

where c, s = solute concentrations in the liquid [ML^{-3}] and solid [MM^{-1}] phases, respectively; μ_w, μ_s = first-order rate constants for solutes in the liquid and solid phases [T^{-1}], respectively; γ_w, γ_s = zero-order rate constants for the liquid [$ML^{-3}T^{-1}$] and solid [T^{-1}] phases, respectively; ρ = soil bulk density [ML^{-3}]; S = sink term in the water flow equation (1); c_r = concentration of the sink term [ML^{-3}]; D_{ij}^{w} = dispersion coefficient tensor [L^2T^{-1}] for the liquid phase.

The HYDRUS models allow solutes to exist also in the gaseous phase, as well as permit them to be involved in sequential first-order decay chain reactions. The different terms of (13) then become:

$$\begin{aligned} M_k &= \theta c_k + \rho s_k + a_v g_k \qquad k = 1, N_s \\ J_{Si,k} &= -\theta D_{ij,k}^{w} \frac{\partial c_k}{\partial x_j} - a_v D_{ij,k}^{g} \frac{\partial g_k}{\partial x_j} + q_i c_k \\ \phi_k &= -(\mu_{w,k} + \mu'_{w,k})\theta c_k - (\mu_{s,k} + \mu'_{s,k})\rho s_k - (\mu_{g,k} + \mu'_{g,k}) a_v g_k - S c_{r,k} + \\ &\quad + \vartheta(\mu'_{w,k-1}\theta c_{k-1} + \mu'_{s,k-1}\rho s_{k-1} + \mu'_{g,k-1} a_v g_{k-1}) + \gamma_{w,k}\theta + \gamma_{s,k}\rho + \gamma_{g,k} a_v \end{aligned} \tag{15}$$

where g = solute concentration in the gaseous phase [ML^{-3}]; μ_g = first-order rate constant for solutes in the gas phase [T^{-1}]; μ'_w, μ'_s, and μ'_g = similar first-order rate constants providing connections between individual species in the decay chain; γ_g = zero-order rate constant for the gas phase [$ML^{-3}T^{-1}$]; a_v = air content [L^3L^{-3}]; D_{ij}^{g} = diffusion coefficient tensor [L^2T^{-1}] for the gas phase; and N_s = number of solutes involved in the chain reaction. The subscripts w, s, and g correspond with the liquid, solid and gas phases, respectively; while the subscript k represents the kth chain number. The parameter ϑ in (15) is zero for $k = 1$. The indicial notation used above assumes summations over indices i and j (i, j = 1,2,3), but not over index k.

Interactions between the solid and liquid phases in the HYDRUS models may be described using nonlinear equations, while interactions between the liquid and gaseous phases are assumed to be linear and instantaneous, i.e.,

$$s = \frac{K_d c^{\beta}}{1 + \eta c^{\beta}} \tag{16}$$

$$g = k_g c \tag{17}$$

respectively, where K_d, β, η, k_g = empirical coefficients [L^3M^{-1}], [-], [L^3M^{-1}], [-], respectively. The Freundlich, Langmuir, and linear adsorption equations are special cases of equation (16). When $\beta = 1$, equation (16) reduced to the Langmuir equation, when $\eta = 0$ the equation becomes a Freundlich isotherm, and when $\beta = 1$ and $\eta = 0$, equation (16) defines a linear adsorption isotherm. Solute transport without adsorption is described with $K_d = 0$. We note here that the SWMS_2D and SWMS_3D models considered only linear sorption (i.e., $\beta = 1$ and $\eta = 0$).

Typical examples of HYDRUS applications involving sequential first-order decay chains are the transport of nitrogen species (e.g., Hanson et al. 2006), pesticides, chlorinated aliphatic hydrocarbons (Schaerlaekens et al. 1999; Casey & Šimůnek 2001), hormones (Casey et al. 2003), and

explosives (Dontsova et al. 2006). The ability of HYDRUS to account for diffusion in the gas phase was further used by Wang et al. (1997, 2000) to simulated the transport of volatile contaminants (fumigants) such as methyl bromine and 1,3-dichloropropene.

4.1.2 *Nonequilibrium transport*

HYDRUS-1D and the two-dimensional module of HYDRUS (2D/3D) offer a large number of options to deal with nonequilibrium flow and transport. A complete list of these options is given by Šimůnek et al. (in press) who divided them into three groups: (a) physical nonequilibrium transport models, (b) chemical nonequilibrium transport models, and (c) physical and chemical nonequilibrium transport models. Physical nonequilibrium models include (1) the Mobile-Immobile Water Model, (2) the Dual-Porosity Model, (3) the Dual-Permeability Model, and (4) the Dual-Permeability Models with Immobile Water. Chemical nonequilibrium models include (1) the One Kinetic Site Model, (2) the Two-Site Model, and (3) the Two-Kinetic Sites Model. Finally, physical and chemical nonequilibrium transport models include (1) the Dual-Porosity Model with one Kinetic Site, and (2) the Dual-Permeability Model with Two-Site Sorption.

An example of **physical nonequilibrium** transport as implemented in HYDRUS is the Mobile-Immobile Water Model that uses the concept of two-region, dual-porosity type solute transport (van Genuchten & Wierenga 1976). Solute exchange between the two liquid regions is then modeled as a first-order process:

$$\frac{\partial \theta_{im} c_{im}}{\partial t} + (1 - f_{mo}) \rho \frac{\partial s_{im}}{\partial t} = \Gamma_s - \phi_{im}$$
$$\Gamma_s = \alpha (c_{mo} - c_{im}) + \Gamma_w c^* \tag{18}$$

where c_{mo} and c_{im} = concentrations of the mobile and immobile regions [ML^{-3}], respectively; s_{im} = sorbed concentration of the immobile region [-]; ϕ_{im} = sink/source term that accounts for various zero- and first-order or other reactions in immobile region [$ML^{-3}T^{-1}$]; f_{mo} = fraction of sorption sites in contact with the mobile water content [-]; α = mass transfer coefficient [T^{-1}]; and Γ_s = mass transfer term for solutes between the mobile and immobile regions [$ML^{-3}T^{-1}$]. The first equation of (18) is a mass balance for the immobile (micropore) domain, while the second equation (Γ_s) describes the rate of mass transfer between the mobile and immobile domains. The second (advective) term of Γ_s in Eq. (18) is equal to zero for the Mobile-Immobile Model since that model does not consider water flow between the two regions. In the Dual-Porosity Model, c^* is equal to c_{mo} for $\Gamma_w > 0$ and c_{im} for $\Gamma_w < 0$. Note that although the HYDRUS models do consider diffusive transport in the gas phase also for nonequilibrium applications, for simplicity we did not include this process in (18) and subsequent equations.

Still much more sophisticated models for physical nonequilibrium transport may be formulated. For example, Pot et al. (2005), Köhne et al. (2006), and Šimůnek et al. (in press) considered not only a dual-permeability model with flow and transport in both regions, but also a dual-permeability model that additionally divides the matrix domain into mobile and immobile subregions. Pot et al. (2005) and Köhne et al. (2006) used the HYDRUS-1D model with these dual-permeability submodels to successfully simulate bromide transport in laboratory soil columns at different flow rates or for transient flow conditions, respectively.

An example of **chemical nonequilibrium** transport in HYDRUS is the Two-Site Model that assumes that the sorption sites can be divided into two fractions:

$$s = s^e + s^k \tag{19}$$

where sorption, s^e [MM^{-1}], on one fraction of the sites (often referred to as type-1 sites) is assumed to be instantaneous, while sorption, s^k [MM^{-1}], on the remaining (type-2) sites is considered to be time-dependent. Sorption on the type-2 nonequilibrium sites is assumed to be a first-order kinetic rate process:

$$\rho \frac{\partial s^k}{\partial t} = \alpha_k \rho [s_e^k - s^k] - \phi_k \tag{20}$$

where α_k = first-order rate constant describing the kinetics of the sorption process [T^{-1}]; s_e^k = sorbed concentration that would be reached at equilibrium with the liquid phase concentration [MM^{-1}]; ϕ_k = sink/source term that accounts for various zero- and first-order or other reactions at the kinetic sorption sites [$ML^{-3}T^{-1}$].

Although we listed the Two-Kinetic Sites Model above among the chemical nonequilibrium models, this model could equally well be viewed as a **physical and chemical nonequilibrium** transport model. This because both physical and chemical nonequilibrium processes could be responsible for solutes associated with each of the two sites. For example, while the chemical process of attachment/detachment could be associated with one fraction of sorption sites, the physical process of straining could be associated with the second fraction of sorption sites. In the Two-Kinetic Sites Model the sorption sites are again divided into two fractions:

$$s = s_1^k + s_2^k \tag{21}$$

but now both sorption sites are assumed to be kinetic. In (21), s_1^k and s_2^k = concentrations of the first and second fraction of kinetic sorption sites [MM^{-1}], respectively. The kinetic processes are then described as follows:

$$\begin{aligned} \rho \frac{\partial s_1^k}{\partial t} &= k_{a1}\theta c - k_{d1}\rho s_1^k - \phi_{k1} \\ \rho \frac{\partial s_2^k}{\partial t} &= k_{a2}\theta c - k_{d2}\rho s_2^k - \phi_{k2} \end{aligned} \tag{22}$$

where k_{a1}, k_{a2} = attachment coefficients for the first and second fraction of kinetic sorption sites [T^{-1}], respectively; k_{d1}, k_{d2} = detachment coefficients for the first and second fraction of kinetic sorption sites [T^{-1}], respectively; and ϕ_{k1}, ϕ_{k2} = sink/source terms for the first and second fraction of kinetic sorption sites [T^{-1}], respectively. As mentioned above, the two kinetic sites may be used to describe different processes. While the first kinetic process could be used for chemical attachment, the second kinetic process could represent physical straining of colloids or pathogenic microorganisms (e.g., Bradford et al. 2004, 2007; Gargiulo et al. 2007). Note that in (22) we do not give the nonlinear blocking coefficients to account for, for example, Langmuirian blocking to attachment or depth-dependent straining that are considered in the HYDRUS models (e.g., Bradford et al. 2004).

Details about the remaining nonequilibrium models, as well as of various applications, can be found in the HYDRUS manuals as well as in Šimůnek et al. (in press).

4.2 *Multicomponent solute transport*

The transport of reactive contaminants in the subsurface is generally affected by a large number of nonlinear and often interactive physical, chemical and biological processes. Proper simulating these processes requires a reactive transport code that couples the physical processes of water flow and advective-dispersive transport with a range of biogeochemical processes. The soil solution is always a mixture of many ions which may be involved in mutually dependent chemical processes, such as complexation reactions, cation exchange, precipitation-dissolution, sorption-desorption, volatilization, redox reactions, and degradation (Šimůnek & Valocchi 2002; Appelo & Postma 2005). The transport and transformation of many contaminants is further mediated by subsurface aerobic or anaerobic bacteria. Bacteria for example can catalyze redox reactions in which organic compounds (e.g., chlorinated hydrocarbons) act as the electron donor and inorganic substances (oxygen, nitrate, sulfate, or metal oxides) as the electron acceptor.

To be able to evaluate these type of systems, several different multicomponent modules were developed for the HYDRUS models. In addition to the single component solute transport module described above, HYDRUS-1D also includes the UNSATCHEM module (Šimůnek & Suarez 1993, 1994) that considers the transport of major ions, and their mutual reactions, such as dissolution-precipitation, cation exchange, or aqueous complexation. HYDRUS-1D has also been coupled

with PHREEQC, resulting in the HP1 model (Jacques & Šimůnek 2005; Jacques et al. 2006) that accounts for a wide range of instantaneous and kinetic chemical and biological reactions, including aqueous complexation, precipitation-dissolution, cation exchange, surface complexation, and redox reactions. Finally, the CW2D wetland module (Langergraber & Šimůnek 2005, 2006) in HYDRUS (2D/3D) considers biochemical processes associated with the flow of wastewater through constructed or natural wetlands. Below we briefly review these models and indicate some of their applications.

Šimůnek & Valocchi (2002) divided geochemical transport models into two major groups: those with specific chemistry and more general models. Models with specific chemistry are generally restricted to certain prescribed chemical systems and thus are usually constrained to specific applications. The major ion chemistry UNSATCHEM module in the HYDRUS-1D software package and the CW2D wetland module in HYDRUS (2D/3D) are typical examples of models with specific chemistry. On the other hand, models with generalized chemistry provide users with much more freedom in simulating chemical systems, and thus permit a much broader range of applications. Users then either can select species and reactions from large geochemical databases, or are able to define their own species with particular chemical properties and reactions. The HP1 model represents a typical example of a model with generalized chemistry.

4.2.1 *Major ion chemistry*

As an alternative to the single-component approach, a major ion chemistry module based on the UNSATCHEM model (Šimůnek & Suarez 1993, 1994) was incorporated in HYDRUS-1D. This module considers the transport of major ions and carbon dioxide in soils. CO_2 transport is assumed to be governed by diffusion in both the liquid and gas phases, by advection in the liquid phase, and by respiration by both soil microorganisms and plant roots (Šimůnek & Suarez 1993):

$$\frac{\partial(c_a a_v + c_w\theta)}{\partial t} = \frac{\partial}{\partial z}\left(a_v D_a \frac{\partial c_a}{\partial z}\right) + \frac{\partial}{\partial z}\left(\theta D_w \frac{\partial c_w}{\partial z}\right) - \frac{\partial q_a c_a}{\partial z} - \frac{\partial q c_w}{\partial z} - S c_w + P \qquad (23)$$

$$P = \gamma_s + \gamma_p$$

where c_w, c_a = volumetric concentrations of CO_2 in the dissolved and gas phases [L^3L^{-3}], respectively; D_a = effective soil matrix diffusion coefficient of CO_2 in the gas phase [L^2T^{-1}]; D_w = effective soil matrix dispersion coefficient of CO_2 in the dissolved phase [L^2T^{-1}]; q_a = soil air flux [LT^{-1}]; q = soil water flux [LT^{-1}]; Sc_w = dissolved CO_2 removed from the soil by root water uptake [$L^3L^{-3}T^{-1}$]; P = CO_2 production/sink term [$L^3L^{-3}T^{-1}$]; and γ_s, γ_p = CO_2 production rates by soil microorganisms or plant roots [$L^3L^{-2}T^{-1}$], respectively (Šimůnek and Suarez, 1993).

The major variables of the UNSATCHEM chemical system are Ca, Mg, Na, K, SO_4, Cl, NO_3, H_4SiO_4, alkalinity, and CO_2. The model accounts for equilibrium chemical reactions between these components such as complexation, cation exchange and precipitation-dissolution, e.g.,

$$\begin{aligned} K_{\mathrm{CaCO_3}} &= \frac{(\mathrm{Ca^{2+}})(\mathrm{CO_3^{2-}})}{(\mathrm{CaCO_3^0})} \\ K_{\mathrm{Mg,Ca}} &= \frac{\overline{\mathrm{Mg}}^{2+}(\mathrm{Ca^{2+}})^{1/2}}{\overline{\mathrm{Ca}}^{2+}(\mathrm{Mg^{2+}})^{1/2}} \\ K_{SP}^{C} &= (\mathrm{Ca^{2+}})(\mathrm{CO_3^{2-}}) \end{aligned} \qquad (24)$$

respectively, where the parentheses denote ion activities, K_{CaCO3} = equilibrium constant of the $CaCO_3^0$ complexed species [-]; $K_{\mathrm{Mg,Ca}}$ = Gapon selectivity coefficient for Mg^{2+} and Ca^{2+} [-]; and K_{SP}^{C} = calcite solubility product [-]. Similar algebraic equations hold for all species of the chemical system.

For the precipitation-dissolution of calcite ($CaCO_3$) and the dissolution of dolomite ($MgCO_3 \cdot 3H_2O$), either equilibrium or multicomponent kinetic expressions can be used, including

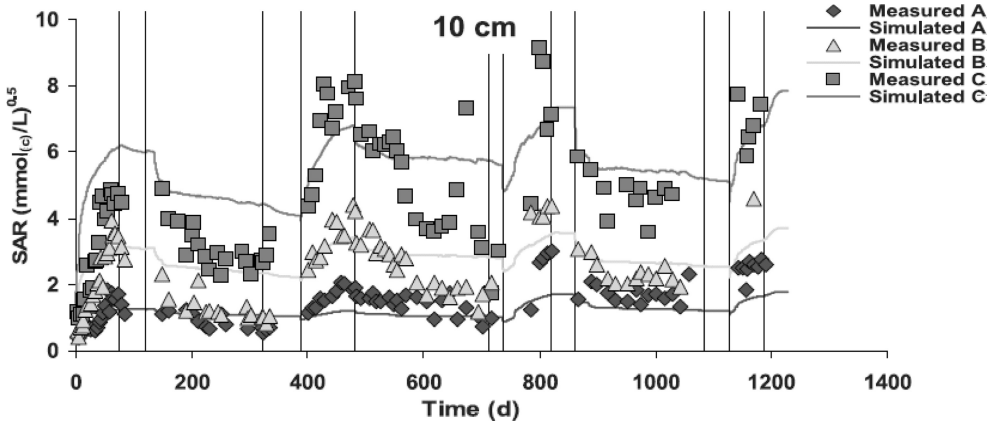


Figure 5. Measured and simulated sodium adsorption ratios (SAR) for lysimeters A, B, and C (Gonçalves et al. 2005).

both forward and backward reactions. Other precipitation-dissolution reactions considered in the UNSATCHEM models involve gypsum ($CaSO_4 \cdot 2H_2O$), hydromagnesite ($Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$), nesquehonite ($MgCO_3 \cdot 3H_2O$), and sepiolite ($Mg_2Si_3O_{7.5}(OH) \cdot 3H_2O$). Since the ionic strength of soil solutions can vary considerably in time and space and often reach high values, both the modified Debye-Hückel (Truesdell & Jones 1974) and Pitzer (Pitzer 1979) expressions were incorporated into the model, thus providing options for calculating single-ion activities. The module also considers the effects of solution composition on the unsaturated soil hydraulic properties. The new UNSATCHEM module of HYDRUS-1D enables quantitative predictions of processes involving major ions, such as simulations of the effects of salinity on root water uptake and plant growth, evaluation of alternative irrigation, salinity and crop management practices, evaluation of water suitability for irrigation, and estimation of the amount of water and amendment required to reclaim soil profiles to desired levels of salinity and ESP (exchangeable sodium percentage).

A recent application of HYDRUS-1D and its major ion chemistry module is given by Gonçalves et al. (2005) who simulated solute transport in three lysimeters irrigated with different quality waters over a time period of three years. HYDRUS-1D successfully described field measurements of not only the overall salinity, but also of individual soluble cations as well as SAR (Sodium Adsorption Ratio) (Fig. 5) and ESP (Gonçalves et al. 2005).

4.2.2 *Chemical and biological reactions in HP1*

HYDRUS-1D was recently coupled also with the PHREEQC geochemical code (Parkhurst & Appelo 1999) to create a new comprehensive simulation tool, HP1 (acronym for HYDRUS1D-PHREEQC) (Jacques & Šimůnek 2005; Jacques et al. 2006; Šimůnek et al. 2006a). This new code contains modules simulating (1) transient water flow in variably-saturated media, (2) the transport of multiple components, (3) mixed equilibrium/kinetic biogeochemical reactions, and (4) heat transport. HP1 is a significant expansion of the individual HYDRUS-1D and PHREEQC programs by preserving most of their original features. The code still uses the Richards equation for simulating variably-saturated water flow and advection-dispersion type equations for heat and solute transport. However, the loosely coupled program can simulate also a broad range of low-temperature biogeochemical reactions in water, the vadose zone and in ground water systems, including interactions with minerals, gases, exchangers and sorption surfaces based on thermodynamic equilibrium, kinetic, or mixed equilibrium-kinetic reactions. HP1 uses the operator-splitting approach with no iterations during one time step (a non-iterative sequential modeling approach). Jacques et al. (2006) evaluated the accuracy of the operator-splitting approach for a kinetic reaction network (i.e., sequential and parallel kinetic degradation reactions) by comparing HP1 with an

analytical solution for TCE-degradation, as well as for mixed equilibrium and kinetic reactions involving different flow conditions (steady-state and transient).

Jacques & Šimůnek (2005), and Šimůnek et al. (2006a) and Jacques et al. (2006; in press), demonstrated the versatility of HP1 on several examples, which included (a) the transport of heavy metals (Zn^{2+}, Pb^{2+}, and Cd^{2+}) subject to multiple cation exchange reactions, (b) transport with mineral dissolution of amorphous SiO_2 and gibbsite ($Al(OH)_3$), (c) heavy metal transport in a medium with a pH-dependent cation exchange complex, (d) infiltration of a hyperalkaline solution in a clay sample (this example considers kinetic precipitation-dissolution of kaolinite, illite, quartz, calcite, dolomite, gypsum, hydrotalcite, and sepiolite), (e) long-term transient flow and transport of major cations (Na^+, K^+, Ca^{2+}, and Mg^{2+}) and heavy metals (Cd^{2+}, Zn^{2+}, and Pb^{2+}) in a soil profile, (f) cadmium leaching in acid sandy soils, (g) radionuclide transport, and (h) long term uranium migration in agricultural field soils following mineral P-fertilization.

4.2.3 *Wetland module*

A multi-component reactive transport model CW2D (**C**onstructed **W**etlands **2D**) (Langergraber & Šimůnek 2005, 2006) was developed to model the biochemical transformation and degradation processes in subsurface-flow constructed wetlands. The model was incorporated into the HYDRUS (2D/3D) variably-saturated water flow and solute transport software package. Constructed wetlands have become increasingly popular for removing organic matter, nutrients, trace elements, pathogens, or other pollutants from wastewater and/or runoff water. Such wetlands involve a complex mixture of water, substrate, plants, litter, and a variety of microorganisms to provide optimal conditions for improving water quality. The water flow regime in subsurface-flow constructed wetlands can be highly dynamic and requires the use of transient variably-saturated flow model. The biochemical components defined in CW2D include dissolved oxygen, three fractions of organic matter (readily- and slowly-biodegradable, and inert), four nitrogen compounds (ammonium, nitrite, nitrate, and dinitrogen), inorganic phosphorus, and heterotrophic and autotrophic micro-organisms. Organic nitrogen and organic phosphorus were modeled as part of the organic matter. The biochemical degradation and transformation processes were based on Monod-type rate expressions, such as for NO_3-based growth of heterotrophs on readily biodegradable COD (denitrification):

$$r = \mu_{\mathrm{DN}} \frac{K_{\mathrm{DN,O_2}}}{K_{\mathrm{DN,O_2}} + c_{\mathrm{O_2}}} \frac{c_{\mathrm{NO_3}}}{K_{\mathrm{DN,NO_3}} + c_{\mathrm{NO_3}}} \frac{K_{\mathrm{DN,NO_2}}}{K_{\mathrm{DN,NO_2}} + c_{\mathrm{NO_2}}} \frac{c_{\mathrm{CR}}}{K_{\mathrm{DN,CR}} + c_{\mathrm{CR}}} f_{\mathrm{N,DN}} c_{XH} \tag{25}$$

We refer to Langergraber & Šimůnek (2005, 2006) for a detailed discussion of the terms in (25). All process rates and diffusion coefficients were assumed to be temperature dependent. Heterotrophic bacteria were assumed to be responsible for hydrolysis, mineralization of organic matter (aerobic growth) and denitrification (anoxic growth), while autotrophic bacteria were assumed to be responsible for nitrification, which was modeled as a two-step process. Lysis was considered to be the sum of all decay and sink processes. Langergraber & Šimůnek (2005, 2006) demonstrated the model for one- and two-stage subsurface vertical flow constructed wetlands. Model simulations of water flow, tracer transport, and selected biochemical compounds were compared against experimental observations.

5 NUMERICAL SOLUTIONS

All models discussed in this manuscript are numerical models that use various techniques to solve the governing partial differential equations. While spatial derivatives are invariably approximated using the Galerkin-type linear finite element schemes, temporal derivatives in the models are approximated using finite differences. We always used a fully implicit (backward) finite difference

scheme for water flow, and a Crank-Nicholson time-centered scheme for heat and solute transport. The mass-conservative method proposed by Celia et al. (1990) was used in all HYDRUS models. This method has been shown to provide excellent results in terms of minimizing the mass balance error. Time and space discretization of the Richards equation generally leads to a nonlinear system of algebraic equations. These equations are then linearized and solved using the Picard iteration method. Higher-order approximations for the time derivative as derived by van Genuchten (1976) are used for discretization of the transport equations. Depending upon the size of the problem, the matrix equations resulting from discretization of the governing equations are solved using either Gaussian elimination for banded matrices, or the conjugate gradient method for symmetric matrices and the ORTHOMIN method for asymmetric matrices (Mendoza et al. 1991).

The operator-splitting approach with no iterations during one time step (non-iterative sequential approach) is used in all multicomponent models (i.e., for the major ion chemistry module of HYDRUS-1D, in HP1, and in the wetland CW2D module of HYDRUS (2D/3D)).

Three-dimensional applications in general require a large number of finite elements to discretize realistically large transport domains. Even with the fast personal computers currently available, it is still very difficult to solve within a reasonable time problems having more than about half million nodes. To decrease the required computational time, Hardelauf et al. (2007) parallelized SWMS_3D to develop PARSWMS that distributes problems with a large number of elements over multiple processors working in parallel. The PARSWMS code was developed for the LINUX or UNIX workstations with installed free-wares MPI, PETSc and PARMETIS. Hardelauf et al. (2007) demonstrated that doubling the number of processors may lead to a decrease in the computational time of up to 48%.

6 GRAPHICAL USER INTERFACES

One major problem often preventing the use of numerical codes is the extensive work required for data preparation, finite element grid design, and graphical presentation of the output results. Hence, a more widespread use of numerical models requires techniques which make it easier to create, manipulate and display large data files, and which facilitate interactive data management (Šimůnek et al. 1999). Introducing such techniques frees users from cumbersome manual data processing, and enhances the efficiency in which programs are being implemented for a particular example. To avoid or simplify the preparation and management of relatively complex input data files and to graphically display final simulation results, we were very keen in developing interactive graphics-based user-friendly interfaces (GUIs) for the MS Windows environments. From 1995 on this has resulted into several of the software packages described above. While earlier versions (until 1998) were still 16-bit applications, all software packages from 1998 on are 32-bit applications fully compatible with all Windows operating systems released after Windows 95.

7 CONCLUSIONS

Over the last 15 years the close collaboration between the University of California Riverside, and the U.S. Salinity Laboratory, and more recently with PC-Progress in Prague, Czech Republic, and SCK• CEN, Mol, Belgium, resulted in the development of a large number of computer tools that are currently being used worldwide for a variety of applications involving the vadose zone. The numerical models may be used to simulate not only one- and multi-dimensional water flow, heat transport and solute transport in variably saturated media, but also the transport of carbon dioxide, major ions, trace elements, colloids, viruses, bacteria, and wastewater. The popularity of the modeling tools is in large part due to their ease of use because of the availability of interactive graphical user interfaces. While the earlier models have proved to be useful for analyzing subsurface flow and processes, several more complex models and modules are being developed at present. One example is the HP1 code that couples the HYDRUS-1D water flow and solute transport model with the PHREEQC geochemical model. Since the resulting combined program brings together the

state-of-the-art in both flow and transport modeling and modeling of geochemical reactions, we believe that HP1 represents a very flexible and promising tool for evaluating complex interactions of various processes in the vadose zone.

ACKNOWLEDGEMENT

This paper is based on work supported in part by the Terrestrial Sciences Program of the Army Research Office (Terrestrial Processes and Landscape Dynamics and Terrestrial System Modeling and Model Integration), by the National Science Foundation Biocomplexity programs #04-10055 and NSF DEB 04-21530, and by SAHRA (Sustainability of semi-Arid Hydrology and Riparian Areas) under the STC Program of the National Science Foundation, Agreement No. EAR-9876800.

REFERENCES

Appelo, C. A. J. & Postma, D. 2005. *Geochemistry, Groundwater and Pollution*, 2nd ed., A. A. Balkema Publ., Leiden, The Netherlands, 649 p.

Bradford, S. A., Bettehar, M., Šimůnek, J. & van Genuchten, M. Th. 2004. Straining and attachment of colloids in physically heterogeneous porous media, *Vadose Zone J.* 3(2), 384–394.

Bradford, S. A., Šimůnek, J., Bettahar, M., van Genuchten, M. Th. & Yates, S. R. 2006. Significance of straining in colloid deposition: evidence and implications. *Water Resour. Res.*, 42, W12S15, doi:10.1029/2005WR004791, 16 pp.

Brooks, R. H. & Corey, A. T. 1964. Hydraulic properties of porous media. *Hydrol. Paper No. 3*, Colorado State Univ., Fort Collins, CO.

Casey, F. X. M. & Šimůnek, J. 2001. Inverse analyses of the transport of chlorinated hydrocarbons subject to sequential transformation reactions. *J. of Environ. Quality, 30*(4), 1354–1360.

Casey, F. X. M., Larsen, G. L., Hakk, H. & Šimůnek, J. 2003. Fate and transport of 17β-Estradiol in soil-water systems. *Environ. Sci. Technol., 37*(11), 2400–2409.

Castiglione, P., Mohanty, B. P., Shouse, P. J., Šimůnek, J., van Genuchten, M. Th., & Santini, A. 2003. Lateral water diffusion in an artificial macroporous system: Theory and experimental evidence. *Vadose Zone J., 2*, 212–221.

Celia, M. A., Bououtas, E. T. & Zarba, R. L. 1990. A general mass-conservative numerical solution for the unsaturated flow equation. *Water Resour. Res., 26*, 1483–1496.

de Vries, D. A. 1958. Simultaneous transfer of heat and moisture in porous media. Transactions, American Geophysical Union, 39(5), 909–916.

Dontsova, K. M., Yost, S. L., Šimůnek, J., Pennington, J. C. & Williford, C. 2006. Dissolution and transport of TNT, RDX, and Composition B in saturated soil columns. *J. of Environ. Quality, 35*, 2043–2054.

Durner, W. 1994. Hydraulic conductivity estimation for soils with heterogeneous pore structure. *Water Resour. Res., 32*(9), 211–223.

Gargiulo, G., Bradford, S. A., Šimůnek, J., Ustohal, P., Vereecken, H. & Klumpp, E. 2007. Transport and deposition of metabolically active and stationary phase *Deinococcus Radiodurans* in unsaturated porous media. *Environ. Sci. and Technol., 41*(4), 1265–1271.

Gerke, H. H., & van Genuchten, M. Th., 1993. A dual-porosity model for simulating the preferential movement of water and solutes in structured porous media. *Water Resour. Res*. 29, 305–319.

Gonçalves, M. C., Šimůnek, J., Ramos, T. B., Martins, J. C., Neves, M. J. & Pires, F. P. 2006. Multicomponent solute transport in soil lysimeters irrigated with waters of different quality. *Water Resour. Res., 42*, W08401, doi:10.1029/2006WR004802, 17 pp.

Hanson, B. R., Šimůnek, J. & Hopmans, J. W. 2006. Numerical modeling of urea-ammonium-nitrate fertigation under microirrigation. *Agric. Water Management, 86*, 102–113.

Hansson, K., Šimůnek, J., Mizoguchi, M., & Lundin, L.-Ch. 2004. Water flow and heat transport in frozen soil: Numerical solution and freeze/thaw applications. *Vadose Zone Journal, 3*(2), 693–704.

Harbaugh, A.W., Banta, E. R., Hill, M. C. & McDonald, M. G. 2000. MODFLOW-2000, the U.S. Geological Survey modular ground-water model user guide to modularization concepts and the ground-water flow process. Denver, CO, Reston, VA, U.S. Geological Survey.

Hardelauf, H., Javaux, M., Herbst, M., Gottschalk, S., Kasteel, R., Vanderborght, J. & Vereecken, H. 2007. PARSWMS: A parallelized model for simulating 3-D water flow and solute transport in variably saturated soils. *Vadose Zone Journal, 6*(2), 255–259.

Haws, N. W., Rao, P. S. C., Šimůnek, J. & Poyer I. C. 2005. Single-porosity and dual-porosity modeling of water flow and solute transport in subsurface-drained fields using effective field-scale parameters. *J. of Hydrology, 313*(3–4), 257–273.

Hendrickx, J. M. H., Dekker, L. W., & Boersma, O. H. 1993. Unstable wetting fronts in water repellent field soils. *J. Environ. Qual.* 22, 109–118.

Jacques, D. & Šimůnek, J. 2005. User Manual of the Multicomponent Variably-Saturated Flow and Transport Model HP1, Description, Verification and Examples. Version 1.0, *SCK•CEN-BLG-998,* Waste and Disposal, SCK•CEN, Mol, Belgium, 79 pp.

Jacques, D., Šimůnek, J., Mallants, D. & van Genuchten, M. Th. 2006. Operator-splitting errors in coupled reactive transport codes for transient variably saturated flow and contaminant transport in layered soil profiles. *J. Contam. Hydrology, 88*, 197–218.

Jacques, D., Šimůnek, J., Mallants, D. & van Genuchten, M. Th. 2008. Modeling coupled hydrological and chemical processes in the vadose zone: A case study on long term uranium migration following mineral p-fertilization, *Vadose Zone J.*, Special Issue "Vadose Zone Modeling", (in press).

Kodešová, R., Kozák, J., Šimůnek, J. & Vacek, O. 2005. Field and numerical study of chlorotoluron transport in the soil profile: Comparison of single and dual-permeability model. *Plant, Soil and Environment, 51*(7), 310–315.

Köhne, J. M., Köhne, S., Mohanty, B. P. & Šimůnek, J. 2004. Inverse mobile-immobile modeling of transport during transient flow: Effect of between-domain transfer and initial water content. *Vadose Zone J., 3*(4), 1309–1321.

Köhne, J. M. & Mohanty, B. P. 2005. Water flow processes in a soil column with a cylindrical macropore: Experiment and hierarchical modeling. *Water Resour. Res*., 41, W03010, doi:10,1029/2004WR003303.

Köhne, J. M., Mohanty, B. P. & Šimůnek, J. 2006. Inverse dual-permeability modeling of preferential water flow in a soil column and implications on field-scale solute transport prediction. *Vadose Zone Journal*, 5, 59–76.

Kosugi K. 1996. Lognormal distribution model for unsaturated soil hydraulic properties. *Water Resour. Res., 32*(9), 2697–2703.

Kung, K-J. S. 1990. Preferential flow in a sandy vadose zone. 2. Mechanism and implications. *Geoderma*, 46, 59–71.

Langergraber, G. & Šimůnek, J. 2005. Modeling Variably-Saturated Water Flow and Multi-Component Reactive Transport in Constructed Wetlands. *Vadose Zone Journal, 4*, 924–938.

Langergraber, G. & Šimůnek, J. 2006. The Multi-component Reactive Transport Module CW2D for Constructed Wetlands for the HYDRUS Software Package, Manual – Version 1.0, *HYDRUS Software Series 2*, Department of Environmental Sciences, University of California Riverside, Riverside, CA, 72 pp.

Mendoza, C. A., Therrien, R. & Sudicky, E. A. 1991. *ORTHOFEM User's Guide, Version 1.02*. Waterloo Centre for Groundwater Research, Univ. of Waterloo, Waterloo, Ontario, Canada.

Parkhurst, D. L. & Appelo, C. A. J. 1999. User's guide to PHREEQC (version 2) – A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. Water Resources Investigation, Report 99-4259, Denver, Co, USA, 312 pp.

Pitzer, K. S. 1979. *Activity Coefficients in Electrolyte Solutions*. Chap. 7, CRC Press, Boca Raton, Fl.

Pot, V., Šimůnek, J., Benoit, P., Coquet, Y., Yra, A. & Martínez-Cordón, M. -J. 2005. Impact of rainfall intensity on the transport of two herbicides in undisturbed grassed filter strip soil cores. *J. of Contaminant Hydrology, 81*, 63–88.

Saito, H. J., Šimůnek, J. & Mohanty, B. 2006. Numerical analyses of coupled water, vapor and heat transport in the vadose zone. *Vadose Zone Journal, 5*, 784–800.

Scanlon, B., Keese, K., Reedy, R. C., Šimůnek, J. & Andraski, B. 2003. Variations in flow and transport in thick desert vadose zones in response to paleoclimatic forcing (0–90 kyr): Monitoring, modeling, and uncertainties. *Water Resour. Res.* 39(7), 1179, doi:10.1029/2002WR001604, 13.1–13.7.

Schaerlaekens, J., Mallants, D., Šimůnek, J., van Genuchten, M. Th. & Feyen, J. 1999. Numerical simulation of transport and sequential biodegradation of chlorinated aliphatic hydrocarbons using CHAIN_2D. *J. of Hydrological Processes, 13*(17), 2847–2859.

Šejna, M. & Šimůnek, J. 2007. HYDRUS (2D/3D): Graphical User Interface for the HYDRUS Software Package Simulating Two- and Three-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Media, published online at www.-pc-progress.cz, PC-Progress, Prague, Czech Republic.

Seo, H. S., Šimůnek, J. & Poeter, E. P. 2006. Application of the Unsaturated Flow (UNSF) Package, Modflow Conference Proceedings, ed. E. Poeter, Ch. Zheng, M. Hill & J. Doherty, International Ground Water Modeling Center, Colorado School of Mines, May 22–24, Golden, CO.

Seo, H. S., Šimůnek, J. & Poeter, E. P. 2007. Documentation of the HYDRUS package for MODFLOW-2000, the U.S. Geological Survey Modular Ground-Water Model. *GWMI 2007-01*. International Ground Water Modeling Center, Colorado School of Mines, Golden, Colorado, 96 pp.

Šimůnek, J. & Suarez, D. L. 1993. Modeling of carbon dioxide transport and production in soil: 1. Model development. *Water Resour. Res., 29*(2), 487–497.

Šimůnek, J. & Suarez, D. L. 1994. Two-dimensional transport model for variably saturated porous media with major ion chemistry. *Water Resour. Res., 30*(4), 1115–1133.

Šimůnek, J., Vogel, T. & van Genuchten, M. Th. 1994. SWMS_2D simulating water flow and solute transport in two-dimensional variably saturated media. Version 1.2, *IGWMC-FOS 61,* International Ground Water Modeling Center, Colorado School of Mines, Golden, Colorado, 196 pp.

Šimůnek, J., Huang, K. & van Genuchten, M. Th. 1995. The SWMS_3D code for simulating water flow and solute transport in three-dimensional variably saturated media. Version 1.0, *Research Report No. 139,* U.S. Salinity Laboratory, USDA, ARS, Riverside, California, 155 pp.

Šimůnek, J., Šejna, M. & van Genuchten, M. Th. 1999. The HYDRUS-2D software package for simulating two-dimensional movement of water, heat, and multiple solutes in variably saturated media. Version 2.0, *IGWMC-TPS-53,* International Ground Water Modeling Center, Colorado School of Mines, Golden, Colorado, 251 pp.

Šimůnek, J., Wendroth, O., Wypler, N., & van Genuchten, M. Th. 2001. Nonequilibrium water flow characterized from an upward infiltration experiment. *European J. of Soil Sci., 52*(1), 13–24.

Šimůnek, J. & Valocchi, A. J. 2002. Geochemical Transport. In: *Methods of Soil Analysis*, Part 1, Physical Methods. Chapter 6.9, Eds. J. H. Dane & G. C. Topp, Third edition, SSSA, Madison, WI, 1511–1536.

Šimůnek, J. 2003. *HYDRUS-2D Code Modification: Modeling Overland Flow and Dynamic Interactions Between Plants and Water Flow in a Hillslope Transect*. Idaho National and Environmental Laboratory, Final report for project under contract number 00021013, 55 pp.

Šimůnek, J., Jarvis, N. J., van Genuchten, M. Th. & Gärdenäs, A. 2003. Nonequilibrium and preferential flow and transport in the vadose zone: Review and case study. *Journal of Hydrology*, 272, 14–35.

Šimůnek, J., van Genuchten, M. Th. & Šejna, M. 2005. The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media. Version 3.0, *HYDRUS Software Series 1*, Department of Environmental Sciences, University of California Riverside, Riverside, CA, 270 pp.

Šimůnek, J., Jacques, D., van Genuchten, M. Th. & Mallants, D. 2006a. Multicomponent geochemical transport modeling using the HYDRUS computer software packages. *J. Am. Water Resour. Assoc., 42*(6), 1537–1547.

Šimůnek, J., van Genuchten, M. Th. & Šejna, M. 2006b. *The HYDRUS Software Package for Simulating Two- and Three-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Media*. Technical Manual, Version 1.0, PC Progress, Prague, Czech Republic, pp. 241.

Šimůnek, J. & van Genuchten, M. Th. 2008. Modeling nonequilibrium flow and transport with HYDRUS. *Vadose Zone Journal*, Special Issue "Vadose Zone Modeling", (in press).

Šimůnek, J., van Genuchten, M. Th. & Šejna, M. 2008. Development and applications of HYDRUS and related software packages. *Vadose Zone Journal*, Special Issue "Vadose Zone Modeling", (in press).

Truesdell, A. H. & Jones, B. F. 1974. WATEQ, a computer program for calculating chemical equilibria of natural waters. *J. Res. U. S. Geol. Surv., 2*(2), 233–248.

Twarakavi, N. K. C., Šimůnek, J. & Seo, H. S. 2008. Evaluating interactions between groundwater and vadose zone using HYDRUS-based flow package for MODFLOW, *Vadose Zone J.*, Special Issue "Vadose Zone Modeling", (in press).

van Genuchten, M. Th. 1976. On the accuracy and efficiency of several numerical schemes for solving the convective-dispersive equation, in *Finite Elements in Water Resources*, edited by W. G. Gray et al., Pentech Press, London, pp. 1.71–1.90.

van Genuchten, M. Th. & Wierenga. P. J. 1976. Mass transfer studies in sorbing porous media, I. Analytical solutions. *Soil Sci. Soc. Am. J*. 40, 473–481.

van Genuchten, M. Th. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J., 44*, 892–898.

Wang, D., Yates, S. R. & Gan, J. 1997. Temperature effect on methyl bromide volatilization in soil fumigation. *Journal of Environmental Quality, 26*(4), 1072–1079.

Wang, D., Knuteson, J. A. & Yates, S. R. 2000. Two-dimensional model simulation of 1,3-dichloropropene volatilization and transport in a field soil. *Journal of Environmental Quality, 29*(2), 639–644.

Zhang, P., Šimůnek, J., & Bowman, R. S. 2004. Nonideal transport of solute and colloidal tracers through reactive zeolite/iron pellets. *Water Resour. Res., 40*, doi:10.1029/2003WR002445.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Limitations of depth-averaged modelling of shallow flows with recirculations

Peter K. Stansby
School of Mechanical, Aerospace and Civil Engineering, University of Manchester, UK

ABSTRACT: Depth-averaged modelling of shallow flows is widely used with greater computational efficiency than 3-D modelling. Limitations for recirculating flows are reviewed here and some new analysis is presented. The shallow wakes of a conical island of small slope have been predicted by a 3-D, hydrostatic-pressure, boundary-layer model using a 3-D, two-mixing-length, eddy-viscosity, turbulence model. Wakes may be stable or vortex shedding. With the horizontal mixing length about six times the vertical Prandtl values good predictions were obtained. A 2-D, depth-averaged model may have the same form of horizontal mixing but the bed shear stress can only be given by an empirical formula based on parallel flow conditions. Depth-averaged wakes show a much greater tendency to be unstable because in 3-D the horizontal and vertical strain rates modify eddy viscosity and as a consequence increase bed shear stress above that given by a formula based on parallel flow.

1 INTRODUCTION

There has been considerable analysis, experimentation and modelling of shallow wakes, jets and mixing layers, as reviewed in Jirka(2001). (As has become common, modelling here implies numerical simulation.) This relates not only to hydraulic engineering in rivers, estuaries and coasts but also to atmospheric and oceanic flows. The question of stability has been particularly prominent; that is to say whether or not coherent structures, or vortices, develop in a horizontal plane, as this strongly influences large-scale mixing processes. For mixing layers and jets this is due to shear layer instability and for wakes there is the additional influence of vorticity generation at a wet/dry (solid/fluid) interface. Observations suggest flow structures may be quite two-dimensional in shallow flow, e.g. Lloyd and Stansby(1997b). This has encouraged modelling by the two-dimensional (2-D), depth-averaged equations which have had various successes. Various formulations for horizontal diffusion/mixing have been used, ranging from the simple Elder approximation to the large eddy simulation (LES) assumption (in two dimensions). Models have also differentiated between turbulence generated by the bed and horizontal strain rates and their influence has been added to give turbulence production (Rastogi and Rodi 1978, Chu and Babarutsi 1989).

We consider shallow wakes where the bed slopes are small so that turbulence is (predominantly) bed-generated and particular effects associated with local vorticity generation by steep (in the extreme vertical) sides are avoided. This suggests that the flow may be idealised as an attached boundary-layer flow driven by horizontal pressure gradients (proportional to surface elevation gradients). This situation relates quite generally to most river, estuary and coastal flows.

We focus on the case of a conical island with sloping sides with an angle of 8°, investigated experimentally in Lloyd and Stansby(1997a), as this has the benefit of a simple geometry and boundary conditions giving a range of stable and vortex-shedding wake features. The surface was smooth and the channel bed horizontal. The flow may be conveniently defined by a stability parameter $St = C_f D_{mid}/h$ where D_{mid} is the island diameter at mid depth and C_f and h are values of

friction coefficient and depth which would occur at the island centre without the island in position, see also Chen and Jirka(1995) for cylinder flows. The early computations in Lloyd and Stansby showed limitations in depth-averaged modelling particularly in relation to the prediction of stable wakes and predictions by a 3-D boundary-layer (3-DBL) approach (following Stansby 1997) were also uncertain. Following that study a step back was taken and experiments were undertaken in laminar flow with oscillatory flow around similar islands (Lloyd, Stansby and Chen 2001). In addition the numerical modelling was improved, importantly by making the time-stepping second-order accurate to reduce/remove any influence of wave damping and numerical diffusion effects (Stansby and Lloyd 2001). This situation is much more complex than that for a current and a range of highly complex vortex-shedding formations occur depending substantially on the amplitude of the ambient flow. The 3-DBL model nevertheless gave good predictions, certainly picking up all the complex vortex interactions, and the more stable wake formations. This is significant in the present context as horizontal diffusion was due to molecular viscosity only and the mathematical formulation based on an attached boundary layer driven by hydrostatic pressure is to a large degree justified. Depth-averaged simulations were markedly less successful and unreported. As stated in Stansby and Lloyd(2001), the major remaining issue for turbulent flows was a suitable formulation for horizontal mixing.

This was investigated by Stansby(2003), returning to turbulent wake flows due to currents. A 3-D mixing-length, eddy-viscosity formulation was used, similar to that used in LES and also described, but not applied, in Rodi(1984). A novel feature was that two mixing lengths were used, the horizontal acting on the horizontal strain rates and the vertical on the vertical strain rates. The vertical mixing length follows the classical Prandtl(1927) formulation (see Rodi 1984). The horizontal mixing length was made a multiple of the vertical with the actual value determined through experimental comparison and a value of about 6 gave good predictions. That the horizontal turbulence length scale should be greater than the vertical in shallow flows is well known and the horizontal mixing length was about half the water depth (except near the bed). These mixing lengths are much smaller than the coherent structures (recirculations or vortices) computed directly, as required for turbulence modelling in unsteady flow.

Bed shear stress was calculated from the vertical velocity gradient at the bed, giving a friction coefficient based on the depth-averaged velocity to compare with the ambient steady-flow 'channel' value (given by the Blasius formula for smooth surfaces). Notably the friction coefficient was increased above the Blasius value in the wake region, sometimes markedly. This is associated with horizontal and vertical strain rates modifying eddy viscosity and hence the velocity gradients near the bed. Only in a small region near the upstream stagnation point did the friction coefficient decrease below the Blasius value.

This definition of horizontal mixing for 3-DBL modelling may be applied in depth-averaged models with the bed shear stress given by an empirical formula based on parallel flow conditions. Stansby(2006) has presented the limitations of depth-averaged modelling for these shallow wakes and the purpose of this paper is to provide a broad overview with some further insights and results.

2 MATHEMATICAL FORMULATIONS

The 3-DBL formulation of Stansby(2003) is presented first as the 2-D depth-averaged formulation of Stansby(2006) follows from it.

2.1 *3-DBL formulation*

The continuity and Navier-Stokes equations in Cartesian form are converted to σ coordinates to automatically fit the bed and the moving water surface, enabling high mesh resolution to be produced at the bed for boundary-layer resolution. The σ coordinate is defined as $\sigma = \frac{z-\eta}{h}$ where η is surface elevation, h is water depth and z is the vertical coordinate. The velocities in Cartesian coordinates are u, v, w in the x, y, z directions. In the σ coordinate system $u_\sigma = u, v_\sigma = v, w_\sigma = \omega/h$

(defined below) and the σ subscript is dropped hereafter for u, v and w. With the assumption for hydrostatic pressure p, $\frac{\partial p}{\partial \sigma} = -\rho g h$, $\frac{\partial p}{\partial x} = -\rho g \frac{\partial \eta}{\partial x}$ and $\frac{\partial p}{\partial y} = -\rho g \frac{\partial \eta}{\partial y}$ where ρ is water density and g is gravitational acceleration.

For local continuity,

$$\frac{\partial \eta}{\partial t} + \frac{\partial (hu)}{\partial x_\sigma} + \frac{\partial (hv)}{\partial y_\sigma} + \frac{\partial \omega}{\partial \sigma} = 0 \tag{1}$$

x momentum

$$\frac{\partial (hu)}{\partial t} + \frac{\partial (hu^2)}{\partial x_\sigma} + \frac{\partial (huv)}{\partial y_\sigma} + \frac{\partial (\omega u)}{\partial \sigma} = -gh\frac{\partial \eta}{\partial x} + \frac{\partial}{\partial \sigma}\left(\frac{\nu_E}{h}\frac{\partial u}{\partial \sigma}\right) + \frac{\partial}{\partial x}\left(2\nu_E h \frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(\nu_E h\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right)\right) \tag{2}$$

and y momentum

$$\frac{\partial (hv)}{\partial t} + \frac{\partial (huv)}{\partial x_\sigma} + \frac{\partial (hv^2)}{\partial y_\sigma} + \frac{\partial (\omega v)}{\partial \sigma} = -gh\frac{\partial \eta}{\partial y} + \frac{\partial}{\partial \sigma}\left(\frac{\nu_E}{h}\frac{\partial v}{\partial \sigma}\right) + \frac{\partial}{\partial x}\left(\nu_E h\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)\right) + \frac{\partial}{\partial y}\left(2\nu_E h \frac{\partial v}{\partial y}\right) \tag{3}$$

where kinematic eddy viscosity has turbulent and molecular components such that $\nu_E = \nu_t + \nu$, respectively. Note that advection terms are defined in σ coordinates and horizontal diffusion terms in Cartesian coordinates (see Stansby 2003). The vertical velocity is defined by

$$\omega = h\frac{d\sigma}{dt} = w - u\left(\sigma\frac{\partial h}{\partial x} + \frac{\partial \eta}{\partial x}\right) - v\left(\sigma\frac{\partial h}{\partial y} + \frac{\partial \eta}{\partial y}\right) - \left(\sigma\frac{\partial h}{\partial t} + \frac{\partial \eta}{\partial t}\right) \tag{4}$$

with $\omega = 0$ when $\sigma = 0, -1$, corresponding to the water surface and the bed.

The depth-integrated continuity equation is given by

$$\frac{\partial \eta}{\partial t} + \int_{-1}^{0} hu\, d\sigma + \int_{-1}^{0} hv\, d\sigma = 0 \tag{5}$$

The vertical mesh is compressed close to the bed using semi-parabolic stretching. The mesh size at the bed should be small enough for $z^+ = (z - z_b)\sqrt{u_*^2 + v_*^2}/\nu$ to be less than 5 in the laminar sublayer, so that the no-slip condition may be applied directly at the smooth bed. Here z_b is the bed elevation and $u_* = \sqrt{\nu \frac{\partial u}{\partial z}}$ and $v_* = \sqrt{\nu \frac{\partial v}{\partial z}}$ are the friction velocities in the x and y directions. Zero stress is imposed at the water surface.

The mixing-length model has been mainly applied to attached wall boundary layers with a mixing length $l = \kappa(z - z_b)$ for $(z - z_b)/\delta < \lambda/\kappa$ and $l = \lambda\delta$ for $\lambda/\kappa < (z - z_b)/\delta < 1$, where κ is the von Karman constant, typically 0.41, $(z - z_b)$ is distance from the wall, δ is boundary layer thickness and λ is a constant, typically 0.09. In order to return conditions for the viscous sublayer, the mixing length has to be 'damped' close the wall and the standard approach is to use the van Driest formula: $l = \kappa(z - z_b)[1 - \exp(-z^+/A)]$, where $z^+ = \sqrt{u_*^2 + v_*^2}(z - z_b)/\nu$ and the constant $A = 26$. In the case of the shallow-water flows here due to a steady current, the boundary-layer thickness may be assumed to be equal to the water depth h. The eddy viscosity is then defined as $\nu_t = l^2\sqrt{\left(\frac{\partial u}{\partial z}\right)^2 + \left(\frac{\partial v}{\partial z}\right)^2}$, e.g. Rodi(1984). A finite mixing length at the water surface is of course physically unrealistic but has no effect on vertical diffusion as the zero stress condition is imposed.

Rodi(1984) suggested a form of the strain rate parameter for eddy viscosity which is also widely applied in LES. This is given by $S=(2s_{ij}s_{ij})^{1/2}$ and $\nu_t=l^2S$, where $s_{ij}=\frac{1}{2}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)$ is the rate-of-strain tensor. Note that overbars are often included to represent 'mean' (or filtered in LES terms) quantities but here we are only concerned with slowly varying ('mean') values and the overbars are omitted. The rate of production of turbulent kinetic energy is given by $P=\nu_t S^2$ and energy transfer is only from the large-scale motions, computed directly, to the small-scale turbulence, which is modelled. We now propose different length scales for horizontal motion l_h and vertical motion l_v, splitting the rate of strain components for horizontal and vertical turbulence, giving a formula for turbulent viscosity:

$$\nu_t=\left(l_h^4\left[2\left(\frac{\partial u}{\partial x}\right)^2+2\left(\frac{\partial v}{\partial y}\right)^2+\left(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\right)^2\right]+l_v^4\left[\left(\frac{\partial u}{\partial z}\right)^2+\left(\frac{\partial v}{\partial z}\right)^2\right]\right)^{1/2} \tag{6}$$

$$\text{assuming } \left(\frac{\partial w}{\partial z}\right)^2 \ll \left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial v}{\partial y}\right)^2, \frac{\partial w}{\partial y} \ll \frac{\partial v}{\partial z} \text{ and } \frac{\partial w}{\partial x} \ll \frac{\partial u}{\partial z}$$

for shallow-water flows where vertical velocity w is small. The vertical length scale l_v is the standard form as defined above. The horizontal length scale is different, usually larger, than this and the simplest assumption is to assume direct proportionality defined by $l_h=\beta l_v$. The constant β has to be determined from comparison with experiment. In the case of parallel, or nearly parallel, flow, eddy viscosity reverts to its standard boundary-layer form. With $l_h=l_v$ it reverts to its correct mathematical 3-D form (with negligible vertical velocity). Although it has been stated that finite l_v is unphysical at the water surface, finite l_h is not and the basic assumption is that l_h is uniform over most of the water depth, except near the bed.

The numerical scheme which uses 2nd order time stepping has been described previously (Stansby and Lloyd 2001, Stansby 2003) and here we only mention pertinent aspects. Terms in hu and hv from Eqs.(2) and (3) are substituted into the depth-integrated continuity equation (5) and η is solved for semi-implicitly. u and v are then obtained; ω is obtained from the local continuity Eq.(1) and the Cartesian w from Eq.(4). ω is determined by horizontal velocities and only feeds back into the momentum equations. In this sense vertical motion departing from the surface and bed boundary-fitted coordinates is a secondary flow.

2.2 *2-D depth-averaged formulation*

Here a vertical profile of horizontal velocity must be given and the 'log' profile for fully developed turbulent boundary layers is implicitly assumed. The depth-averaged equations are given below.

For depth-averaged continuity,

$$\frac{\partial \eta}{\partial t}+\frac{\partial (h\bar{u})}{\partial x}+\frac{\partial (h\bar{v})}{\partial y}=0 \tag{7}$$

x momentum

$$\frac{\partial (h\bar{u})}{\partial t}+\alpha\frac{\partial (h\bar{u}^2)}{\partial x}+\alpha\frac{\partial (h\bar{u}\,\bar{v})}{\partial y}=-gh\frac{\partial \eta}{\partial x}-\frac{\tau_{bx}}{\rho}+\frac{\partial}{\partial x}\left(2\nu_E h\frac{\partial \bar{u}}{\partial x}\right)+\frac{\partial}{\partial y}\left(\nu_E h\left(\frac{\partial \bar{u}}{\partial y}+\frac{\partial \bar{v}}{\partial x}\right)\right) \tag{8}$$

and y momentum

$$\frac{\partial (h\bar{v})}{\partial t}+\alpha\frac{\partial (h\bar{u}\,\bar{v})}{\partial x}+\alpha\frac{\partial (h\bar{v}^2)}{\partial y}=-gh\frac{\partial \eta}{\partial y}-\frac{\tau_{by}}{\rho}+\frac{\partial}{\partial x}\left(\nu_E h\left(\frac{\partial \bar{v}}{\partial x}+\frac{\partial \bar{u}}{\partial y}\right)\right)+\frac{\partial}{\partial y}\left(2\nu_E h\frac{\partial \bar{v}}{\partial y}\right) \tag{9}$$

where the overbar denotes depth-averaged. The factor α for the advection terms is the small correction of 1.016 due the 'log' velocity profile, and its deviation from unity has negligible effect.

The bed shear stress is given by $(\tau_{bx}, \tau_{by}) = \frac{1}{2}\rho C_f(\bar{u}, \bar{v})\sqrt{\bar{u}^2 + \bar{v}^2}$ (note there was a typographical error in Stansby(2006) with the factor $\frac{1}{2}$ missing) where $C_f = 0.0559 Re_h^{-0.25}$ (the Blasius formula) and the depth Reynolds number $Re_h = \sqrt{\bar{u}^2 + \bar{v}^2}h/\nu$. The eddy viscosity $\nu_E = \nu_t + \nu$ equivalent to Eq.(6) is given by

$$\nu_t = \left(l_h^4 \left[2\left(\frac{\partial \bar{u}}{\partial x}\right)^2 + 2\left(\frac{\partial \bar{v}}{\partial y}\right)^2 + \left(\frac{\partial \bar{v}}{\partial x} + \frac{\partial \bar{u}}{\partial y}\right)^2 \right] + (\gamma u_* h)^2 \right)^{1/2} \tag{10}$$

where $u_* = \sqrt{|\tau_b|/\rho}$ and γ is the Elder constant of approximately 0.067 to define the depth-averaged vertical mixing. The horizontal mixing length $l_h = \beta\lambda h$.

The numerical scheme is again 2nd order accurate in time with η determined semi-implicitly by the Crank-Nicolson method, following substitution of $h\bar{u}$ and $h\bar{v}$ from Eqs.(8) and (9) into (7), and advection and diffusion by the Adams-Bashforth scheme. QUICK upwinding is used for advection. Spatial discretisation is on a staggered mesh, effectively a finite-volume method. This closely mirrors the 3-D scheme.

3 DEPTH-AVERAGED AND 3-DBL RESULTS

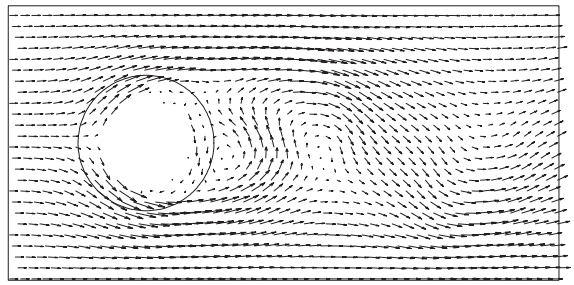
The predominant factor in determining the nature of the wake is the stability parameter $St = C_f D_{mid}/h$. For the 8° slope island under consideration there is a critical value close to 0.4 above which the wake is stable and below which vortex shedding occurs in some form. The 3-DBL model was compared with experiment in Stansby(2003) showing how the critical St was closely reproduced with $\beta = 6$ and the form of vortex shedding was at least qualitatively reproduced for lower values of 0.26 and 0.36. The island base diameter was 0.75 m and the flow parameters are given in Table 1. The outflow depth h_{out} and the inflow flux Uh_{out} are prescribed. h is the depth adjacent to the island and $U = 0.1$ m/s for all cases. The computational domain is 4.56 m long and 1.52 m wide; the island centre is 0.76 m from the upstream end. The bed is horizontal to correspond with experiment. The depth Reynolds number, $R_h = Uh_{out}/\nu$, for the experiments (and 3-DBL modelling) is always greater than 1000 giving turbulent flow. St values larger than in the experiments will be required for depth-averaged modelling and, although R_h is reduced below 500, turbulent flow conditions are assumed for consistent comparison. The free-stream Froude number, $Fr = U_c/\sqrt{gh}$, where $U_c (= Uh_{out}/h)$ is what the mean velocity at the cylinder centre would be without the cylinder in position, is subcritical. A horizontal mesh of 300 × 100 was used with 10 cells in the vertical, the cell size at the bed giving $z^+ < 5$. Mesh refinement was shown to have negligible influence on the results in Stansby(2003). In all cases asymmetry was set up by setting the cross velocity at inflow, $\nu/U = 0.5\sin(\pi t)$ for $t < 1$s. Without this asymmetry the same wakes

Table 1. Flow parameters for the island of 8° slope. Cases 1–3 are for 3-DBL modelling and cases 1–8 for 2-D depth-averaged modelling.

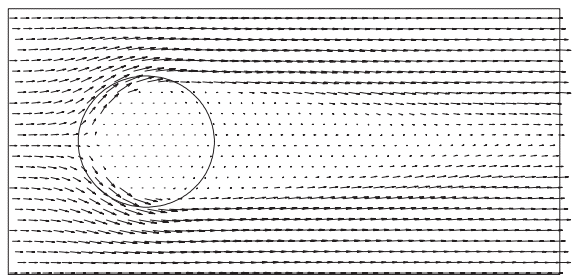
Case	h_{out}(m)	h (m)	$D_{mid}(m)$	R_h	C_f	Fr	St
1	0.019	0.0200	0.607	1900	0.0085	0.21	0.26
2	0.0145	0.0160	0.636	1450	0.0090	0.23	0.36
3	0.01325	0.0148	0.644	1325	0.0093	0.23	0.40
4	0.009	0.0107	0.674	900	0.0102	0.26	0.64
5	0.008	0.0099	0.680	800	0.0105	0.26	0.72
6	0.0065	0.00875	0.688	650	0.0111	0.25	0.87
7	0.0055	0.0080	0.693	550	0.0115	0.25	1.00
8	0.0045	0.00715	0.700	450	0.0121	0.24	1.18

structures resulted, but more slowly with vortex shedding. Froude numbers based on local velocity magnitude and depth were shown to be everywhere subcritical for all cases.

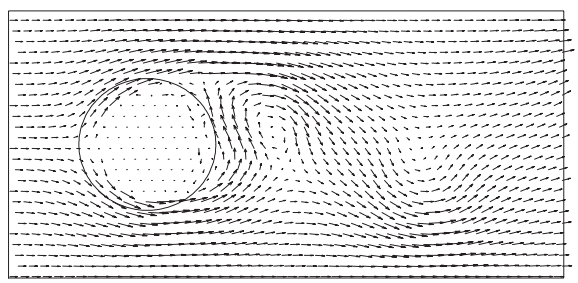
For the stable wake with $St = 0.40$ in the 3-DBL model, the distance to the downstream stagnation point from the island centre (the wake length l_w) was 1.51 m underestimating the experimental value of 1.75 m by 14%. Snapshots of surface vector are shown in Figure 1 for $St = 0.26$ and 0.40 from the 3-DBL model and for$St = 0.40$ and 0.87 from the depth-averaged model. Strong vortex shedding



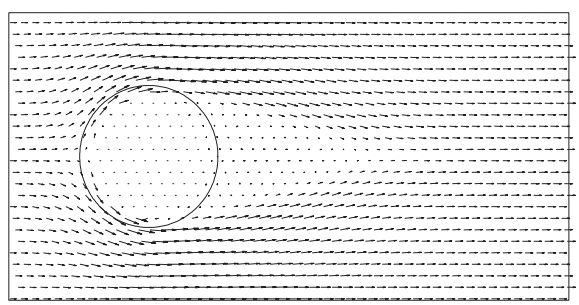
3-D: $St = 0.26$



3-D: $St = 0.40$



2-D: $St = 0.40$



2-D: $St = 0.87$

Figure 1. Surface vector plots for 8° slope island for 3-DBL and 2-D models. The dashed line shows the island base of 0.75 m diameter, from Stansby(2006) with permission of ASCE.

is clearly seen at $St = 0.40$ for the latter and the stable wake with $St = 0.87$ has a much reduced wake length of 1.07 m. To give some quantitative indication of the strength of vortex shedding the cross velocity (v) variation with time for a point 0.73 m downstream of the island centre was output. Cross-flow wake fluctuation is prominent at this position when vortex shedding occurs. Amplitudes of velocity are shown in Figure 2 for the 3-DBL case and the depth-averaged case. The critical values of St mentioned above for $\beta = 6$ can be seen. for the depth-averaged case increasing β to a high value of 11 reduces the critical St to about 0.6 and reducing it to zero increases it to about 1.1. The amplitude for the lower St values is highly modulated as shown by the error bars. Clearly the depth-averaged model gives poor predictions of critical St and of wake length when stable. However when prominent vortex shedding occurs the velocity amplitude is quite similar to the 3-D case with the same β value of 6. To demonstrate this further Figure 3 shows corresponding cross velocity variations with time for the depth-averaged and 3-DBL cases with $St = 0.26$. The amplitudes are very similar and the frequency for the depth-averaged case is slightly higher. The transient variations to a (nearly) periodic state are however somewhat different. To compare further, surface velocity vector plots are shown in Figure 4 at the same phase (when the cross velocity in Figure 3 is maximum) for the depth-averaged and 3-D cases. The wakes are quite similar although with generally higher wake velocities for the depth-averaged case, noticeably close to the wet/dry interface which is consistent with a greater friction coefficient in the 3-DBL case.

Friction coefficients were investigated for $St = 0.4$ using the 3-DBL model in Stansby(2003). Values based on local depth-averaged velocity magnitude were compared with the onset Blasius value and shown to be much greater in the wake region, up to at least ten times greater with a stable wake. It is now pertinent to compare with the local Blasius value, as used in depth-averaged modelling. Contour plots of this ratio are shown in Figure 5 to be still greater than two over most of the wake region and to be ten or more on the edge of the wake and near the wet/dry interface. These regions coincide with high horizontal shear and strain rate and thus high eddy viscosity.

It might be thought that 2-D depth-averaged models would be more realistic for small bed slopes being closer to gradually varying uniform conditions. The case of an island with a slope of 4° is thus considered. The island base diameter is now 1.5 m. The bed slope is set to give the normal depth without the island in position. The flow parameters are given in Table 2 with the same definitions as for Table 1. Note that h and h_{out} are now closer as the bed slope is set for normal conditions. Snapshots of surface velocity vector plots from the 3-DBL model are shown in Figure 6 for $S = 0.35$,

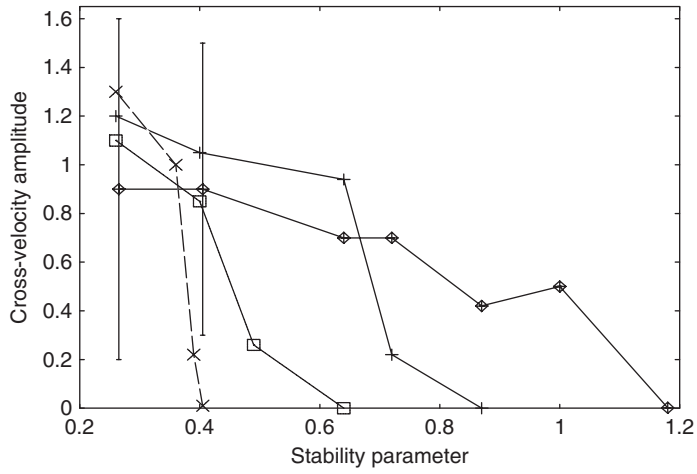


Figure 2. Variation of cross velocity amplitude, normalised by $U = 0.1$ m/s, against stability parameter for island with 8° slope, from Stansby(2006) with permission of ASCE. – – – 3-DBL; $\times\ \beta = 6$; —— 2-D; $\square \beta = 11$; $+\ \beta = 6$; $\diamond \beta = 0$ (error bars indicate range of amplitude modulation).

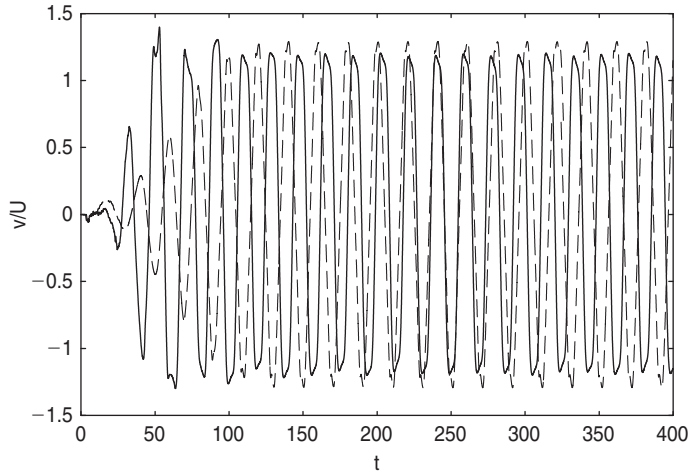


Figure 3. Variations of v/U with time t(s) at 0.73 m downstream of the island centre for the 8° slope island comparing 2-D and 3-DBL models. —— 2D; – – –3D, from Stansby(2006) with permission of ASCE.

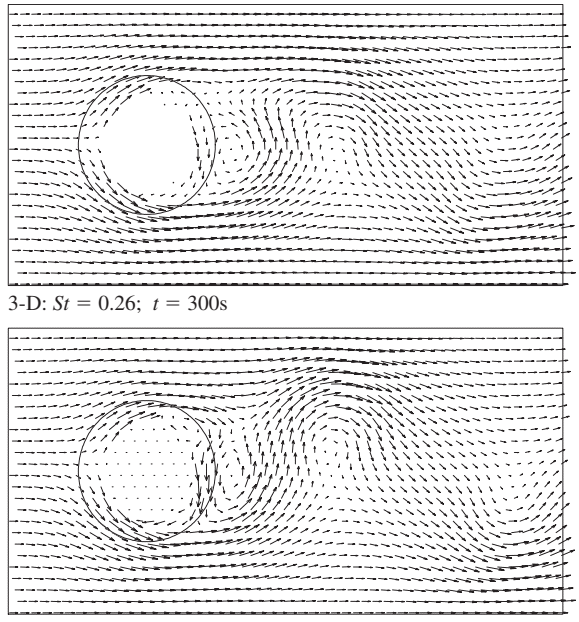


Figure 4. Surface vector plots for 8° slope island from 3-DBL and 2-D models at $St = 0.26$. The dashed line shows the island base of 0.75 m diameter, from Stansby(2006) with permission of ASCE.

0.40 and 0.53. There is strong vortex shedding for $St = 0.35$ as there was with the 8° slope island. However there is a difference in that there is now weak downstream wake unsteadiness for $St = 0.40$ and 0.53. This was observed for $St = 0.39$ for the 8° slope island, also in experiments for a vertical cylinder (Chen and Jirka 1995). It is referred to as an unsteady wake bubble. Basically it appears as a much more local phenomenon in terms of St for the island of 8° slope than it is for the island of 4°

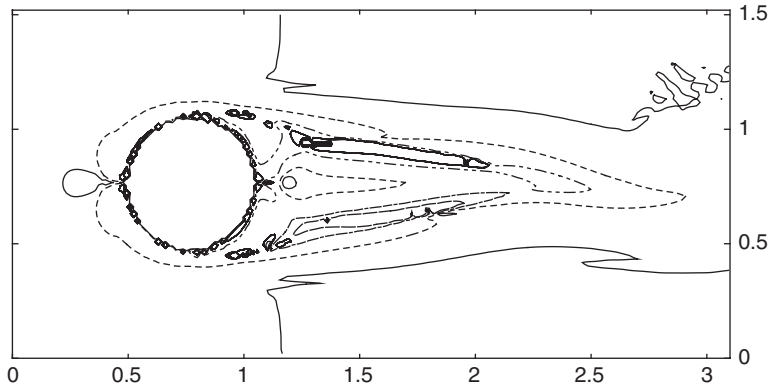


Figure 5. Contour plots of ratio of bed shear from the 3-DBL model to that given by the Blasius formula based on local conditions, for $St = 0.40$ and $\beta = 6$: —— 1; - - - - - 2; — - - — 5; — — — 10.

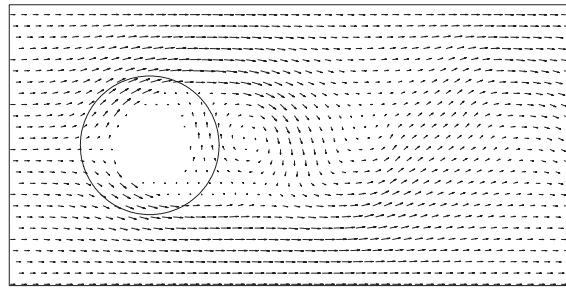
Table 2. Flow parameters for the island of 4° slope. Cases 1–7 are for 3-DBL modelling and cases 1–10 for 2-D depth-averaged modelling.

Case	h_{out}(m)	h (m)	D_{mid}(m)	R_h	C_f	Fr	St
1	0.031	0.0312	1.05	3100	0.0075	0.18	0.25
2	0.025	0.0255	1.13	2500	0.0079	0.20	0.35
3	0.023	0.0234	1.16	2300	0.0081	0.20	0.40
4	0.022	0.0224	1.18	2200	0.0082	0.21	0.43
5	0.020	0.0204	1.21	2000	0.0084	0.22	0.50
6	0.019	0.0194	1.22	1900	0.0085	0.225	0.53
7	0.017	0.0175	1.25	1700	0.0087	0.235	0.62
8	0.015	0.0156	1.28	1500	0.0090	0.245	0.74
9	0.013	0.0133	1.31	1300	0.0093	0.27	0.92
10	0.011	0.0113	1.34	1100	0.0097	0.29	1.15

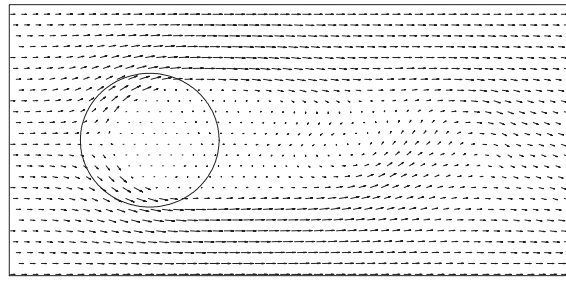
slope. For the depth-averaged model with 4° slope the critical St is again markedly overestimated and plots of cross velocity amplitude at 1.46 m downstream of the island centre against St clearly show this in Figure 7, although again, when there is prominent vortex shedding, the magnitudes are quite similar. The length of the depth-averaged stable wake bubble is again much smaller than for the 3-DBL model.

4 DIAGNOSIS USING THE 3-DBL MODEL

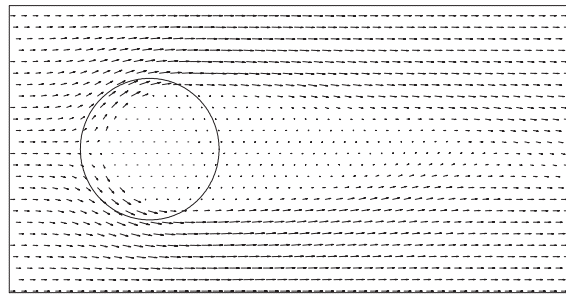
Prominent differences between 3-DBL and depth-averaged modelling is that for the latter the bed shear stress is from an empirical uniform flow formula and there is no secondary vertical flow. It is possible to diagnose the influence of each in the stable wake by modifying the 3-DBL model; the 8° slope island with $St = 0.4$ is used. First setting the horizontal length scale to zero removes the influence of horizontal strain rates from eddy viscosity, including vertical mixing and the critical St increases to 0.55 and wake length reduces to 1.31 m. These changes are consistent with changes in the 2-D depth-averaged model, as set out in Table 3. Setting the boundary-fitted vertical velocity ω to zero, rather than determining from Eq.(1), does not change the critical St but reduces the wake length slightly to 1.43 m. The former effect is clearly more significant and it appears that the 2-D



3-D: $St = 0.35$



3-D: $St = 0.40$



3-D: $St = 0.53$

Figure 6. Surface vector plots for 4° slope island from 3-DBL model. The dashed line shows the island base of 1.5 m diameter.

depth-averaged model overestimates critical *St* and underestimates wake length mainly because the influence of horizontal and vertical strain rates on eddy viscosity and hence bed shear is missing.

To show the nature of the weak vertical velocity, velocity vector plots in the centreline vertical *x*-*z* plane downstream of the island are shown in Figure 8. Note that there is considerable vertical distortion and typically 1 in 7 vectors are plotted. The prominent feature here is the downwards velocity at the downstream stagnation point (the end of the wake bubble). The cross *y*-*z* vertical plane vectors though this point are shown in Figure 9. Here there are horizontal cross velocities towards the centreline and the very small magnitudes of vertical velocity should be noted with a maximum of 1.3×10^{-4} m/s.

5 CONCLUSIONS

A mixing-length, eddy-viscosity formulation for horizontal mixing has been incorporated in a depth-averaged, 2-D shallow-flow model as a reduced form of a 3-DBL model with a 3-D,

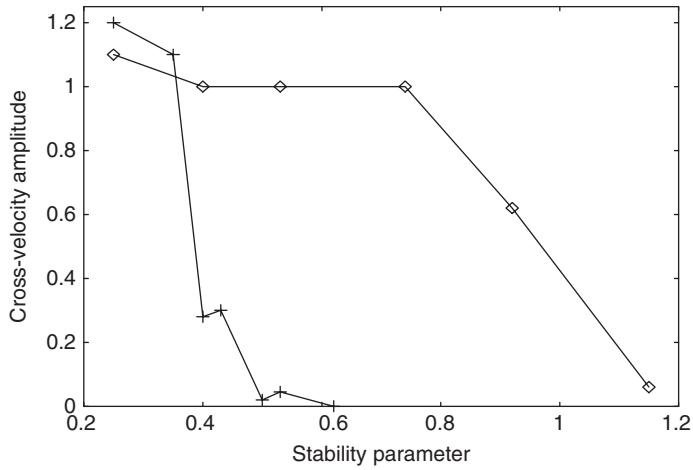


Figure 7. Cross velocity amplitude, normalised by $U = 0.1$ m/s, against stability parameter for island with 4° slope, both with $\beta = 6$. + 3-DBL; ◇ 2-D.

Table 3. The influence of horizontal mixing length factor β and vertical σ-velocity ω on the stability parameter St and wake length l_w.

	Ω	β	Critical St	Wake length l_w (m)
Experiment			0.40	1.75
3-D model	from Eq.(1)	6	0.40	1.51
3-D model	from Eq.(1)	0	0.55	1.31
3-D model	$\omega = 0$	6	0.40	1.43
2-D model	–	6	0.8	1.1
2-D model	–	0	1.1	0.91

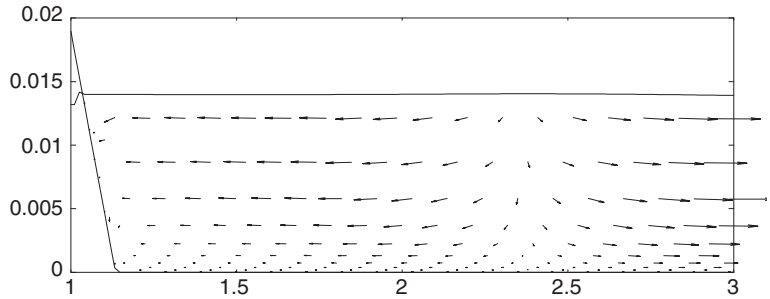


Figure 8. Velocity vector plot in vertical (x–z) plane downstream of the 8° slope island on its centreline. Maximum velocity magnitudes: u, 0.050 m/s; w, 0.00024 m/s. Dimensions in metres.

two- mixing-length, eddy-viscosity, turbulence model. While prediction of stable wakes of a conical island of small 8° slope by the 3-DBL model was good, prediction by the depth-averaged model was poor, with the critical stability parameter being considerably overestimated and the wake length underestimated. There were equivalent results for the conical island of smaller 4° slope. However,

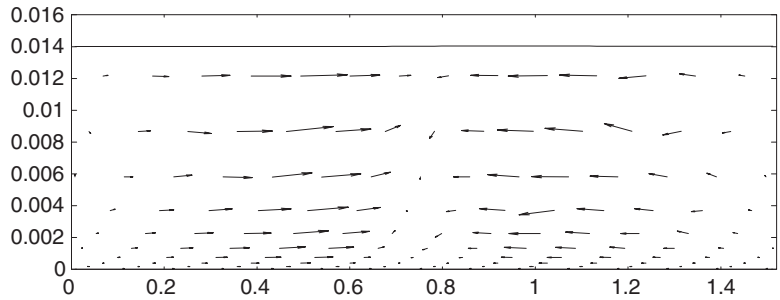


Figure 9. Velocity vector plot in vertical cross-stream (y–z) plane through the wake stagnation point at 1.51 m downstream of the 8° slope island centre. Maximum velocity magnitudes: v, 0.021 m/s; w, 0.00013 m/s. Dimensions in metres.

when vortex shedding was prominent in both depth-averaged and 3-DBL model wakes, their structures were quite similar. In the 3-DBL model, wake stability is increased with increased bed shear resulting from horizontal and vertical strain rates modifying eddy viscosity. This will be a general characteristic of flows with recirculation, or horizontal shear, and thus its inevitable omission in depth-averaged modelling will be a general limitation.

REFERENCES

Chen, D. and Jirka, G.H. 1995 Experimental study of plane turbulent wakes in a shallow water layer, Fluid Dyn. Res., **16**, 11–41.

Chu, V.H. and Babarutsi, S. 1989 Modelling of turbulent mixing layers in shallow open channel flows, Proc. 23rd IAHR Congress, Ottawa, 191–198.

Ingram, R.G. and Chu, V.H. 1987 Flow around islands in Rupert Bay: An investigation of the bottom friction effect, J. Geophys. Res., **92**(C13), 14521–14533.

Jirka, G.H. 2001 Large scale structures and mixing processes in shallow flows, J. Hydr. Res., **39**, 567–573.

Lloyd, P.M. and Stansby, P.K. 1997a Shallow-water flow around model conical islands of small side slope. I: Surface-piercing, J. Hydr. Engrg., **123**, 1057–1068.

Lloyd, P.M. and Stansby, P.K. 1997b Shallow-water flow around model conical islands of small side slope. II: Submerged, J. Hydr. Engrg., **123**, 1068–1077.

Lloyd, P.M., Stansby, P.K. and Chen, D. 2001 Wake formation around islands in oscillatory laminar shallow-water flows: Part I Experimental investigation, J. Fluid Mech., **429**, 217–238.

Prandtl, L. 1927 Über die ausgebildete turbulenz, ZAMM, **5**, 137–138.

Rastogi, A.K. and Rodi, W. 1978 Predictions of heat and mass transfer in open channels, J. Hydr. Engrg., **104**, 397–419.

Rodi, W. 1984 Turbulence models and their applications in hydraulics, IAHR Monograph, 2nd edition.

Stansby, P.K. 1997 Semi-implicit finite-volume shallow-water flow and solute transport solver with $k-\varepsilon$ turbulence model, Int. J. Numer. Meth. Fluids, **25**, 285–313.

Stansby, P.K. 2003 A mixing-length model for shallow turbulent wakes, J. Fluid Mech., **495**, 369–384.

Stansby, P.K. 2006 Limitations of depth-averaged modelling for shallow wakes, J. Hydr. Engrg., **132**, 737–740.

Stansby, P.K. and Lloyd, P.M. 2001 Wake formation around islands in oscillatory laminar shallow-water flows: Part II Boundary-layer modelling, J. Fluid Mech., **429**, 239–254.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Depth averaged turbulence models and source terms

Luis Cea
Environmental and Water Engineering Group, Universidad de A Coruña, Spain

María-Elena Vázquez-Cendón
Applied Mathematics Department, Universidad de Santiago de Compostela, Spain

ABSTRACT: Depth averaged models are widely used in engineering practice in order to model environmental flows in river and coastal regions, as well as shallow flows in hydraulic structures. This paper deals with two issues of depth averaged numerical modelling: turbulence models and the discretisation of source terms. The most important and widely used depth averaged turbulence models are reviewed and discussed, and the upwind discretisation of the bed slope source term is justified. Several practical applications are presented, and the numerical results are compared with experimental data.

1 INTRODUCTION

This paper is focused on the numerical modelling of quasi-2D free surface turbulent flows. The two-dimensional character of a free surface flow is usually enforced by a horizontal length scale much larger than the vertical one, and by a velocity field quasi-homogeneous over the water depth. Under these conditions the 3D Reynolds averaged Navier-Stokes equations can be simplified in order to obtain the depth averaged shallow water equations. Here, the term *shallow* refers to a small ratio between the vertical and horizontal length scales. Shallow flows appear in many engineering applications, mainly in river and coastal engineering, but also in certain hydraulic structures like open channels or sedimentation tanks, just to cite some examples.

The modelling of turbulence in shallow water flows has not been treated so profusely as in other fluid dynamics areas. Some depth averaged turbulence models have been proposed for the 2D-SWE. Those models derive from well known RANS turbulence models, including somehow the effects of bed friction in the turbulence field. Special mention should be given to the depth averaged $k-\varepsilon$ model proposed by Rastogi and Rodi in 1978 (Rastogi and Rodi 1978), which was the first depth averaged two-equation eddy viscosity model, and it is still the most commonly used with the 2D-SWE when turbulent effects are included in the computation.

The main effect of turbulence is to diffuse the velocity field. This turbulent diffusion, which is a *physical* effect, is added in the solution to the *numerical* diffusion inherent to the upwind schemes which are used for solving hyperbolic equations. For this reason, when modelling turbulence, it is even more important to use high resolution schemes for the convective flux, in order to reduce the numerical diffusion to the minimum without numerical instabilities. A first order scheme, which might be suitable for some problems, should not be used when turbulent effects are significant, because the numerical diffusion might be of the same order or even larger than the turbulent diffusion. In this paper we use an hybrid second-order/first-order scheme (first order in the water depth and second order in the unit discharge). The hybrid scheme uses a second order discretisation for the two unit discharge components, whilst keeping a first order discretisation for the water depth. In such a way the numerical diffusion is considerably reduced, without a significant reduction on

the numerical stability of the scheme. In order to avoid spurious oscillations of the free surface when the bathymetry is irregular, an upwind discretisation of the bed slope source term is used (García-Navarro and Vázquez-Cendón 2000; Bermúdez and Vázquez-Cendón 1994).

In order to show the capabilities of depth averaged models in the computation of turbulent flows, a finite volume model has been used to model the flow in a coastal estuary, in an open channel with a 90° bend and in a vertical slot fishway. The estuary considered has extensive flat marsh areas which flood and dry periodically due to the tidal driven flow. This makes it possible to test the turbulence models in the presence of wet-dry fronts in complex two-dimensional geometries with a very irregular bathimetry. The 90° bend has a simple geometry, but at the same time it presents a recirculation bubble, which makes it well suited for testing and comparing different turbulence models. The flow in the fishway is highly turbulent and it has strong recirculation regions, which makes it a perfect test case for the turbulence models. All the numerical results are compared with extensive experimental data.

2 DEPTH AVERAGED MODELS FOR FREE SURFACE FLOW

The depth averaged shallow water equations, also known as 2D shallow water equations (2D-SWE), are obtained after vertical integration of the 3D shallow water equations (3D-SWE). The depth averaged formulation has been successfully applied to different problems, obtaining quite accurate results with a relatively low computational cost when compared with the 3D approach. It has also the advantage of being very robust for computing accurately the water depth, even in unsteady problems with free surface shocks, as it happens to be the case in hydraulic jumps and in dam break simulations. The treatment of unsteady wet-dry fronts, which appear typically in coastal regions and in flooding problems, is also much simpler and stable than in the 3D approach. The 2D depth averaged formulation has been extensively used to model the dam break problem (Brufau and García-Navarro 2000), the propagation and runup of shallow water waves (Cea et al. 2004; Hubbard and Dodd 2002), flooding and drying problems (Brufau et al. 2004), free surface flow in hydraulic structures (Cea et al. 2007; Molls and Chaudhry 1995), flow in rivers and coastal regions (Cea et al. 2006; Ding et al. 2004; Winterwerp et al. 2002), and sediment transport in open channels and reservoirs (Wu 2004).

The mathematical derivation of the 2D-SWE can be found in many hydrodynamic books, and has already been presented by many authors. Minor differences appear from one derivation to another, but basically the process consists in assuming an hydrostatic pressure distribution, integrating the horizontal RANS equations over the water depth, applying Leibnitz's rule, and using the kinematic free surface and bed surface conditions. Several approximations are done through the mathematical derivation. It is important to have in mind these approximations in order to understand the limitations of the equations and to interpret the results obtained from them correctly. The approximations made are summarised in the following:

- Hydrostatic pressure:
  The hydrostatic pressure distribution is mainly a consequence of assuming a separation of the vertical and horizontal characteristic scales, which is a typical characteristic of quasi-2D flows.
- Homogeneous behaviour over the water depth:
  This approximation implies that the values of both the velocity and the Reynolds stresses are almost independent of the vertical coordinate. This is equivalent to neglect the lateral and longitudinal dispersion effects, and it cannot be assumed as frequently as the hydrostatic pressure assumption. The error introduced due to this approximation is strongly dependent on the existence of secondary vertical currents, which appear typically when the curvature effects in the velocity field are important. An attempt to account for the dispersion effects in the 2D-SWE is to increase the effective eddy viscosity. However, in that case the value of the effective viscosity must be fixed on empirical basis, and it depends on the problem geometry, as well as on the flow conditions.

- Depth averaged viscous and turbulent stresses:
  Slightly different formulations of the viscous and turbulent diffusion terms can be found in the literature depending on the author's preference. The most common formulation is given by:

$$\frac{\partial}{\partial x_j}\left((\nu+\nu_t)h\frac{\partial U_i}{\partial x_j}\right) \qquad i=1,2 \tag{1}$$

  With the former approximations, the turbulent depth averaged shallow water equations are obtained as:

$$\frac{\partial h}{\partial t}+\frac{\partial hU_j}{\partial x_j}=0 \tag{2}$$

$$\frac{\partial hU_i}{\partial t}+\frac{\partial hU_iU_j}{\partial x_j}=-gh\frac{\partial h}{\partial x_i}-gh\frac{\partial z_b}{\partial x_i}-\frac{\tau_{b,i}}{\rho}+\frac{\partial}{\partial x_j}\left(\nu h\frac{\partial U_i}{\partial x_j}\right)-\frac{\partial h\overline{u_i'u_j'}}{\partial x_j} \qquad i=1,2$$

  Additional terms can be added to Equation 2 in order to account for differences in atmospheric pressure and wind shear stress in the free surface. Those terms have not been included in Equation 2 for clarity in the notation.

3 DEPTH AVERAGED TURBULENCE MODELS FOR SHALLOW FLOWS

The depth averaged horizontal Reynolds stresses $\overline{u'^2}, \overline{u'v'}, \overline{v'^2}$ appearing in the 2D-SWE need to be computed by means of a depth averaged turbulence model, which are usually derived from RANS turbulence models by introducing in some way the effects of bed friction and shallowness in the turbulence field. The most simple turbulence models used with the 2D-SWE are the depth averaged parabolic model and the depth averaged mixing length model. Both of them consider an equilibrium state of turbulence, and compute the eddy viscosity with simple algebraic expressions assuming a turbulent length scale dependent on the water depth. Both are local models, in the sense that the eddy viscosity depends only on the local flow field, and it is not transported with the flow. The parabolic model accounts only for the turbulence generated by bed friction, while the mixing length model accounts also for the horizontal strain-rate production of turbulence. In order to account for the production, transport, and dissipation of turbulence, Rastogi and Rodi (Rastogi and Rodi 1978) proposed a depth averaged $k-\varepsilon$ model for the 2D-SWE, which is the most used depth averaged model nowadays in practical applications when turbulent effects are significant. Aiming to account for non-isotropic turbulence, Cea et al. (Cea et al. 2007) proposed and used in several applications a depth averaged algebraic stress model as an extension of the $k-\varepsilon$ model. All the former turbulence models are described in this section.

3.1 *The depth averaged parabolic eddy viscosity model*

This is the simplest turbulence model for shallow flows. It assumes an equilibrium state of the flow, in which the vertical shear stress and the hydrostatic pressure gradient are in perfect balance. By assuming uniform flow the convection terms in the RANS equations can be neglected, and the x-momentum equation reduces to a balance between vertical shear and pressure gradient:

$$\frac{1}{\rho}\frac{\partial P}{\partial x}=\frac{\partial \tau_{xz}}{\partial z}=\nu_t'\frac{\partial \overline{u}}{\partial z} \tag{3}$$

where ρ is the density of water, $P=\rho g(z_s-z)$ is the hydrostatic pressure, z_s is the free surface elevation, and $\nu_t'(z)$ is the eddy viscosity (not its depth averaged value). Assuming a logarithmic velocity profile in depth, the model predicts a parabolic profile in the vertical coordinate for the eddy viscosity (Cea 2005), which is given by:

$$\nu_t'(z)=u_f\kappa z\left(1-\frac{z}{h}\right) \tag{4}$$

where u_f is the bed friction velocity $\kappa = 0.41$ is the von Karman's constant, and h is the water depth. The depth averaged eddy viscosity is obtained after vertical integration of Equation 4 over the water depth, which yields:

$$v_t = \frac{1}{h}\int_{z_b}^{z_s} u_f \kappa z \left(1 - \frac{z}{h}\right) dz = \frac{1}{6}\kappa u_f h = 0.068 u_f h \tag{5}$$

The parabolic eddy viscosity model does not account for the effect of horizontal velocity gradients, but only for the turbulence generated by bed friction. It leads to very low eddy viscosity values when the turbulence production due to horizontal shear is important. It does not account for transport and dissipation processes. Despite its simplicity, it has been used by Hsieh and Yang (Hsieh and Yang 2003) and Duan (Duan 2004) in the computation of simple channel flows. It is also used by more complex models in order to account for the turbulence produced by bed friction.

3.2 *The depth averaged mixing length model*

This is a depth averaged version of the original mixing length model. In order to account for both the horizontal and the vertical production of turbulence, the total eddy viscosity is split into an horizontal v_t^h and a vertical v_t^v component. The horizontal eddy viscosity accounts for the turbulence produced by horizontal shear. It is computed as:

$$v_t^h = l_s^2 \sqrt{2 S_{ij} S_{ij}} = l_s^2 \sqrt{2(S_{11}^2 + S_{22}^2 + 2S_{12}^2)} \tag{6}$$

where l_s is the characteristic horizontal turbulent length scale, and S_{ij} is the horizontal mean strain-rate tensor computed from the depth averaged velocity as:

$$S_{ij} = \frac{1}{2}\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\right) \tag{7}$$

On the other hand, the vertical eddy viscosity is generated by the vertical velocity gradient produced by bed friction. Therefore, it is computed from the parabolic eddy viscosity model as:

$$v_t^v = \frac{1}{6}\kappa u_f h \tag{8}$$

The total eddy viscosity is evaluated from the horizontal and vertical values as:

$$v_t = \sqrt{(v_t^h)^2 + (v_t^v)^2} = l_s^2 \sqrt{2 S_{ij} S_{ij} + \left(\frac{1}{6}\frac{\kappa u_f h}{l_s^2}\right)^2} \tag{9}$$

In the inner part of the domain the turbulent length scale l_s is assumed to be dependent on the water depth, since it is a restriction on the size of the turbulent eddies. In order to estimate l_s, a fictitious turbulent length derived from the parabolic eddy viscosity model is averaged over the vertical coordinate, which yields (Cea 2005):

$$l_s \approx 0.267 \kappa h \tag{10}$$

It should be remarked that, as it was pointed out in section 2, there might be 2D turbulent structures with a l_s larger than the water depth. Therefore, the assumption of l_s dependent on the water depth can often lead to an underestimation of v_t, specially in flows in which turbulence is mostly 2D.

Equation 10 is not valid near the walls, as it would predict very large length scales. Instead, it is more correct to use the wall distance as the turbulent length scale in the near wall regions. The final expression for the eddy viscosity given by the model is:

$$\nu_t = l_s^2 \sqrt{2\left(\frac{\partial U}{\partial x}\right)^2 + 2\left(\frac{\partial V}{\partial y}\right)^2 + \left(\frac{\partial U}{\partial y} + \frac{\partial V}{\partial x}\right)^2 + \left(2.34\frac{u_f}{\kappa h}\right)^2} \tag{11}$$

$$l_s = \min\left(0.267\,\kappa h, \kappa d_{wall}\right)$$

where d_{wall} is the distance to the nearest wall. In the absence of horizontal velocity gradients, Equations 11 and 5 are equal. Hence, the depth averaged mixing length model tends to the parabolic eddy viscosity model when turbulence is mainly produced by bed friction. This is the case of flows with a very low water depth, a relatively rough bed surface, and small horizontal velocity gradients. The depth averaged mixing length model has been used in channel and river flow computations in (Cea et al. 2007; Cea et al. 2006; Jia and Wang 1999), among others.

3.3 *The $k-\varepsilon$ model of Rastogi and Rodi*

The first $k-\varepsilon$ model for shallow water flows was proposed by Rastogi and Rodi (Rastogi and Rodi 1978). Instead of solving for the three-dimensional turbulence and dissipation, it solves for their depth averaged value. The equations of the model are very similar to those of the standard 2D $k-\varepsilon$ model. The main difference is that it is necessary to introduce a production term in order to account for the production of turbulence due to bed friction. The equations of the model are given by:

$$\frac{\partial hk}{\partial t} + \frac{\partial U_j hk}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\left(\nu + \frac{\nu_t}{\sigma_k}\right) h\frac{\partial k}{\partial x_j}\right) + hP_k + hP_{kv} - h\varepsilon \tag{12}$$

$$\frac{\partial h\varepsilon}{\partial t} + \frac{\partial U_j h\varepsilon}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\left(\nu + \frac{\nu_t}{\sigma_\varepsilon}\right) h\frac{\partial \varepsilon}{\partial x_j}\right) + hc_{\varepsilon 1}\frac{\varepsilon}{k}P_k + hP_{\varepsilon v} - hc_{\varepsilon 2}\frac{\varepsilon^2}{k}$$

$$\nu_t = c_\mu \frac{k^2}{\varepsilon} \qquad P_k = 2\nu_t(S_{uu}^2 + S_{vv}^2 + 2S_{uv}^2)$$

$$P_{kv} = c_k \frac{u_f^3}{h} \qquad c_k = \frac{1}{c_f^{1/2}}$$

$$P_{\varepsilon v} = c_\varepsilon \frac{u_f^4}{h^2} \qquad c_\varepsilon = 3.6\frac{c_{\varepsilon 2} c_\mu^{1/2}}{c_f^{3/4}}$$

$$c_\mu = 0.09 \qquad c_{\varepsilon 1} = 1.44 \quad c_{\varepsilon 2} = 1.92 \quad \sigma_k = 1.0 \quad \sigma_\varepsilon = 1.31$$

where c_f is the bed friction coefficient ($u_f^2 = c_f|\mathrm{U}|^2$). The five constants of the model $c_\mu, \sigma_k, \sigma_\varepsilon, c_{1\varepsilon}, c_{2\varepsilon}$ are assumed to have the same values as in the original $k-\varepsilon$ model. The term P_k accounts for the production of turbulent energy due to horizontal velocity gradients, and it has the same expression as in the 2D model. The source terms P_{kv} and $P_{\varepsilon v}$ are responsible for modelling the 3D turbulence generated by bed friction.

In uniform channel flow conditions turbulence is in an equilibrium state. In that case the production of turbulent kinetic energy is due to bed friction, and the $k-\varepsilon$ equations reduce to:

$$P_{kv} = \varepsilon \qquad P_{\varepsilon v} = c_{\varepsilon 2}\frac{\varepsilon^2}{k} \tag{13}$$

Using the expressions for P_{kv} and $P_{\varepsilon v}$ given by Equation 12, the values for the turbulent energy and dissipation predicted by the model in uniform channel flow are given by:

$$k_u = \frac{c_{\varepsilon 2}}{c_\varepsilon} c_k^2 u_f^2 = \frac{|\mathrm{U}|^{1/2} u_f^{3/2}}{1.08} \qquad \varepsilon_u = c_k \frac{u_f^3}{h} = \frac{|\mathrm{U}| u_f^2}{h} \tag{14}$$

and the eddy viscosity is equal to:

$$\nu_{t,u} = c_\mu c_k^3 \frac{c_{\varepsilon 2}^2}{c_\varepsilon^2} u_f h = 0.08 u_f h \tag{15}$$

which is a very similar value to that one given by the depth averaged parabolic and mixing length models (Equation 5). On the other hand, when the horizontal shear is much larger than the vertical shear, the vertical production terms are negligible compared to the horizontal shear production ($P_{kv} \approx 0, P_{\varepsilon v} \approx 0$), and the model reduces to the standard 2D $k - \varepsilon$ model.

A slight modification of the coefficient c_ε has been used in recent works by Minh-Duc et al. (Minh-Duc et al. 2004). The expression used is given by:

$$c_\varepsilon = \frac{1}{(e_* \sigma_t)^{1/2}} \frac{c_{\varepsilon 2} c_\mu^{1/2}}{c_f^{3/4}} \tag{16}$$

where σ_t is a Prandtl/Schmidt number relating the eddy viscosity and diffusivity for the transport of scalars ($\sigma_t = 0.7$ in (Minh-Duc et al. 2004)), and e_* is an adjustable empirical parameter. The constant 3.6 used in the definition of c_ε in Equation 12, corresponds to a value of $e_* = 0.11$ in Equation 16. In the results presented in this paper the original definition of c_ε was used, since Equation 16 would need a previous calibration for each practical application.

3.4 *The modified constants of Booij*

The production terms P_{kv} and $P_{\varepsilon v}$ in Equation 12 depend on two coefficients, c_k and c_ε. The value of these coefficients determine the turbulence level in uniform channel flow conditions (Equation 14). Booij (Booij 1989) proposed the following modification in the value of c_k and c_ε:

$$c_k^B = \frac{1}{10} c_k \qquad c_\varepsilon^B = \frac{1}{44} c_\varepsilon \tag{17}$$

where the superindex B refers to Booij. The effect of diminishing the coefficients c_k and c_ε in such a way is that the production terms of both k and ε due to bed friction (P_{kv} and $P_{\varepsilon v}$) are strongly reduced. Using the modified coefficients of Booij to evaluate the turbulent kinetic energy and dissipation in uniform channel flow (Equation 14) gives:

$$k_u^B = 0.44 k_u \qquad \varepsilon_u^B = 0.10 \varepsilon_u \qquad \nu_{t,u}^B = 1.94 \nu_{t,u} \tag{18}$$

where the values k_u, ε_u, $\nu_{t,u}$ are those given by the original constants used by Rastogi and Rodi (Rastogi and Rodi 1978). Hence, in uniform channel flow conditions, the modification proposed by Booij reduces the turbulent kinetic energy by a factor of 0.44, but it almost duplicates the eddy viscosity. On the other hand, in flows dominated by horizontal shear the differences with the Rastogi and Rodi model smear, because the term P_{kv} becomes much smaller than P_k in Equation 12.

3.5 *The $k - \varepsilon$ model of Babarutsi and Chu*

Babarutsi and Chu (Babarutsi and Chu 1991) proposed a two-length-scale depth averaged $k - \varepsilon$ model. The main difference with Rastogi's model is the way in which the effects of the 3D turbulence produced by bed friction are introduced. Babarutsi and Chu propose to split the total eddy viscosity into a 3D part and a 2D part as:

$$\nu_t = \nu_t^{2D} + \nu_t^{3D} \tag{19}$$

The 3D eddy viscosity accounts for the small scale bed generated turbulence. The equation for ν_t^{3D} used by Babarutsi and Chu is given by:

$$\nu_t^{3D} = 0.08 u_f h \tag{20}$$

which is the same value given by the model of Rastogi and Rodi for uniform channel flow (Equation 15). The 2D part is computed from k' and ε', which account for the large-scale 2D turbulence generated by horizontal shear, as:

$$\nu_t^{2D} = c_\mu \frac{k'^2}{\varepsilon'} \tag{21}$$

It should be remarked that the turbulent kinetic energy k' does not include the 3D bed generated turbulence. Hence, the transport equations used to compute k' and ε' are similar to those used in Rastogi's model (Equation 12), but the vertical production terms are zero ($P_k v = P_\varepsilon v = 0$) and a new term F' is introduced in order to account for the transfer of energy between the large 2D turbulent scales and the small 3D scales. The equations for k' and ε' are given by:

$$\begin{aligned}
\frac{\partial hk'}{\partial t} + \frac{\partial U_j hk'}{\partial x_j} &= \frac{\partial}{\partial x_j}\left(\left(\nu + \frac{\nu_t}{\sigma_k}\right) h \frac{\partial k'}{\partial x_j}\right) + h(P_{k'} - F') - h\varepsilon' \\
\frac{\partial h\varepsilon'}{\partial t} + \frac{\partial U_j h\varepsilon'}{\partial x_j} &= \frac{\partial}{\partial x_j}\left(\left(\nu + \frac{\nu_t}{\sigma_\varepsilon}\right) h \frac{\partial \varepsilon'}{\partial x_j}\right) + hc_{\varepsilon 1}\frac{\varepsilon'}{k'}(P_{k'} - (1 - c_{\varepsilon 3})F') - hc_{\varepsilon 2}\frac{\varepsilon'^2}{k'}
\end{aligned} \tag{22}$$

A value of $c_{\varepsilon 3} = 0.8$ is proposed in (Babarutsi and Chu 1991). Note that in Equation 22 the diffusion term is modelled with the total eddy viscosity ν_t. On the other hand, the production term $P_{k'}$ is computed with Equation 12, but using only the 2D eddy viscosity ν_t^{2D}. All the constants of the model are the same as in Rastogi's model. The sink term F' is associated with the negative work done by the large scale turbulent fluctuations against the bed friction, i.e. the bed friction takes out energy from the large 2D turbulent structures. It is computed as:

$$F' = \frac{c_f\left[\overline{u'^2}(2U^2 + V^2) + 2\overline{u'v'}UV + \overline{v'^2}(U^2 + 2V^2)\right]}{h\sqrt{U^2 + V^2}} \tag{23}$$

where the Reynolds stresses $\overline{u'^2}$, $\overline{u'v'}$, $\overline{v'^2}$ are computed from the Boussinesq approximation using the large scale eddy viscosity ν_t^{2D}.

In uniform channel flow, the production of large scale turbulence due to horizontal shear is zero ($P_{k'} = 0$). The large scale turbulent kinetic energy equation reduces to $0 = -F' - \varepsilon'$. Therefore, all the large scale turbulent variables are zero, i.e:

$$k'_u = 0 \qquad \varepsilon'_u = 0 \qquad \nu_{t,u}^{2D} = 0 \qquad \nu_{t,u} = \nu_t^{3D} = 0.8 u_f h \tag{24}$$

where the subindex u refers to uniform channel flow conditions. In this case all the turbulence is 3D, and generated by bed friction. The total eddy viscosity is similar to that one given by the model of Rastogi. On the other hand, when the bed friction velocity u_f tends to zero, the flow is dominated by horizontal shear. In such a case the term F' and the 3D eddy viscosity may be neglected ($F' \approx 0$, $\nu_t^{3D} \approx 0$), all the turbulence is 2D, and generated by horizontal shear. The equations for k' and ε' reduce to the standard 2D $k - \varepsilon$ model.

This model has been used by Babarutsi et al. in (Babarutsi et al. 1996), with the constant $c_{\varepsilon 3} = 1$, to compute shallow recirculating flows dominated by friction. It has also been used by Babarutsi and Chu to model shallow mixing layers (Babarutsi and Chu 1998), where they obtained the same results with $c_{\varepsilon 3} = 1$ and with $c_{\varepsilon 3} = 0.8$.

3.6 *Differences between the 3 versions of the $k-\varepsilon$ model*

The main differences between the previous depth averaged $k-\varepsilon$ models appear in friction dominated flows. The modification of Booij reduces the bed generated turbulence, and as a consequence, the large scale turbulence as well as the total turbulent diffusion of the model are increased. We have not found concluding results showing that the modification proposed by Booij improves the results of the original model of Rastogi in general flow conditions. Furthermore, the work by Minh-Duc et al. (Minh-Duc et al. 2004) seems to show that the optimum value of the constant c_ε might be problem dependent.

The version of Babarutsi and Chu is more interesting from a conceptual point of view, since it distinguishes between two different turbulent length scales. The main differences with Rastogi's model appear in shear flows dominated by friction. Both models converge to the standard 2D $k-\varepsilon$ model for zero bed friction. On the other hand, in uniform channel flow, where all the turbulence is generated by bed friction, both models give the same eddy viscosity value (Equations 15 and 24). It is in horizontal shear layers dominated by bed friction where both models show some differences. Altai et al. (Altai et al. 1999) estimate that, in order to produce the modelled effects introduced by the term F', the water depth should be approximately 100 times smaller than the horizontal turbulent length scale, i.e. the effects of F' are only significant in flows which are strongly dominated by friction and at the same time have large 2D turbulent structures. Babarutsi et al. (Babarutsi et al. 1996) compared the results given by the models of Babarutsi and Rastogi in a recirculating shallow water flow dominated by friction. Despite the different eddy viscosity predicted by the models, no significant differences in the mean flow field, neither in the recirculation length, were found, and therefore, no conclusion was obtained regarding which model performs better. Both models were compared again by Babarutsi and Chu (Babarutsi and Chu 1998) in order to model transverse mixing layers in shallow flows dominated by friction. In that case they found that the results given by the model of Babarutsi were in better agreement with the experimental data.

In the applications presented in this work no differences are expected to be found in the mean flow field obtained with the different versions of the model. This is because the flow in the vertical slot fishway, as well as in the open channel 90° bend, is not dominated by bed friction, while in the Crouch estuary, although the turbulence is mainly generated by bed friction, there are not strong horizontal shear stresses.

3.7 *Limiter to the production of turbulent kinetic energy*

It is well known that Boussinesq assumption does not work well when the turbulence is strongly non-isotropic, as it happens in swirling flows, in regions with strong adverse pressure gradients, or near stagnation points. The main cause of these problems is that the eddy viscosity is assumed to be isotropic. In depth averaged models, the non-isotropy between the vertical and horizontal directions is somehow introduced in the equations by the different ways of accounting for the production of 2D horizontal turbulence and 3D turbulence generated by bed friction. However, it still remains the problem with non-isotropy in the horizontal plane.

It is well known that in stagnation points the production of turbulent kinetic energy is overpredicted by the $k-\varepsilon$ model, which gives much larger values of turbulent kinetic energy than the measured ones (Menter 1993). For this reason, a limiter in the production term is introduced in almost all the eddy viscosity models in order to avoid too large values of the turbulent energy in stagnation regions. Menter (Menter 1993) limits the ratio between the production and dissipation of turbulence as:

$$P = \min(P_k, c_l \varepsilon) \tag{25}$$

where c_l is a constant with a value between 10 and 20. This limiter is a rough estimation based on empirical observations which does not imply a good modelling of the stagnation region, but it can improve the numerical results, avoiding an extremely large turbulence level which may propagate and affect the solution in the whole numerical domain.

In this work the turbulent production limiter has been applied independently to both, the horizontal shear production and the bed friction production as:

$$P_k = \min(P_k, c_l\varepsilon)$$
$$P_{kv} = \min(P_{kv}, c_l\varepsilon) \tag{26}$$

with a value of $c_l = 10$. The limit imposed in the bed friction production P_{kv} is specially suited for problems with wet-dry fronts. The term P_{kv} can take quite large values when the water depth is very small. In that case, the second limiter in Equation 26 avoids instabilities in the solution, and allows the use of smaller values of the wet-dry tolerance parameter (see section 4).

3.8 *Realizability in the eddy viscosity models*

Another important drawback to Boussinesq assumption, which is somehow related to the excessive turbulence prediction in non-isotropic turbulence, is that it can produce negative values of the normal Reynolds stresses, specially in stagnation regions. This was pointed out by Durbin in (Durbin 1996). In order to avoid negative normal stresses, a constraint should be used in the eddy viscosity value. This procedure also helps to avoid the excessive turbulent production near stagnation regions.

In 2D flow, the constraint over the eddy viscosity which assures that all the normal turbulent stresses remain positive is given by:

$$\nu_t < \frac{k}{3}\left(\frac{2}{S_{ij}S_{ij}}\right)^{1/2} \tag{27}$$

The limiter given by Equation 27 should always be used in all turbulence models based on Boussinesq assumption.

3.9 *A depth averaged algebraic stress turbulence model*

One of the main weakness of eddy viscosity models is Boussinesq approximation, which assumes an isotropic eddy viscosity when computing the Reynolds stresses. This assumption causes a reduction in the accuracy of the model when the turbulence field is strongly non-isotropic, as for example in swirling and recirculating flows, in stagnation regions, or in flows with strong adverse pressure gradients. An attempt to improve the turbulence modelling under these conditions, without a significant increase in the computational cost, can be done using an Algebraic Stress Model.

In (Cea et al. 2007) an explicit depth averaged Algebraic Stress Model (DASM) which accounts for the turbulence production due to bed friction was proposed, showing a numerical stability and computational cost similar to that of the $k-\varepsilon$ model. The DASM proposed in (Cea et al. 2007) can be written in matrix form as:

$$\begin{pmatrix}\overline{u'^2}\\ \overline{v'^2}\\ \overline{u'v'}\end{pmatrix} = \frac{1-c_2}{c_{11}}\frac{k}{\varepsilon}\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}\begin{pmatrix}\overline{u'^2}\\ \overline{v'^2}\\ \overline{u'v'}\end{pmatrix} + \begin{pmatrix}b_1\\ b_2\\ b_3\end{pmatrix} \tag{28}$$

with the coefficients a_{ij} and b_i equal to:

$$\begin{aligned}
a_{11} &= -\frac{4}{3}\frac{\partial U}{\partial x} & a_{12} &= \frac{2}{3}\frac{\partial V}{\partial y} & a_{13} &= -\frac{4}{3}\frac{\partial U}{\partial y} + \frac{2}{3}\frac{\partial V}{\partial x}\\
a_{21} &= \frac{2}{3}\frac{\partial U}{\partial x} & a_{22} &= -\frac{4}{3}\frac{\partial V}{\partial y} & a_{23} &= -\frac{4}{3}\frac{\partial V}{\partial x} + \frac{2}{3}\frac{\partial U}{\partial y}\\
a_{31} &= -\frac{\partial V}{\partial x} & a_{32} &= -\frac{\partial U}{\partial y} & a_{33} &= -\left(\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y}\right)
\end{aligned} \tag{29}$$

$$
\begin{aligned}
b_1 &= \frac{2}{3}k + \frac{k}{\varepsilon}\frac{(1-c_2)}{c_{11}}\left(\frac{2}{3}P_{uu,V} - \frac{1}{3}P_{vv,V}\right) \\
b_2 &= \frac{2}{3}k + \frac{k}{\varepsilon}\frac{(1-c_2)}{c_{11}}\left(\frac{2}{3}P_{vv,V} - \frac{1}{3}P_{uu,V}\right) \\
b_3 &= \frac{k}{\varepsilon}\frac{(1-c_2)}{c_{11}}P_{uv,V}
\end{aligned}
\tag{30}
$$

The following constants are used in the model:

$$
c_1 = 1.8 \qquad c_2 = 0.6 \qquad c_{11} = c_1 + \frac{P_k}{\varepsilon} - 1 \tag{31}
$$

The turbulent kinetic energy production P_k can be computed using the eddy viscosity assumption as:

$$
P_k^{vt} = 2\nu_t\left[\left(\frac{\partial U}{\partial x}\right)^2 + \frac{1}{2}\left(\frac{\partial U}{\partial y} + \frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2\right] + P_{kv} \tag{32}
$$

with the eddy viscosity ν_t computed from the $k-\varepsilon$ model. In this way, the system of equations 28 can be solved explicitly, improving in this way the convergence of the numerical scheme. The approximation given by Equation 32 is not expected to influence importantly the results, since it is just used in order to compute the constant c_{11} and therefore, its influence is smoothed. Therefore, Equation 32 will be used for computing the turbulent kinetic energy production term in the coefficient $\boldsymbol{c_{11}}$ (Equation 31)

The only thing which remains to specify in the present model is the evaluation of the vertical production terms due to bed friction ($P_{uu,V}, P_{vv,V}, P_{uv,V}$). Those terms are evaluated as (Cea et al. 2007):

$$
P_{uu,V} = 2\frac{u_f^2 u_x^2}{h|\mathbf{U}|} \qquad P_{vv,V} = 2\frac{u_f^2 u_y^2}{h|\mathbf{U}|} \qquad P_{uv,V} = -2\frac{u_f^2 u_x u_y}{h|\mathbf{U}|} \tag{33}
$$

It is necessary to use a realizability condition with the DASM, since there is nothing in the equations of the model which assures that the normal Reynolds stresses will remain positive. Hence, the following restrictions should be imposed in the model:

$$
\overline{u'^2} \geq 0 \qquad \overline{v'^2} \geq 0 \qquad \overline{u'^2} + \overline{v'^2} \leq 2k \tag{34}
$$

The third condition in Equation 34 assures that the vertical normal stress $\boldsymbol{\overline{w'^2}}$ remains positive, and avoids excessively large values of the horizontal normal stresses. If any of the normal stresses is negative its value is set to zero. At the same time, in order to keep constant the total turbulent kinetic energy, it is necessary to subtract from the other normal stresses the same amount of energy. This can be easily implemented in the following way:

$$
\begin{aligned}
\overline{u'^2} &= \max\left\{0, \min\left(\overline{u'^2}, \frac{2k}{2k-\overline{v'^2}}\overline{u'^2}, \frac{2k}{2k-\overline{w'^2}}\overline{u'^2}\right)\right\} \\
\overline{v'^2} &= \max\left\{0, \min\left(\overline{v'^2}, \frac{2k}{2k-\overline{u'^2}}\overline{v'^2}, \frac{2k}{2k-\overline{w'^2}}\overline{v'^2}\right)\right\}
\end{aligned}
\tag{35}
$$

4 FINITE VOLUME DISCRETISATION OF THE TURBULENT SHALLOW WATER EQUATIONS

This section presents an unstructured finite volume model for solving the 2D-SWE and turbulence models presented in sections 3 and 3.9. A comprehensive description of finite volume methods for fluid dynamics can be found in (LeVeque 2002; Toro 1999).

The discretisation of the spatial domain will be made with edge-type finite volumes, which were introduced by Bermúdez et al. in (Bermúdez et al. 1998). These kind of control volumes are generated from a previous triangulation of the numerical domain, as sketched in Figure 1. The nodes (N_i and N_j) are placed at the midpoint of the edges of the triangulation, and the cells C_i are built from the vertex and the barycentre of the triangles. The control volumes built in such a way have 4 faces except in the boundaries, where they have **3** faces. This approach permits an easy definition of the normal vector at the boundary faces, avoiding the indetermination which may appear in certain boundary nodes when using vertex-type control volumes in complex geometries (Dervieux and Desideri 1992).

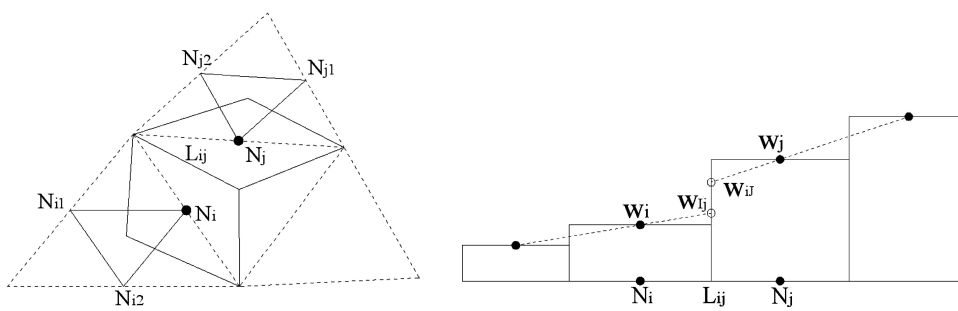
The 2D-SWE can be written in vectorial form as:

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\partial \mathbf{F_x}}{\partial x} + \frac{\partial \mathbf{F_y}}{\partial y} = \mathbf{S} + \mathbf{G} + \mathbf{D} \tag{36}$$

$$\mathbf{w} = \begin{pmatrix} h \\ q_x \\ q_y \end{pmatrix} \qquad \mathbf{F_x} = \begin{pmatrix} q_x \\ \frac{q_x^2}{h} + \frac{gh^2}{2} \\ \frac{q_x q_y}{h} \end{pmatrix} \qquad \mathbf{F_y} = \begin{pmatrix} q_y \\ \frac{q_x q_y}{h} \\ \frac{q_y^2}{h} + \frac{gh^2}{2} \end{pmatrix}$$

$$\mathbf{S} = \begin{pmatrix} 0 \\ -gh\frac{\partial z_b}{\partial x} \\ -gh\frac{\partial z_b}{\partial y} \end{pmatrix} \qquad \mathbf{G} = \begin{pmatrix} 0 \\ -\frac{\tau_{bx}}{\rho} \\ -\frac{\tau_{by}}{\rho} \end{pmatrix} \qquad \mathbf{D} = \begin{pmatrix} 0 \\ \frac{\partial}{\partial x_j}\left(\nu_e h \frac{\partial U_x}{\partial x_j}\right) \\ \frac{\partial}{\partial x_j}\left(\nu_e h \frac{\partial U_y}{\partial x_j}\right) \end{pmatrix}$$

where ν_e is the effective viscosity ($\nu_e = \nu + \nu_t$). The vectors $\mathbf{F_x}$ and $\mathbf{F_y}$ account for the convective flux. The vector $\mathbf{S}$ accounts for the bed slope, $\mathbf{G}$ for the bed friction, and $\mathbf{D}$ for the turbulent and



(a) Triangles used to compute the gradients $\nabla \mathbf{w}_i$ and $\nabla \mathbf{w}_j$

(b) Linear reconstruction from nodes to faces.1D view.

Figure 1. Second order scheme. Reconstruction of the conservative variables from the cell nodes to the cell faces.

viscous diffusion. An explicit discretisation of Equation (36) is given by:

$$\frac{\mathbf{w}^{n+1}-\mathbf{w}^n}{\Delta t}+\frac{\partial \mathbf{F_x}}{\partial x}(\mathbf{w}^n)+\frac{\partial \mathbf{F_y}}{\partial y}(\mathbf{w}^n)=\mathbf{S}^n+\mathbf{G}^n+\mathbf{D}^n \tag{37}$$

where the super-index n refers to time t^n. The extension of the scheme to second order in time can be easily implemented for example by a two-step method (Cea 2005; Brufau 2000).

Integration of Equation (37) over a control volume $\boldsymbol{C_i}$ and yields:

$$\frac{\mathbf{w}_i^{n+1}-\mathbf{w}_i^n}{\Delta t}A_i+\sum_{j\in K_i}\int_{L_{ij}}(\mathbf{F_x}\tilde{n}_x+\mathbf{F_y}\tilde{n}_y)\,dL=\int_{C_i}(\mathbf{S^n}+\mathbf{G^n}+\mathbf{D^n})\,dA \tag{38}$$

where $\mathbf{w}_i^n$ is the mean value of $\mathbf{w}$ in the cell C_i at time t^n, L_{ij} is the common face to the volumes C_i and C_j, $\tilde{\mathbf{n}}=(\tilde{n}_x,\tilde{n}_y)$ is the unit normal vector to the cell face, and K_i accounts for all the cells C_j which share any face with the cell C_i.

4.1 *Convective flux*

The convective flux in Equation (38) is discretised with a Roe upwind scheme (Roe 1986). Harten's regularisation (Harten et al. 1983) is applied when necessary. In order to obtain second order accuracy in space, the left and right states of the Riemann problem are obtained after a linear reconstruction of the variables $\mathbf{w}$ from the nodes to the cell faces, which can be computed as:

$$\mathbf{w}_{Ij}=\mathbf{w}_i+\frac{1}{2}\Delta_i^* \qquad \mathbf{w}_{iJ}=\mathbf{w}_j+\frac{1}{2}\Delta_j^* \tag{39}$$

where Δ_i^* Δ_j^* are the limited slopes (Toro 2001) at the nodes $\mathbf{N}_i$ and $\mathbf{N}_j$, $\mathbf{w}_{Ij}$ is the extrapolated value of $\mathbf{w}_i$ to the cell face $\mathbf{L}_{ij}$, and $\mathbf{w}_{ij}$ is the extrapolated value of $\mathbf{w}_j$ to the cell face $\mathbf{L}_{ij}$ (Figure 1(b)). The limited slopes can be computed as:

$$\boldsymbol{\Delta}_i^*=\begin{cases}\max\left[0,\min\left(\nabla\mathbf{w}_i\mathbf{r}_{ij},\Delta_{ij}\right)\right] & \text{if} \quad \Delta_{ij}>0\\ \min\left[0,\max\left(\nabla\mathbf{w}_i\mathbf{r}_{ij},\Delta_{ij}\right)\right] & \text{if} \quad \Delta_{ij}<0\end{cases} \quad \text{with} \quad \Delta_{ij}=\mathbf{w}_j-\mathbf{w}_i \tag{40}$$

with an analogous expression for $\boldsymbol{\Delta}_{\boldsymbol{j}}^*$. In Equation (40) $\mathbf{r_{ij}}$ is the distance vector between the two nodes N_i and N_j. The gradients $\nabla\mathbf{w}_i$ and $\nabla\mathbf{w}_j$ are computed from the values of $\mathbf{w}$ at the nodes (N_i,N_{i_1},N_{i_2}) and (N_j,N_{j_1},N_{j_2}) respectively (Vázquez-Cendón 1996) (Figure 1(a)). The limited slopes computed from Equation (40) reproduce the Minmod limiter (Toro 2001).

The boundary integral of the convective flux in Equation (38) is approximated by the numerical flux $\boldsymbol{\phi_{ij}}$ as:

$$\int_{L_{ij}}(\mathbf{F}_x\tilde{n}_x+\mathbf{F}_y\tilde{n}_y)\,dL\approx\phi_{ij}(\mathbf{w}_L,\mathbf{w}_R,\mathbf{n}_{ij}) \tag{41}$$

with:

$$\boldsymbol{\phi_{ij}}(\mathbf{w}_L,\mathbf{w}_R,\mathbf{n}_{ij})=\frac{\mathbf{Z}(\mathbf{w}_L,\mathbf{n}_{ij})+\mathbf{Z}(\mathbf{w}_R,\mathbf{n}_{ij})}{2}-\frac{1}{2}|\mathbf{Q}|(\mathbf{w}_L,\mathbf{w}_R,\mathbf{n}_{ij})(\mathbf{w}_R-\mathbf{w}_L) \tag{42}$$

$$\mathbf{Z}=\mathbf{F_x}n_x+\mathbf{F_y}n_y \qquad |\mathbf{Q}|=\mathbf{X}|\mathbf{D}|\mathbf{X}^{-1}$$

$$\mathbf{X}=\begin{pmatrix}0 & 1 & 1\\ -\tilde{n}_y & \tilde{U}_x+\tilde{c}\tilde{n}_x & \tilde{U}_x-\tilde{c}\tilde{n}_x\\ \tilde{n}_x & \tilde{U}_y+\tilde{c}\tilde{n}_y & \tilde{U}_y-\tilde{c}\tilde{n}_y\end{pmatrix} \qquad |\mathbf{D}|=\begin{pmatrix}|\tilde{\lambda}_1| & 0 & 0\\ 0 & |\tilde{\lambda}_2| & 0\\ 0 & 0 & |\tilde{\lambda}_3|\end{pmatrix}$$

$$\tilde{\lambda}_1 = n_x\tilde{U}_x + n_y\tilde{U}_y \qquad \tilde{\lambda}_2 = \tilde{\lambda}_1 + \tilde{c}\sqrt{n_x^2+n_y^2} \qquad \tilde{\lambda}_3 = \tilde{\lambda}_1 + \tilde{c}\sqrt{n_x^2+n_y^2} \tag{43}$$

where $\mathbf{w}_L = \mathbf{w}_{Ij}$ and $\mathbf{w}_R = \mathbf{w}_{Ij}$. In Equation (41–42) $\mathbf{n}_{ij}$ is a normal vector to the cell face L_{ij} with the same length as the cell face.

4.2 *Bed slope source term*

When the upwind scheme given by Equations (39–43) is used with a centred discretisation of the bed slope source term $\mathbf{S}$, spurious oscillations are generated under hydrostatic flow conditions. In order to avoid this unphysical oscillations, the bed slope source term must be discretised with an upwind scheme (Fernández-Nieto 2003; García-Navarro and Vázquez Cendón 2000; Vázquez-Cendón 1999; Bermúdez et al. 1998). A suitable upwind discretisation of $\mathbf{S}$ which is free of spurious oscillations when used with the first order scheme of Roe is given by (Fernández-Nieto 2003; Bermúdez et al. 1998):

$$\mathbf{S}_i = \mathbf{S}_i^C - \frac{1}{A_i}\sum_{j\in K_i}\frac{d_{\perp,ij}}{2}|\mathbf{n}_{ij}||\mathbf{Q}|_{ij}\mathbf{Q}_{ij}^{-1}\tilde{\mathbf{S}}_{ij} \tag{44}$$

$$\mathbf{S}_i^C - \frac{1}{A_i}\sum_{j\in K_i}\frac{d_{\perp,ij}}{2}|\mathbf{n}_{ij}|\tilde{\mathbf{S}}_{ij} \qquad \tilde{\mathbf{S}}_{ij} = -g\frac{h_i+h_j}{2}\frac{z_{b,j}-z_{b,i}}{d_{\perp,ij}}\begin{pmatrix}0\\ \tilde{n}_{x,ij}\\ \tilde{n}_{y,ij}\end{pmatrix}$$

where $\mathbf{S}_i$ is the upwind discretisation of the bed slope source term, $\mathbf{S}_i^C$ is a centred discretisation of the source term in the cell C_i, $\tilde{\mathbf{S}}_{ij}$ is a centred approximation of the source term at the cell face L_{ij}, A_i is the area of the cell C_i, and $d_{\perp,ij} = \mathbf{r}_{ij}\tilde{\mathbf{n}}_{ij}$ is the projection of the distance between the nodes N_i and $N_j(d_{ij})$ on to the unit normal vector $\tilde{\mathbf{n}}_{ij}$.

In order to prove the convenience of upwinding the bed slope source term, the first order scheme of Roe will be applied to hydrostatic flow conditions. In the hydrostatic assumption, the discrete stationary equations reduce to:

$$\sum_{j\in K_i}\phi_{ij} = \mathbf{S}_i\boldsymbol{A}_i \tag{45}$$

The numerical flux given by the scheme for the cell $\boldsymbol{C}_i$ is given by:

$$\boldsymbol{\phi_{ij}} = \frac{\mathbf{Z}_i+\mathbf{Z}_j}{2} - \frac{1}{2}\mathbf{X}_{ij}|\mathbf{D}|_{ij}\mathbf{X}_{ij}^{-1}(\mathbf{w}_j-\mathbf{w}_i) \tag{46}$$

where the subindices ij indicate Roe's mean state, i.e. $\mathbf{X}_{ij} = \mathbf{X}(\tilde{\mathbf{w}}_{ij})$ and $|\mathbf{D}|_{ij} = |\mathbf{D}|(\tilde{\mathbf{w}}_{ij})$. Considering that in hydrostatic conditions $q_x = q_y = 0$ in all cells, the mean state of Roe is reduced to:

$$\tilde{h} = \sqrt{h_i h_j} \qquad \tilde{c} = \sqrt{g\frac{h_i+h_j}{2}} \qquad \tilde{q}_x = 0 \qquad \tilde{q}_y = 0 \tag{47}$$

The values of the different vectors and matrices in Equation 46 are given by:

$$\mathbf{X}_{ij} = \begin{pmatrix}0 & 1 & 1\\ -\tilde{n}_{y,ij} & \tilde{c}_{ij}\tilde{n}_{x,ij} & -\tilde{c}_{ij}\tilde{n}_{x,ij}\\ \tilde{n}_{x,ij} & \tilde{c}_{ij}\tilde{n}_{y,ij} & -\tilde{c}_{ij}\tilde{n}_{y,ij}\end{pmatrix} \quad \mathbf{X}_{ij}^{-1} = \frac{1}{2\tilde{c}_{ij}}\begin{pmatrix}0 & -2\tilde{c}_{ij}\tilde{n}_{y,ij} & 2\tilde{c}_{ij}\tilde{n}_{x,ij}\\ \tilde{c}_{ij} & \tilde{n}_{x,ij} & \tilde{n}_{y,ij}\\ \tilde{c}_{ij} & -\tilde{n}_{x,ij} & -\tilde{n}_{y,ij}\end{pmatrix} \tag{48}$$

$$\mathbf{D}_{ij} = \tilde{c}_{ij}|\mathbf{n}_{ij}|\begin{pmatrix}0&0&0\\0&1&0\\0&0&-1\end{pmatrix} \quad |\mathbf{D}|_{ij} = \tilde{c}_{ij}|\mathbf{n}_{ij}|\begin{pmatrix}0&0&0\\0&1&0\\0&0&1\end{pmatrix} \quad |\mathbf{D}|_{ij}^{-1} = \frac{1}{\tilde{c}_{ij}|\mathbf{n}_{ij}|}\begin{pmatrix}0&0&0\\0&1&0\\0&0&-1\end{pmatrix}$$

$$\mathbf{Z}_i = |\mathbf{n}_{ij}| \begin{pmatrix} 0 \\ g\dfrac{h_i^2}{2}\tilde{n}_{x,ij} \\ g\dfrac{h_i^2}{2}\tilde{n}_{y,ij} \end{pmatrix} \qquad w_i = \begin{pmatrix} h_i \\ 0 \\ 0 \end{pmatrix}$$

Introducing expressions 48 into Equation 46 gives, after some mathematical manipulation, the total flux for the cell C_i as:

$$\sum_{j\in K_i} \phi_{ij} = \sum_{j\in K_i} \frac{|\mathbf{n}_{ij}|}{2} g \frac{h_i^2 + h_j^2}{2} \begin{pmatrix} 0 \\ \tilde{n}_{x,ij} \\ \tilde{n}_{y,ij} \end{pmatrix} - \sum_{j\in K_i} \frac{|\mathbf{n}_{ij}|}{2} \tilde{c}_{ij}(h_j - h_i) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{49}$$

where the first addend accounts for the centred contribution and the second addend for the upwind contribution.

The discretisation of the bed slope term is obtained introducing expression 48 into Equation 44, which yields:

$$\mathbf{S}_i A_i = \sum_{j\in K_i} -\frac{|\mathbf{n}_{ij}|}{2} g \frac{h_i + h_j}{2}(z_{b,j} - z_{b,i}) \begin{pmatrix} 0 \\ \tilde{n}_{x,ij} \\ \tilde{n}_{y,ij} \end{pmatrix} + \sum_{j\in K_i} \frac{|\mathbf{n}_{ij}|}{2} \tilde{c}_{ij}(z_{b,j} - z_{b,i}) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{50}$$

where the first addend accounts for the centred contribution and the second addend for the upwind contribution.

Considering that in hydrostatic flow the water surface elevation is constant ($z_{b,i} + h_i = z_{b,i} + h_j$), it is straightforward to show that the upwind contribution in the convective flux (second addend in Equation 49) and in the source term (second addend in Equation 50) are equal. The balance of the centred contributions follows directly considering that for any closed volume the following property (Equation 51) applies for both the x and the y component of the normal vector:

$$\sum_{j\in K_i} h_i^2 \tilde{n}_{x,ij} |\mathbf{n}_{ij}| = h_i^2 \sum_{j\in K_i} n_{x,ij} = 0 \tag{51}$$

The previous argument is valid for the first order upwind scheme. If the second order extension of Roe scheme is used, the exact balance between the numerical flux and the bed slope source term is broken, even if a second order extrapolation is used to evaluate the bed elevation at the cell faces. This is because the extrapolated value of h_i is different at each cell face, and therefore,

$$\sum_{j\in K_i} h_{Ij}^2 \tilde{n}_{x,ij} |\mathbf{n}_{ij}| \neq 0 \tag{52}$$

where h_{Ij} is the extrapolated value of the flux from the node N_i to the cell face L_{ij}. For this reason, the centred contributions in the flux term discretisation (Equation 49) and in the source term discretisation (Equation 50) do not balance any more. There is actually a deficit in the source term given by:

$$\mathbf{R}_i = \sum_{j\in K_i} -g \frac{|\mathbf{n}_{ij}|}{2} h_{Ij}^2 \begin{pmatrix} 0 \\ \tilde{n}_{x,ij} \\ \tilde{n}_{y,ij} \end{pmatrix} \tag{53}$$

and the following relation applies:

$$\sum_{j\in K_i} \phi_{ij} = \mathbf{S}_i A_i + \mathbf{R}_i \tag{54}$$

We have found that a simple and efficient solution to this problem is to use a first order approximation for the water depth ($h_{Ij} = h_i$), and a second order approximation for the unit discharges. In this case the deficit $\mathbf{R}_i$ is zero:

$$\mathbf{R}_i = \sum_{j \in K_i} -g \frac{|\mathbf{n}_{ij}|}{2} h_{Ij}^2 \begin{pmatrix} 0 \\ \tilde{n}_{x,ij} \\ \tilde{n}_{y,ij} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{55}$$

and the exact balance between flux and source is recovered. This approach leads to an hybrid scheme which is able to compute the hydrostatic flow solution exactly, and at the same time reduces in a very significant way the numerical diffusion.

Hubbard and García-Navarro (Hubbard and García-Navarro 2000) proposed a high order correction of the source term in order to obtain a fully second order scheme which gives an exact balance between convective flux and bed slope in the hydrostatic case. Following the ideas presented in (Hubbard and García-Navarro 2000), when using a second order scheme Equation 53 can be replaced by the more general relation:

$$\mathbf{R}_i = \sum_{j \in K_i} [\mathbf{F}_{x,Ij} n_{x,ij} + \mathbf{F}_{y,Ij} n_{y,ij}] = - \sum_{j \in K_i} [(\mathbf{F}_{x,Ij} - \mathbf{F}_{x,i}) n_{x,ij} + (\mathbf{F}_{y,Ij} - \mathbf{F}_{y,i}) n_{y,ij}] \tag{56}$$

Equation 56 is actually a discretisation of the flux gradient within the node N_i and the face L_{ij}. Hence, it can be rewritten as:

$$\mathbf{R}_i = - \sum_{j \in K_i} \left(\frac{\partial \mathbf{F}_x}{\partial x} + \frac{\partial \mathbf{F}_y}{\partial y} \right)_{iI} |\mathbf{n}_{ij}| \frac{d_{\perp,ij}}{2} = - \sum_{j \in K_i} \mathbf{S}_{iI} |\mathbf{n}_{ij}| \frac{d_{\perp,ij}}{2} \tag{57}$$

where it has been considered that we want to maintain a balance between the convective flux and the bed slope source term in the hydrostatic case. The vector $\mathbf{S}_{iI}$ in Equation 57 accounts for a discretisation of the bed slope term between the node $\mathbf{N}_i$ and the face $\mathbf{L}_{ij}$. Having this in mind, and considering Equation 54, the discretisation of the source term applied to the finite volume defined in Figure 1 is given by:

$$\begin{aligned} \mathbf{S}_i^* &= \mathbf{S}_i + \frac{1}{A_i} \mathbf{R}_i = \mathbf{S}_i - \frac{1}{\mathbf{A}_i} \sum_{j \in K_i} \mathbf{S}_{iI} |\mathbf{n}_{ij}| \frac{d_{\perp,ij}}{2} \\ &= \mathbf{S}_i - \frac{1}{A_i} \sum_{j \in K_i} \left[-|\mathbf{n}_{ij}| \frac{d_{\perp,ij}}{2} g \frac{h_i + h_{Ij}}{2} 2 \frac{z_{b,Ij} - z_{b,i}}{d_{\perp,ij}} \right] \begin{pmatrix} 0 \\ \tilde{n}_{x,ij} \\ \tilde{n}_{y,ij} \end{pmatrix} \end{aligned} \tag{58}$$

where $\mathbf{S}_i^*$ is the new discretisation of the source term. In Equation 58, the upper-case subindices refer to extrapolated variables at the cell faces, and the lower-case subindices refer to node values. It is straightforward to show that the source term $\mathbf{S}_i^*$ balances exactly the convective flux in the hydrostatic case.

All of the applications presented in this paper have been computed with both the hybrid scheme (first order for the water depth and second order for the unit discharges) as well as with the fully second order scheme. No significant differences have been found in the results given by both schemes, probably because the water depth gradients are not large enough. On the other hand, the hybrid scheme has been found to be much more stable than the fully second order scheme, especially when the bathymetry is irregular and there are unsteady wet-dry fronts in the solution.

4.3 *Turbulent and viscous diffusion*

The diffusion term is discretised with a semi-implicit centred scheme. In the following we will refer only to the x-momentum equation, with analogous expressions for the y-momentum equation.

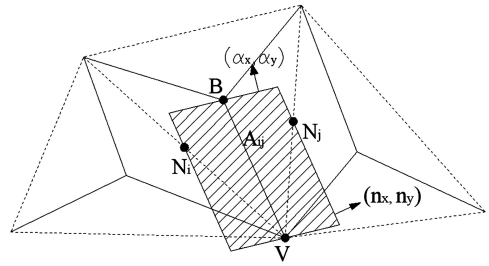


Figure 2. Discretisation of the diffusion term.

Integration of the diffusion term in the x-momentum equation over the cell C_i gives:

$$D_{tot} = \int_{C_i} \frac{\partial}{\partial x}\left(\nu_e h \frac{\partial U_x}{\partial x}\right) + \frac{\partial}{\partial y}\left(\nu_e h \frac{\partial U_x}{\partial y}\right) dA \approx \sum_{j \in K_i} \nu_{e,ij} h_{ij} \left(\frac{\partial U_x}{\partial x} n_x + \frac{\partial U_x}{\partial y} n_y\right)_{ij} \tag{59}$$

When the eddy viscosity is large, it is convenient to implicit the main diagonal of the diffusion source term in order to relax the stability condition over the time step. In order to do so, following the ideas of Davidson (Davidson 1993), the total diffusive flux (D_{tot}) is split in two parts: an orthogonal diffusion ($D_\perp$) and a non-orthogonal diffusion ($D_\parallel$), so that $D_{tot} = D_\perp + D_\parallel$.

The gradient of the velocity at the cell face L_{ij} is computed applying the Gauss theorem on the volume A_{ij}, which is defined by the shaded area in Figure 2. The two edges of the volume A_{ij} which pass through the nodes N_i and N_j are defined by the same normal vector as the cell face L_{ij}, i.e. $\mathbf{n}_{ij}$. The other two edges are parallel to the line which joins the nodes N_i and N_j, and are defined by the normal vector $\boldsymbol{\alpha}_{ij}$, with modulus $|\boldsymbol{\alpha}_{ij}| = |\mathbf{r}_{ij}|$, so that $\boldsymbol{\alpha}_{ij}\mathbf{r}_{ij} = 0$. The area of the volume is given by $A_{ij} = |\mathbf{n}_{ij}| d_{\perp,ij} = \mathbf{r}_{ij}\mathbf{n}_{ij}$. With these definitions, the gradient of U_x at the cell face L_{ij} is evaluated as:

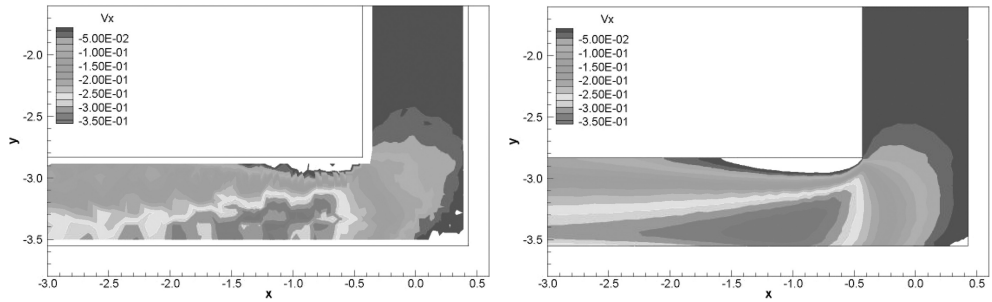
$$\begin{aligned} \left.\frac{\partial U_x}{\partial x}\right|_{ij} &\approx \frac{1}{A_{ij}} \int_{A_{ij}} \frac{\partial U_x}{\partial x} dA = \frac{1}{A_{ij}} \int_L U_x \tilde{n}_x dL \\ &\approx \frac{1}{A_{ij}} (U_{x,j} n_{x,ij} + U_{x,B} \alpha_{x,ij} - U_{x,i} n_{x,ij} - U_{x,V} \alpha_{x,ij}) \end{aligned} \tag{60}$$

with an analogous expression for $\left.\frac{\partial U_x}{\partial y}\right|_{ij}$. The following expression is obtained for the discrete diffusive flux at the cell C_i:

$$\begin{aligned} D_{tot} \approx & \underbrace{\sum_{j \in K_i} \nu_{e,ij} h_{ij} \frac{|\mathbf{n}_{i,j}|}{d_{\perp,ij}} (U_{x,j} - U_{x,j})}_{D_\perp \equiv \text{orthogonal}} \\ & + \underbrace{\sum_{j \in K_i} \nu_{e,ij} h_{ij} \frac{d_{i,j}}{d_{\perp,ij}} (U_{x,B} - U_{x,V})(\tilde{\alpha}_{x,ij} \tilde{n}_{x,ij} + \tilde{\alpha}_{y,ij} \tilde{n}_{y,ij})}_{D_\parallel \equiv \text{orthogonal}} \end{aligned} \tag{61}$$

The non-orthogonal part ($D_\parallel$) is treated explicitly with the rest of source terms, while the orthogonal part ($D_\perp$) is split as:

$$D_{\perp,ij} = \nu_{e,ij} h_{ij} \frac{|\mathbf{n}_{i,j}|}{d_{\perp,ij}} (U_{x,j} - U_{x,i}) = \Gamma_{D_{\perp,ij}} U_{x,j} - \frac{\Gamma_{D_{\perp,ij}}}{h_i} q_{x,i} \tag{62}$$



(a) Experimental. (b) κ–ε model.

Figure 3. Velocity field V_x(m/s) in the 90° bend. White colour accounts for positive V_x.

where $\Gamma_{D_{\perp,ij}} = \frac{v_{e,ij} h_{ij} |\mathbf{n}_{i,j}|}{d_{\perp,ij}}$ is the orthogonal diffusion coefficient. In Equation (62) all the variables are evaluated at time t^n except the unit discharge $q_{x,i}$, which is evaluated at time t^{n+1}. In this way no additional computational cost is introduced, since there is no need to solve any system of equations. In orthogonal meshes the vectors $\mathbf{n}_{ij}$ and $\boldsymbol{\alpha}_{ij}$ are perpendicular and therefore, the non-orthogonal diffusion in Equation (61) vanishes ($D_{\parallel} = 0$).

5 APPLICATIONS AND EXPERIMENTAL VALIDATION

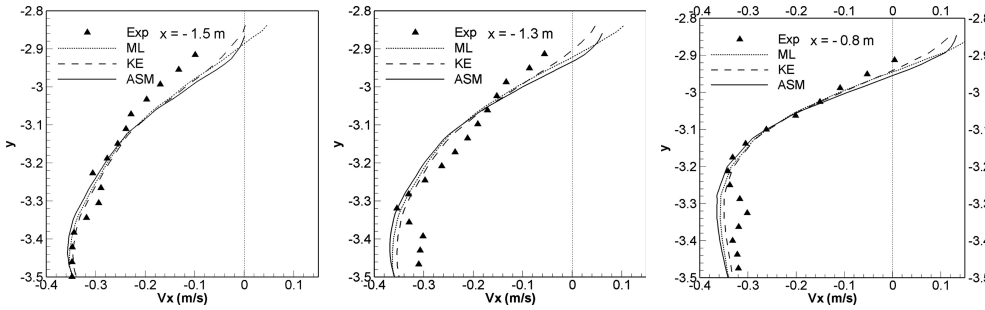
5.1 *Open channel flow in a 90° bend*

Shallow flow in open channels is one potential application of the 2D-SWE. In the absence of abrupt obstacles which might induce 3D flow features, the flow is usually well described by the shallow water approximation. However, strong changes in the direction of the channel can produce 2D recirculation eddies, as well as 3D flow patterns which might deteriorate the results given by the 2D-SWE.

The water discharge in the channel is 29.5 l/s, the mean water depth is around 0.175 m, and the horizontal velocities are of the order of 0.20 m/s. The ratio between water depth and horizontal length scale is approximately $H/L \approx 0.2$, and the Froude number $F_R \approx 0.15$. The most challenging point of the numerical simulation is the correct prediction of the separated region which appears in the inner corner of the bend. It is well known that RANS models usually fail to give accurate predictions of strong separated flows over 3D geometries. However, in this case the 2D forcing produced by the shallowness of the flow is expected to reduce the 3D features of the flow field, helping to improve the accuracy of the results. In order to correctly predict the recirculation eddy, three depth averaged turbulence models have been used in the simulations: the mixing length model, the $k - \varepsilon$ model of Rastogi and Rodi, and the depth averaged algebraic stress model presented in section 3.9.

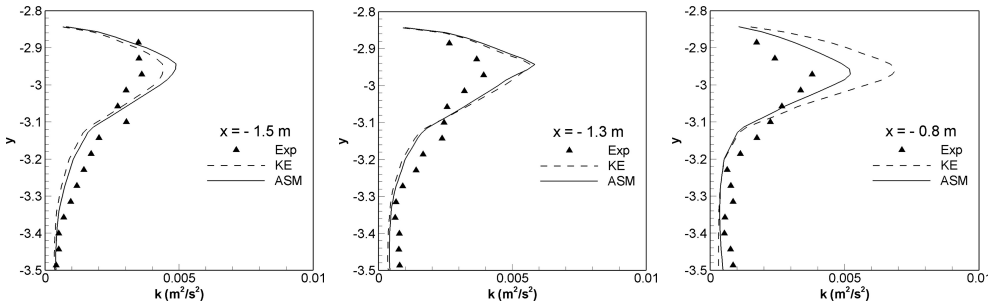
It should be noticed that in uniform channel flow the velocity profile is quite homogeneous in the vertical direction, and the velocity at z_{exp} is approximately equal to the depth averaged velocity. However, near the bend a vertical eddy generates in the flow, which tends to destroy the homogeneity in the vertical direction. This has two main effects. First, the differences between the depth averaged velocity and the experimental velocity measured at z_{exp} may increase, depending on the actual vertical profile. Second, a modelling error is introduced in the depth averaged equations, which neglect the dispersion terms due to the non-homogeneity of the vertical profiles. Despite these 3D flow features appearing in the near bend region, the overall agreement between the experimental and the numerical fields is very satisfactory, as it is shown in Figure 3.

Regarding the performance of the turbulence models, the differences between the computed velocity fields, which appear mainly in the recirculation bubble, are really small compared with



(a) Section $x = -1.5$ m (b) Section $x = -1.3$ m (c) Section $x = -0.8$ m

Figure 4. Longitudinal velocity V_x(m/s) at several cross sections in the 90° bend.



(a) Section $x = -1.5$ m (b) Section $x = -1.3$ m (c) Section $x = -0.8$ m

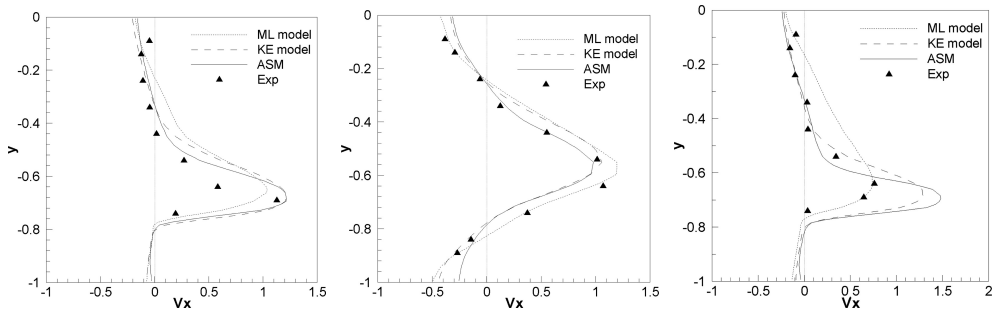
Figure 5. Turbulent kinetic energy k(m^2/s^2) at several cross sections in the 90° bend.

the scatter in the experimental data (Figure 4). The ML model gives a slighter poorer prediction of the bubble length, but its degree of accuracy is rather satisfactory, specially considering the simplicity of the model. This is because the model does not account for transport processes, and therefore, all the turbulence generated in the inner corner of the bend is not convected, neither diffused, downstream the bend, as it is in the DASM and $k-\varepsilon$ models. The DASM and $k-\varepsilon$ models give very similar results. Larger differences appear in the turbulence intensity near the bend (cross section $x = -0.8\,m$), where the maximum turbulent kinetic energy is better predicted by the DASM than by the $k-\varepsilon$ model (Figure 5(c)). However, the differences between models smear out really fast downstream, and no difference at all can be observed at the cross section $x = -1.3$ m (Figure 5(b)). Although both models slightly over-predict the turbulence intensity inside the bubble, the overall agreement and downstream evolution is very satisfactory.

5.2 *Turbulent flow in a vertical slot fishway*

A vertical slot fishway is a channel divided into several pools separated by vertical slots through which the water flows downstream. Although the geometry of a vertical slot fishway is two-dimensional, the ratio between the water depth and the length of the pool is of the order of 0.5, and therefore, it is not straightforward that the 2D-SWE are able to reproduce the flow pattern in a vertical slot fishway. In fact, the hypotheses which are made in the derivation of the 2D-SWE are broken in some regions of the flow, specially near the slot. However, the numerical and experimental data agree quite well.

From the experimental results of several researchers (Puertas et al. 2004; Pena et al. 2004) it follows that for uniform flow conditions the velocity field in vertical slot fishways is almost independent of the total flow discharge. On the other hand, the water depth is proportional to the



(a) Q = 35 l/s. x = 0.16 m (b) Q = 35 l/s. x = 0.46 m (c) Q = 35 l/s. x = 0.16 m

Figure 6. Numerical and experimental depth averaged longitudinal velocity V_x(m/s) at several cross sections.

flow discharge with an almost linear relation. The first feature we should expect from the numerical model is to reproduce this behaviour. In order to prove so, three different discharges covering almost the full range of experimental flow conditions were used in the simulations: 35 l/s, 65 l/s and 105 l/s.

The flow pattern in the pools can be described as a main jet which crosses the pool from the inlet to the outlet slot (Figure 6). At each side of the jet a recirculation eddy appears. The flow separates behind the inlet slot baffie and the so-called lower eddy appears. Going through the outlet slot the flow is forced to reattach. Both eddies are driven by turbulent shear stress, and therefore their prediction is very dependent on the turbulence model used.

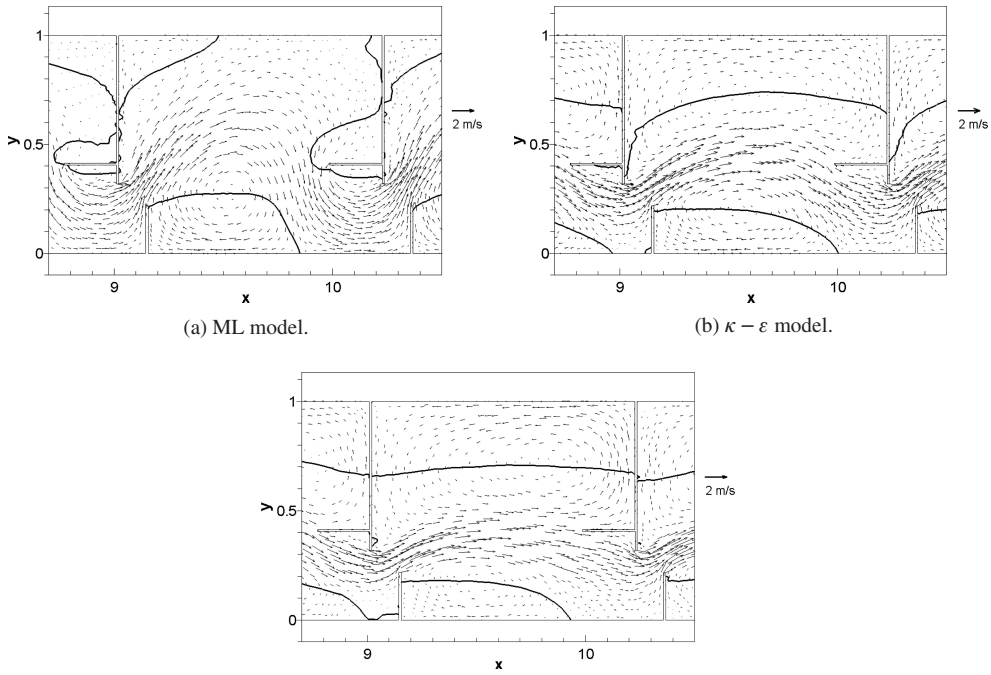
A correct turbulence modelling is of great importance to predict all the previously described features. An excessive turbulence level will diffuse too much the velocity profiles, tending to eliminate the recirculation eddies. On the other hand, a low level of turbulent energy will sharpen the profiles too much, predicting excessively high velocities.

A comparison of the depth averaged longitudinal velocity at several cross sections reveals a quite satisfactory agreement between the experimental and numerical results (Figure 6). The ASM and $k-\varepsilon$ models predict fairly well the width of the recirculation eddies as well as the evolution of the maximum velocity as the flow crosses the pool. Although both models give slightly different results, none of them can be considered in better agreement with the experimental data. The predictions of the ML model are somewhat worse, specially as the discharge increases.

The DASM and $k-\varepsilon$ models predict maximum velocity and recirculation regions which are quite insensitive to the total discharge. This result is in direct agreement with the experimental data. The mixing length (ML) model is more sensitive to the total discharge. This is because the length scale used to compute the eddy viscosity with the ML model depends directly on the water depth (Equation 11). Therefore, with larger depths the turbulent stresses (diffusive forces) increase and the velocity profiles become more flattened. The dependence of the turbulent length scale on the water depth is a weakness of the ML model.

The ASM gives better predictions of the turbulence field than the $k-\varepsilon$ model, specially regarding the Reynolds stresses. The $k-\varepsilon$ model fails to predict the anisotropy between the two horizontal Reynolds stresses, and gives excessively large values of the transversal component $\overline{v'^2}$ (Figure 8). On the other hand, the agreement between the ASM predictions of $\overline{v'^2}$ and the experimental data is remarkable.

Nonetheless, the global agreement between the experimental and numerical normal Reynolds stresses is quite satisfactory. The largest differences appear in the main jet stream, where the ASM tends to underpredict the turbulent energy, while the $k-\varepsilon$ model overpredicts it (Figure 8). The excessively large turbulence level given by the $k-\varepsilon$ model just downstream the inlet slot might be explained by the strong shear strain, with velocity gradients of the order of $20\,s^{-1}$. The flow conditions in this region, with a very strong swirl, as well as separation and reattachment in a very short distance, are beyond the capabilities of the eddy viscosity turbulence models. The

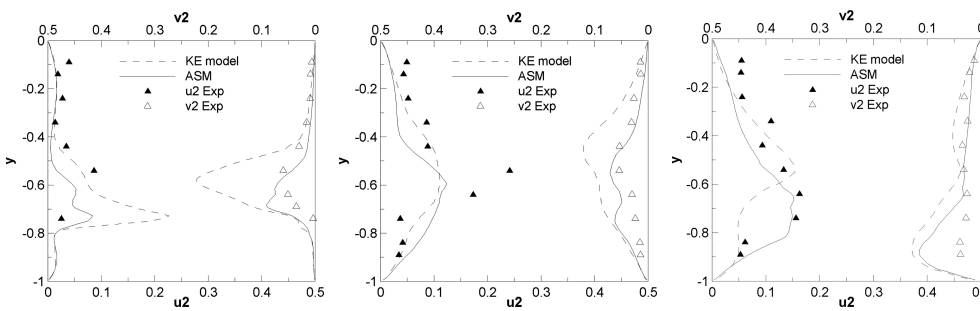


(a) ML model.

(b) $\kappa - \varepsilon$ model.

(c) DASM.

Figure 7. Depth averaged velocity field. $Q = 105$ l/s. The black line separates the regions with positive and negative longitudinal velocity.



(a) Q = 65 l/s. $x = 0.16$ m

(b) Q = 65 l/s. $x = 0.46$ m

(c) Q = 65 l/s. $x = 0.86$ m

Figure 8. Numerical and experimental depth averaged horizontal Reynolds stresses $\overline{u'^2}$ and $\overline{v'^2}$(m^2/s^2) at several cross sections. Q = 651 l/s.

largest source of turbulence in the pool is due to the strong shear strain in the inlet slot and lower eddy. Although the turbulence production in the $k - \varepsilon$ model is controlled by the production limiter (Equation 25) and the realizability condition (Equation 27), the ASM still gives lower values for the turbulent stresses in the slot stagnation regions and in the lower recirculation eddy. The maximum value of the turbulent kinetic energy given by the $k - \varepsilon$ model is inside the eddy region, with values around 0.24 m^2/s^2. With the ASM the peak values are lowered to 0.18 m^2/s^2 and moved downstream.

The strong anisotropy between the longitudinal and transversal Reynolds stresses downstream the inlet slot is well predicted by the ASM model. While the longitudinal component ($\overline{u'^2}$) is

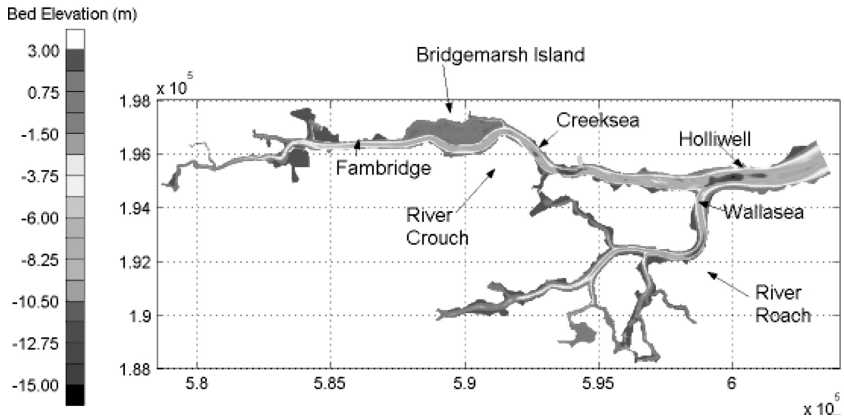


Figure 9. Bathymetry of the Crouch-Roach estuary for the numerical model, relative to mean sea level at the mouth. Locations of tide gauges and current meter deployments are shown.

underpredicted, probably because the turbulence produced by the air bubbles in the slot region is not considered in the model, the transversal component ($\overline{v'^2}$) is quite accurately computed. On the other hand, the $k-\varepsilon$ model overpredicts both components in this region, specially the transversal one.

5.3 *Tidal flow in the Crouch estuary*

The Crouch estuary, Essex, UK (Figure 9), is a challenging test case if we consider its complex bathymetry, with extensive intertidal flats. The tidal range is approximately 4.8 m at the mouth. The maximum depth is about 15 m at high tides. Freshwater influx is negligible in the whole estuary, and therefore, all the water in the estuary is salt water. The large separation between horizontal and vertical length scales justify the use of a 2D-SWE model.

The flow in the model is completely driven by the water surface elevation at the mouth, which is the only open boundary condition to be imposed. The rest of the boundaries are treated as walls. Due to the size and complexity of the computational domain (27.65 km^2) the mesh size near the walls is relatively coarse and therefore, a slip condition has been used at the wall boundaries. Accordingly, the normal velocity as well as the diffusion of k and ε are set to zero at the boundary faces. Nonetheless, most of the time during the simulation the wall boundaries are dry and thus, they do not participate in the solution. On the other hand the wet-dry front, which defines the fluid extension, plays a more important role in the flow field.

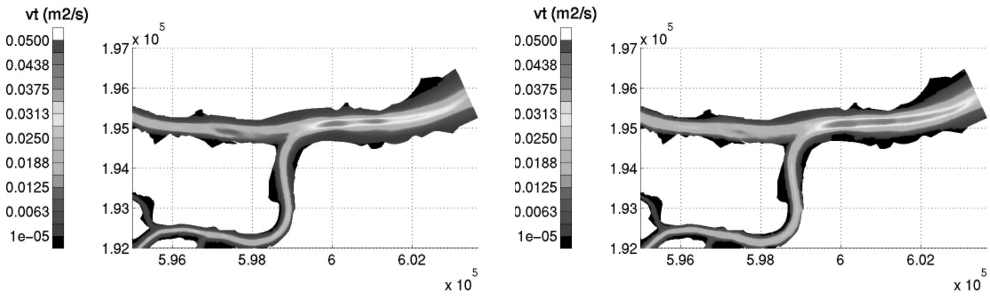
The water surface elevation at the mouth is directly obtained from a tidal gauge. The flow at the mouth is always subcritical. Hence, only the water depth needs to be imposed during the ebb tide (subcritical outlet). During the flood tide (subcritical inlet) the flow is forced to be perpendicular to the boundary, as an additional condition to the imposed water depth.

A Manning coefficient of $n = 0.02\,\mathrm{ms}^{-1/3}$ was used in the inter-tidal and sub-tidal regions, where the bottom sediments are predominantly muddy. There are also several marsh regions, which are assigned a Manning coefficient $n = 0.05\,\mathrm{ms}^{-1/3}$.

After using several meshes with different resolution, a mesh with approximately 50000 volumes was chosen. The numerical mesh used covers the whole estuary using 48995 control volumes. The area of the control volumes in the main channel is about 300 m^2.

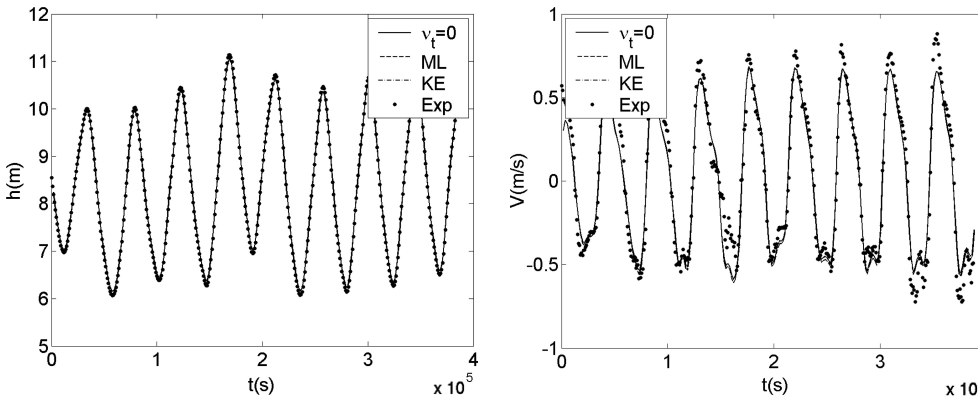
The hybrid scheme was used in the computations, which proved to be much more stable than the fully second order scheme, while producing almost identical results.

Both the ML and $k-\varepsilon$ models give very similar eddy viscosity fields (Figure 10). This is because turbulence is mainly produced by bed friction, and both models account in a similar way for this process. For uniform channel flow the eddy viscosity given by the $k-\varepsilon$ model reduces



(a) ML model. (b) κ–ε model.

Figure 10. Eddy viscosity field at the mouth of the estuary. Ebb tide. ML and $k-\varepsilon$ models. Dry regions are shown in black.



(a) Water depth. (b) Velocity.

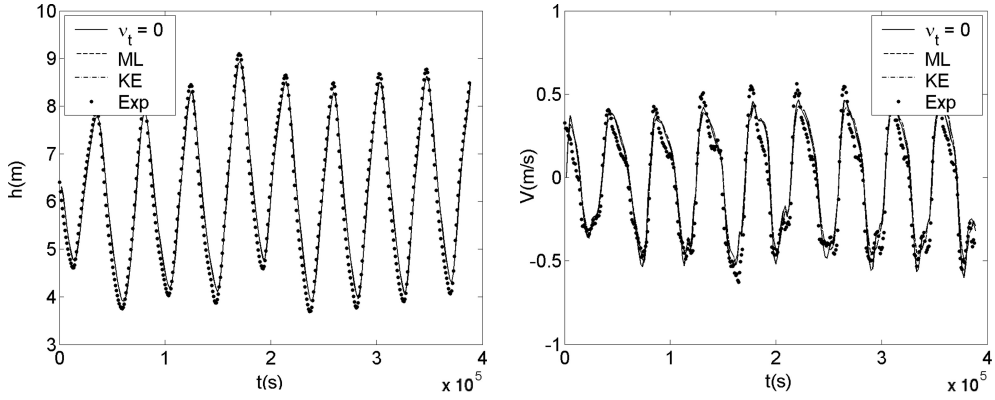
Figure 11. Water depth and horizontal velocity time series at Holliwell. ML and $k-\varepsilon$ turbulence models, and zero eddy viscosity ($\nu_t = 0$).

to $\nu_t^{k-\varepsilon} = 0.08 u_f h$, while the ML model gives $\nu_t^{ML} = 0.068 u_f h$ (see section 3). For this reason the velocity and water depth fields are very insensitive to the turbulence model used.

Even in the mouth of the estuary, where the turbulence intensities are highest (Figure 10), the turbulent horizontal Reynolds number is rather large ($R_t = UL/\nu_t \approx 21000$). This means that the turbulent diffusion forces are small when compared to the convective forces, which diminishes the influence of the turbulence model on the mean velocity field. Velocity and water depth fields independent of the turbulence model were also obtained by Babarutsi et al. (Babarutsi et al. 1996) when modelling shallow recirculating flows dominated by bed friction, by Davies et al. (Davies et al. 1997) when computing the tidal flow in the Irish Sea, and by Lloyd and Stansby (Lloyd and Stansby 1997) when modelling the flow around conical islands.

The velocity field in the whole estuary is very dependent on the bathymetry, with velocities being highest in the deepest regions of the estuary. The sensitivity of the numerical results to the bed friction coefficient and turbulence model is much smaller than the differences between numerical and experimental data (Figures 11 and 12). Some of the disagreements between numerical and experimental results are probably due to errors in the local bathymetry of the model. The fact of comparing the experimental velocity measured at a fixed height of 2 m above the bed, with the numerical depth-averaged velocity, might also account for some of these discrepancies.

Figures 11 and 12 show several numerical and experimental time-series for water depth and current speed. In general water depth is rather well predicted by the model (Figures 11(a) and



(a) Water depth. (b) Velocity.

Figure 12. Water depth and horizontal velocity time series at Fambridge. ML and $k-\varepsilon$ turbulence models, and zero eddy viscosity ($\nu_t = 0$).

12(a)) except at Fambridge, where the tidal amplitude is slightly underpredicted. Regarding the current speed, both ebb and flood flows at Holliwell and Fambridge are well predicted, although the maximum ebb velocity is slightly underpredicted.

6 CONCLUSIONS

In this paper we have presented a comprehensive study on the depth averaged turbulence modelling on shallow flows, including a critical description of the equations and numerical schemes used to solve them, practical applications and experimental validation. Depth averaged mixing length, $k-\varepsilon$ and Algebraic Stresss models have been discussed.

In order to avoid spurious oscillations when the bathimetry is irregular, an upwind discretisation of the bed slope term has been described. Alternatively to a fully second order scheme, we have described a rather simple, free of spurious oscillations and stable scheme, which reduces the numerical diffusion in a significant way, which just uses a second order discretisation for the two unit discharge components, whilst keeping a first order discretisation for the water depth and bed elevation. The resulting hybrid scheme has been used in the practical applications, giving accurate and stable results. Nevertheless, it is important to remark that the water depth gradient in all the applications was quite smooth.

It turned up that, even if the flow is fully turbulent in all the applications, the importance of the turbulence modelling is much dependent on the problem, to the point that in the Crouch estuary turbulence has not importance at all in the global flow field, while in the fishway its role is critical. Still, it should be remarked that the turbulence field might be important in order to compute sediment transport or dispersion of pollutants, which are problems usually dominated by turbulent diffusion.

The comparison between the numerical and experimental results in the applications presented in section 5 showed that the depth averaged shallow water equations, coupled with a suitable turbulence model, may be used in order to compute the free surface flow in vertical slot fishways as well as the tidal flow in complex estuaries with extensive tidal flats. The importance of a correct turbulence modelling in the fishway was confirmed by the numerical results. The ML model failed to predict the velocity field accurately, specially for large water discharges, because the model assumes a turbulent length scale proportional to the water depth, which is not correct in the fishway flow. The DASM and $k-\varepsilon$ models give fairly accurate velocity fields. The original purpose of using the algebraic stress model was to account for the effects of the anisotropic Reynolds stresses

in the flow. When compared to the $k-\varepsilon$ model, which assumes an isotropic eddy viscosity, the DASM improves the prediction of the turbulent stresses. On the other hand, no improvement has been found in the water depth and velocity fields, which are predicted with a similar degree of accuracy with either the DASM or $k-\varepsilon$ the model.

REFERENCES

Altai, S., J. Zhang, and V. H. Chu (1999). Shallow turbulent flow simulation using two-length-scale model. *J. Eng. Mech. 125*(7), 780–788.

Babarutsi, S. and V. H. Chu (1991). A two-length-scale model for quasi-two-dimensional turbulent shear flows. In *Proc. 24th Congr. of IAHR, Vol. C, Madrid, Spain*, pp. 51–60.

Babarutsi, S. and V. H. Chu (1998). Modelling transverse mixing layer in shallow open-channel flows. *J. Hydraul. Eng. 124*(7), 718–727.

Babarutsi, S., M. Nassiri, and V. H. Chu (1996). Computation of shallow recirculating flow dominated by friction. *J. Hydraul. Eng. 122*(7), 367–372.

Bermúdez, A., A. Dervieux, J. A. Desideri, and M. E. Vázquez-Cendón (1998). Upwind schemes for the two-dimensional shallow water equations with variable depth using unstructured meshes. *Comput. Methods Appl. Mech. Eng. 155*, 49–72.

Bermúdez, A. and M. E. Vázquez-Cendón (1994). Upwind methods for hyperbolic conservation laws with source terms. *Comput. Fluids 23*(8), 1049–1071.

Booij, R. (1989). Depth-averaged $k-\varepsilon$ modelling. In *Proc. 23rd IAHR Congr., Delft, The Netherlands*, Volume A, pp. 198–206. IAHR.

Brufau, P. (2000). Simulación bidimensional de flujos hidrodinámicos transitorios en geometrías irregulares. Tesis doctoral. Área de Mécanica de Fluidos. Universidad de Zaragoza.

Brufau, P. and P. García-Navarro (2000). Two-dimensional dam break flow simulation. *Int. J. Numer. Meth. Fluids 33*, 35–57.

Brufau, P., P. García-Navarro, and M. E. Vázquez-Cendón (2004). Zero mass error using unsteady wetting-drying conditions shallow flows over dry of irregular topography. *Int. J. Numer. Meth. Fluids 45*, 1047–1082.

Cea, L. (2005). An unstructured finite volume model for unsteady turbulent shallow water flow with wet-dry fronts: Numerical solver and experimental validation. Doctoral Thesis. Departamento de Métodos Matemáticos y de Representación. Universidad de A Coruña.

Cea, L., A. Ferreiro, M. E. Vázquez-Cendón, and J. Puertas (2004). Experimental and numerical analysis of solitary waves generated by bed and boundary movements. *Int. J. Numer. Meth. Fluids 46*(8), 793–813.

Cea, L., J. French, and M. E. Vázquez-Cendón (2006). Numerical modelling of tidal flows in complex estuaries including turbulence: an unstructured finite volume solver and experimental validation. *Int. J. Numer. Meth. Engineering 67*(13), 1909–1932.

Cea, L., L. Pena, J. Puertas, M. E. Vázquez-Cendón, and E. Peña (2007). Application of several depth-averaged turbulence models to simulate flow in vertical slot fishways. *J. Hydraul. Eng. 133*(2), 160–172.

Cea, L., J. Puertas, and M. E. Vázquez-Cendón (2007). Depth averaged modelling of turbulent shallow water flow with wet-dry fronts. *ARCME In press.*

Davidson, L. (1993). Implementation of a $k-\varepsilon$ model and a Reynolds Stress Model into a multiblock code. Technical Report CRS4-APPMATH-93-21, Applied Mathematics and Simulation Group CRS4, Cagliary, Italy.

Davies, A. M., J. E. Jones, and J. Xing (1997). Review of recent developments in tidal hydrodynamic modeling. II: Turbulence energy models. *J. Hydraul. Eng. 123*(4), 278–292.

Dervieux, A. and J. A. Desideri (1992). Compressible flow solvers using unstructured grids. Rapports de Recherche 1732, INRIA.

Ding, Y., Y. Jia, and S. S. Y. Wang (2004). Identification of Manning's roughness coefficients in shallow water flows. *J. Hydraul. Eng. 130*(6), 501–510.

Duan, J. G. (2004). Simulation of flow and mass dispersion in meandering channels. *J. Hydraul. Eng. 130*(10), 964–976.

Durbin, P. (1996). On the $k-\varepsilon$ stagnation point anomaly. *Int. J. Heat Fluid Flow 17*, 89–90.

Fernández-Nieto, E. D. (2003). Aproximación numérica de leyes de conservación hiperbólicas no homogéneas. Aplicación a las ecuaciones de aguas someras. Tesis doctoral. Universidad de Sevilla.

García-Navarro, P. and M. E. Vázquez-Cendón (2000). On numerical treatment of the source terms in the shallow water equations. *Comput. Fluids 29*, 951–979.

Harten, A., P. Lax, and A. van Leer (1983). On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Rev. 25*, 35–61.

Hsieh, T. Y. and J. C. Yang (2003). Investigation on the suitability of two-dimensional depth-averaged models for bend-flow simulation. *J. Hydraul. Eng. 129*(8), 597–612.

Hubbard, M. E. and N. Dodd (2002). A 2D numerical model of wave runup and overtopping. *Coast. Eng. 47*, 1–26.

Hubbard, M. E. and P. García-Navarro (2000). Flux difference splitting and the balancing of source terms and flux gradients. *J. Comput. Phys. 165*, 89–125.

Jia, Y. and S. S. Y. Wang (1999). Numerical model for channel flow and morphological change studies. *J. Hydraul. Eng. 125*(9), 924–933.

LeVeque, R. J. (2002). *Finite Volume Methods for Hyperbolic Problems*, Volume 31 of *Cambridge Texts in Applied Mathematics*. Cambridge University Press.

Lloyd, P. M. and P. K. Stansby (1997). Shallow water flow around model conical islands of small side slope. Part 1: Surface piercing. *J. Hydraul. Eng. 123*, 1057–1067.

Menter, F. R. (1993). Zonal two-equation $k - \omega$ turbulence models for aerodynamic flows, aiaa paper 93–2906. In 24th *Fluid Dynamics Conference, 6–9 July 1993, Orlando, FL*.

Minh-Duc, B., T. Wenka, and W. Rodi (2004). Numerical modeling of bed deformation in laboratory channels. *J. Hydraul. Eng. 130*(9), 894–904.

Molls, T. and M. H. Chaudhry (1995). Depth-averaged open-channel flow model. *J. Hydraul. Eng. 121*(6), 453–465.

Pena, L., L. Cea, and J. Puertas (2004). Turbulent flow: An experimental analysis in vertical slot fishways. In *Fifth International Symposium on Ecohydraulics, Madrid, Spain*, pp. 881–888. IAHR.

Puertas, J., L. Pena, and T. Teijeiro (2004). An experimental approach to the hydraulics of vertical slot fishways. *J. Hydraul. Eng. 130*(1), 10–23.

Rastogi, A. K. and W. Rodi (1978). Predictions of heat and mass transfer in open channels. *Journal of the Hydraulics Division HY3*, 397–420.

Roe, P. L. (1986). Discrete models for the numerical analysis of time-dependent multidimensional gas dynamics. *J. Comput. Phys. 63*, 458–476.

Toro, E. F. (1999). *Riemann Solvers and Numerical Methods for Fluid Dynamics. Second Edition*. New York: Springer-Verlag.

Toro, E. F. (2001). *Shock-capturing Methods for Free-Surface Shallow Flows*. Chichester, West Sussex PO19 1UD, England: Wiley.

Vázquez-Cendón, M. E. (1996). An efficient upwind scheme with finite volumes of the edge-type for the bidimensional shallow water equations. In F. Benkhanldoun and R. Vilsmeier (Eds.), *Finite Volumes for Complex Applications. Problems and Perspectives*, pp. 605–612. Hermes.

Vázquez-Cendón, M. E. (1999). Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry. *J. Comput. Phys. 148*, 497–526.

Winterwerp, J. C., Z. B. Wang, A. T. M. van Kester, and J. F. Verweij (2002). Far-field impact of water injection dredging in the Crouch River. *Water and Maritime Engineering 154*(4), 285–296.

Wu, W. (2004). Depth-averaged two-dimensional numerical modeling of unsteady flow and nonuniform sediment transport in open channels. *J. Hydraul. Eng. 130*(10), 1013–1024.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Application of distributed hydrologic models

R.F. Vázquez
Centro de Investigación y Tecnología Agroalimentaria de Aragón, Unidad de Suelos y Riegos, Zaragoza, Spain

J. Feyen
Division Soil and Water Management, Department of Land Management and Economics, K.U. Leuven, Heverlee, Belgium

ABSTRACT: Physically based distributed models, when properly applied, have the potential of being operational tools that can be useful to water resources planning and management. They can account for the distributed variation of physical properties of the domain being modelled as well as produce spatially distributed predictions. In theory, these models do not need to be calibrated because their parameters may be assessed directly from field measurements. However, in practice, due to the lack of accurate and well distributed data, sub-grid variability of parameters, the uncertainty attached to the definition of initial and boundary conditions and the discrepancies among the scale to which the mathematical structure of the models were derived, the scale to which field data are collected and the scale to which these models are finally applied, this type of models have to be calibrated. In this context, with respect to the application of distributed hydrologic models on a catchment scale, this study presents 5 study cases to briefly address some modelling aspects such as (i) how best to represent hydrological processes; (ii) how best to identify the parameters of distributed models at different scales; (iii) how best to incorporate new sources of data into the modelling process; and (iv) how to assess general uncertainty attached to the application of distributed hydrologic models on a catchment scale. In doing so, different study sites located within Belgium were modelled. Finally, the paper presents some brief thoughts on the future of distributed hydrologic modelling.

1 INTRODUCTION

1.1 *Numerical hydrologic modelling*

When properly applied, hydrologic models as operational tools can supply useful information about flow and transport processes of water and chemicals, assist in the design of remedial programs and provide useful information for decision-making (Vázquez 2003). In this context, it is necessary to adopt an appropriate modelling protocol that includes several steps to set-up the model of the physical domain being modelled, evaluate its accuracy and reliability for future operational use and, if needed, reformulate the conceptualisation of the physical system and/or structure of the model.

A crucial activity in this protocol is the conceptualisation of the physical system and its main hydrological processes. In this stage, field data are processed and assembled including information that will be used for evaluating the model predictions. Figure 1 illustrates this conceptualisation process in the framework of a regional groundwater simulation (Vázquez 1997). Figure 1a sketches the processing and interpretation of the geologic data of the study site. This analysis indicated that the northern region of the aquifer has confined characteristics whilst the southern region has unconfined properties. Correspondingly, constant head boundary conditions defined by the river network where used in the unconfined region where the upper limit of the aquifer coincides with

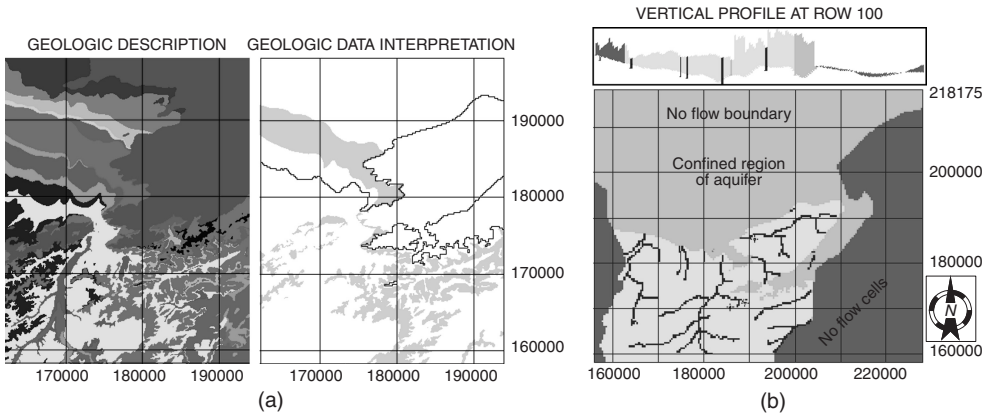


Figure 1. Sketch of (a) the data processing/model conceptualisation approach; and (b) the conceptual distributed (three-dimensional) model of the aquifer (after Vázquez 1997).

the topographic surface. Figure 1b shows the resulting conceptual (three-dimensional) model of the study site as it was implemented in the hydrologic code (Vázquez 1997).

1.2 *Lumped conceptual versus physically-based distributed models*

The rapid increase in computer power and improved insight into hydrological processes has led to the development of models with increasingly complex equations based on physical processes, along with an increasing density of calculation nodes allowing representation of the spatial variability of the system. In this context, several distributed models ranging in sophistication and complexity have been developed such as, MIKE SHE (Refsgaard and Storm 1995), WATBAL (Refsgaard and Knudsen 1996), TOPMODEL (Beven 2001), ECOMAG (Motovilov et al. 1999) and MODFLOW (McDonald & Harbaugh 1988).

The main differences between conceptual and PhB models lie in the way that the processes are described in the model and the way that the catchment spatial variability is treated by the model. PhB models are usually distributed as a means to overcome the problem of sub-catchment variability enabling the variation of parameter values among the different computational grid points (Vázquez et al. 2002a). On the other hand, conceptual rainfall-runoff models do not explicitly take into account the spatial variability of inputs, outputs, or parameters (i.e., they are generally lumped) and for this reason they are usually structured to utilise average values of the catchment characteristics (Célleri et al. 2003).

1.3 *Model calibration*

Lumped conceptual models are usually structured to utilise average values of the catchment characteristics. Due to this fact, model parameters usually do not account for physical properties of the basins and, in consequence, they have to be assessed by model calibration. But, since PhB distributed models are based on the physical description of the main hydrologic processes and are furthermore structured to assimilate easily distributed measurements, is it logic then referring as to the calibration of PhB distributed models?

The scale to which the physical processes that are included in distributed PhB models are described is usually a point scale. The scale to which data are collected is conventionally ranging from point scale to field scale, despite that nowadays, data assimilation from remote sensing and radar sources are becoming more and more feasible, which allows expanding considerably the scale to which data are collected. Nevertheless, these data are still encapsulating some uncertainty since

they are estimates rather than direct measurements (Verhoest et al. 1998). Furthermore, the grid-sizes that are currently being used in distributed models applied to a catchment scale ranges from some tens of meters for small relatively simple catchments (Güntner et al. 1999) up to some thousands of meters for large catchments (Jain et al. 1992, Refsgaard & Knudsen 1996, Refsgaard & Storm 1995, Refsgaard 1997, Christiaens & Feyen 2002, Madsen 2003).

Thus, there is a significant discrepancy among the scale to which the mathematical structure of the codes were derived, the scale to which field data about the modelled system are collected and the scale to which the hydrological codes are finally applied. Owing to this scale discrepancy and to uncertainties attached to data and the definition of initial and boundary conditions, PhB distributed models need indeed to be calibrated somehow to define model parameters that are effective to the modelling scale and which are otherwise difficult to assess, among other reasons, because of the lack of appropriate field measuring methods to the modelling scale that enable direct consideration of sub-grid variability of parameters (Vázquez 2003).

1.4 *Context of this study*

Important issues on the application of distributed hydrologic models are therefore (Beven & Feyen, 2002) (i) how best to represent hydrological processes at different scales; (ii) how best to identify the parameters of distributed models at different scales lacking information and measurement techniques; and (iii) how best to incorporate new sources of data, especially from remote sensing, into the modelling process at different scales. Research on these topics is currently undergoing worldwide considering a very broad and diverse set of approaches.

In this context, this study presents a series of study cases to briefly address these topics. With respect to the first topic, a first study case explores on the feasibility of associating a physical meaning to the parameters of a conceptual model, which would enable not only a better representation of the physical processes that can be attained with a lumped conceptual model applied on a catchment scale but it would enable as well applying the model on ungauged catchments with similar properties. With regard to the second topic on how best to identify parameters of distributed models, study cases 2 and 3 present assessments on the identification of modelling grid-size (scale parameter) and potential evapotranspiration (modelling variable) data, being appropriate for the distributed modelling of a study catchment. With respect to the third topic, study case 4 explores on the incorporation of flow-component data into a catchment scale modelling in order to enhance the physical consistency of the predictions of the distributed model of the study site. Finally, a fifth study case is presented which assesses the general uncertainty attached to the application of a distributed hydrologic model to a catchment scale. This last study case is shown herein to illustrate the need of including uncertainty assessment as a standard component of every modelling protocol.

2 MATERIALS

2.1 *The study sites*

The different methods described in this report were applied on several study sites located in Belgium (Fig. 2a). The distributed modelling addressed in this report however is related only to two study sites: the Gete catchment (Fig. 2b) and the Nete catchment (Fig. 2c).

The Gete catchment (586 km^2), located to the east of Brussels, comprises the sub-catchments of the Grote Gete (326 km^2) and the Kleine Gete (260 km^2). The catchment boundary defined by the topographical divide does not necessary correspond to the actual hydrological divide. However, since no extra information, such as piezometric measurements, was available to assess on the exact location of the hydrological divide, it was assumed to be coincident with the topographical divide of the catchment.

The area is located in the central part of Belgium where the soils usually have a loamy texture, are deep and the average groundwater table is generally at a depth of 3 m to 10 m below surface.

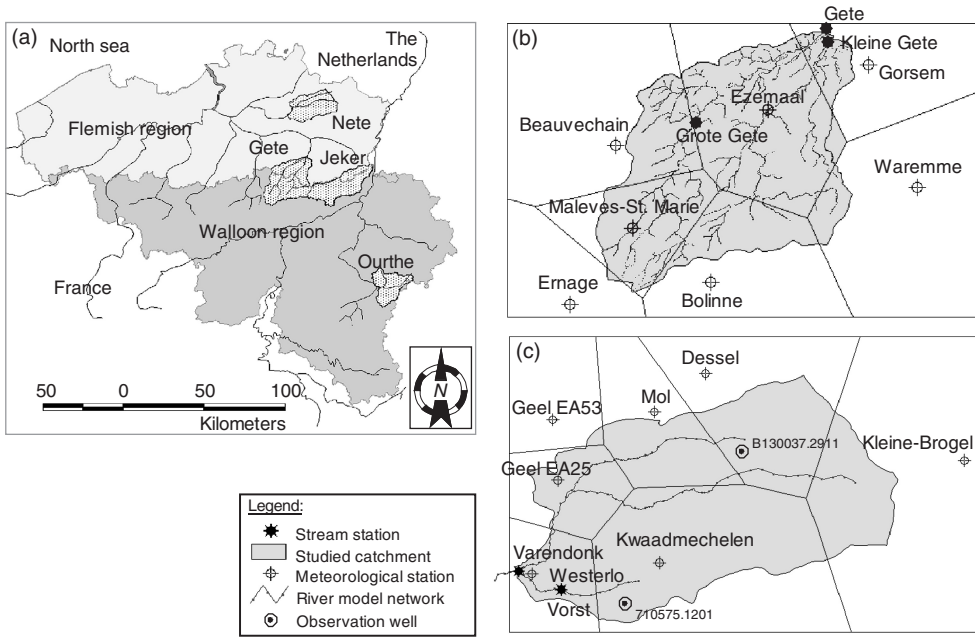


Figure 2. (a) Location of the study sites; and main characteristics of (b) the Gete and (c) the Nete catchments (after Vázquez 2003, Vázquez & Feyen 2004, Vázquez et al., unpubl.).

The elevation of the area varies from approximately 27 m in the northern part to 174 m in the southern parts. Land use is mainly agricultural with both pasture and cultivated fields and some local forested areas (Feyen et al. 2000). Nine soil units can be distinguished according to the Belgian soil classification system, e.g. loamy soils (Aba, Ada and Adc), sand-loamy soils (Lca, Lda and Ldc), clay soils (Eep and Uep) and soils with stony mixtures (Gbb). The dominant soil type in the catchment is the Aba soil unit. The geology of the catchment comprises nine geologic units, some of which occur only in isolated parts of the catchment (Vázquez et al. 2002a). Thiessen polygons based on seven precipitation stations describe the spatially variable precipitation within the catchment (Fig. 2b).

The second study site is the Nete River basin, located on the northern region of Belgium (Fig. 2a). This catchment (356 km^2) comprises the subbasins of the Grote Nete and the Grote Laak up to their confluence point. The elevation of the area varies from approximately 13 m above sea level in the western part to 65 m in the eastern part. Land coverage is mainly agricultural (both pasture and cropland) with some local forested areas although a significant proportion of the area is urbanized (Vázquez et al., unpubl.). According to the Belgian soil map (Vázquez 2003), sandy soils (Zdg, Zbf, Zbm and Zap) are the dominant in the catchment. Light sand loamy soils (Pdc and Sbc) and clayey soils (Uep) can also be distinguished. The average rainfall in the period [1985–1999] is 800 mm. Thiessen polygons based on seven precipitation stations describe the spatially variable precipitation within the catchment (Fig. 2c).

2.2 *The hydrologic models*

The NAM module (DHI, 1999) of the MIKE 11 code (Havnø et al. 1995) was used for the lumped rainfall-runoff simulation of all of the study sites reported in Figure 2a. It performs the simulation by accounting for water content in four different and mutually interrelated reservoirs that conceptualise physical elements of the catchment, namely surface, root zone, groundwater and snow storage.

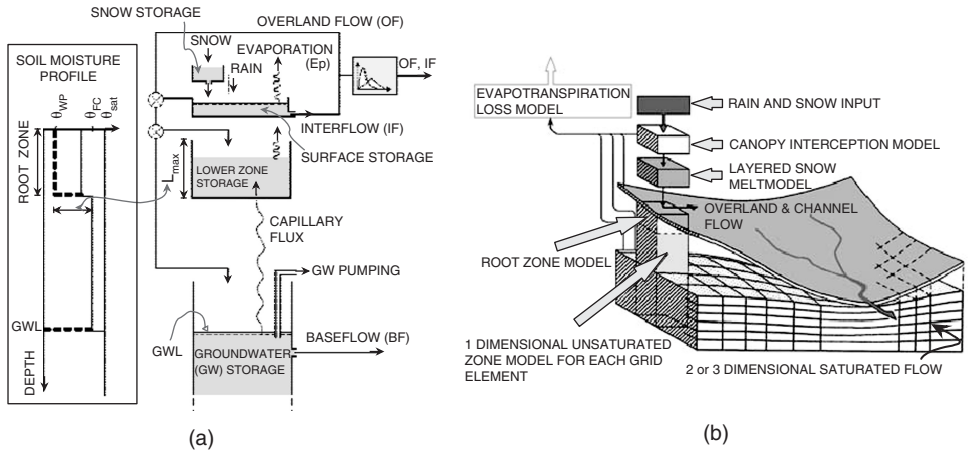


Figure 3. (a) Linear reservoir structure of the NAM model (after Havnø et al. 1995); and (b) schematic representation of the water movement module of the PhB distributed model MIKE SHE (Vázquez 2003).

Irrigation and groundwater pumping can be considered additionally (Fig. 3a). The NAM module is defined as a conceptual model, representing a simplified imitation of the land phase of the hydrologic cycle. As the NAM model is lumped it treats the catchment as a simple unit, with its parameters representing average values for the entire catchment (Célleri et al. 2003).

The distributed model MIKE SHE (Refsgaard and Storm 1995) was used for the integral modelling of the study sites. MIKE SHE is a well-known deterministic (i.e. physically based) model that has been used and described in a wide range of applications (e.g., Refsgaard 1997, Feyen et al. 2000, Christiaens & Feyen 2002, Vázquez & Feyen 2002a). MIKE SHE integrates the entire land phase of the hydrologic cycle and can model interception, actual evapotranspiration (ET_{act}), overland flow, channel flow, flow in the unsaturated zone, flow in the saturated zone and exchange between aquifers and rivers. The MIKE SHE model uses a regular square-grid modelling structure.

3 METHODS

3.1 *Calibration of deterministic models*

The increasing popularity of using hydrological models often leads to a less critical evaluation of the basin being modelled. Often parameter values are selected without taking into account their physical nature and variability, considerations that are essential for PhB distributed models such as MIKE SHE. In lumped conceptual models, the parameters do not have a physical meaning and the parameterisation is an optimisation process not restricted to physical boundaries. The goal of the calibration process for PhB models is to find an optimal set of physically realistic parameter values that simulate the behaviour of the catchment as accurately as possible. To accomplish this, PhB constraints for the values of the parameters have to be specified based on the information available.

Both observations in the field and laboratory experiments on the physical system, as well as mathematical models, are subject to errors (Fig. 4a). Moreover, it is not possible to characterise all the spatial (and temporal) sub-grid variability of the parameters since they represent spatially averaged values rather than point values obtained from field measurements. Thus, optimising the parameters in the calibration process can improve the results.

It is advisable that the most uncertain or critical parameters, (i.e. the parameters with the most significant impact on the results) are defined at the beginning of the calibration process.

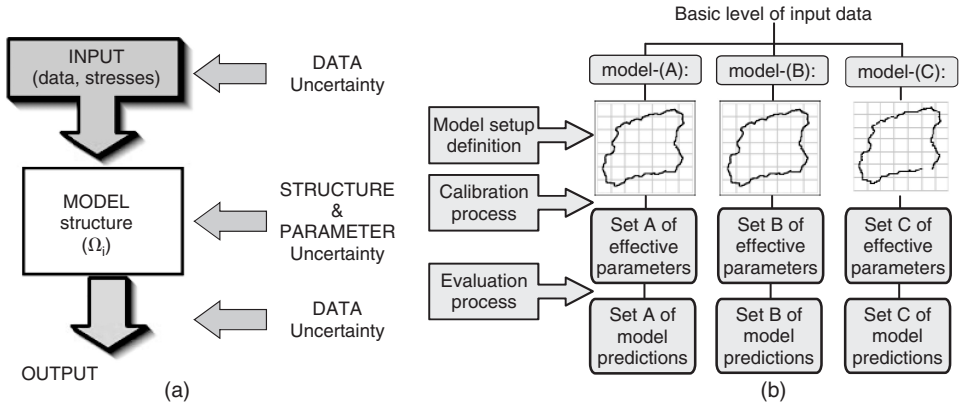


Figure 4. (a) Schematic of the relationship between the model system and the main sources of uncertainty and error in hydrologic modelling (Vázquez 2003); and flowchart of the Multi-calibration (MC) sensitivity analysis (Vázquez 2003 after Vázquez et al. 2002a).

When the model is unable to produce adequate results, an error analysis can possibly reveal the causes of the model failure. The following six sources of error can be distinguished (Fig. 4a, Feyen 2000): (*i*) random and/or systematic errors in the input and evaluation data; (*ii*) errors as a result of a non optimal parameter set; (*iii*) mathematical errors in the model, in relation to the physical and mathematical basis of the computer code; (*iv*) conceptual errors in the model, e.g., incorrect interpretation of the processes, boundary conditions and dimensionality of the system; (*v*) numerical errors inherent in the solution algorithm; and (*vi*) interpretation errors of the predicted results, especially in models lacking post-processing tools and misconception of expected results.

During the calibration process of a hydrological model, only the uncertainty due to the second source of error can be minimised. Therefore, it is important to distinguish between the various sources of uncertainty so as not to violate physical constraints when compensating for other error sources. Calibration can be accomplished with many different techniques, e.g., manually through 'trial and error', automatically through numerical parameter optimisation, or a combination of both techniques. Trial and error requires hydrological insight and experience and, due to its subjective character, it is likely that different model users will find different parameter sets for the same model area subject to identical stresses. In the automatic optimisation process, a mathematical algorithm is used to optimise a given objective function (Madsen 2003). To calibrate the parameters of a model by optimisation of an objective function means minimising the residuals, adopting some statistical technique. A major disadvantage of automatic calibration approaches is that those methods do not take account of the physical meaning of the model parameters.

3.2 *Assessment of model performance*

Model performance in the calibration and evaluation processes can be assessed both qualitatively, based on the graphical analysis of different time series properties, and quantitatively, based on a set of multi-objective statistical measures. In this study, both methods were combined in the context of a multi-criteria protocol (Vázquez et al., unpubl.). The inspected time series simulation properties included: (*i*) cumulative flow volumes; (*ii*) modelled versus observed discharge maxima and minima; and (*iii*) high flow extreme value statistics (Vázquez & Feyen 2007). Unrelated or slightly correlated statistics were preferred for the set of multi-objective statistics. One of these

statistics is the Coefficient of Efficiency (EF_2) (Legates & McCabe 1999), that is frequently used for an estimation of the total (combined systematic and random) average error:

$$EF_2 = 1 - \frac{\sum_{i=1}^{n} (O_i - P_i)^2}{\sum_{i=1}^{n} (O_i - \overline{O})^2} \quad (1)$$

where, P_i is the i-th simulated value, O_i is the i-th observed value, $\overline{O}$ is the average of the observed values and n is the number of observations. The optimal value of EF_2 is 1.0 and the feasible range of variation is $-\infty < EF_2 \leq 1.0$.

The EF_2 index is oversensitive to peak values because of larger random errors associated with higher discharges (Legates & McCabe 1999, Vázquez 2003). As such, it tends to give more weight to errors at high flow conditions. Thus, for a more even assessment of the simulation of both high and low flows, the multi-objective set of statistics was also applied on the logarithmic transformation of the observed and the predicted variables (Vázquez 2003).

4 STUDY CASES

4.1 *Case 1: Assessment on the feasibility of associating a physical meaning to the parameters of a conceptual model*

4.1.1 *Introduction*

The parameters of conceptual rainfall-runoff models usually do not account for physical properties of the basins and, in consequence, they have to be assessed by model calibration. An alternative approach for the determination of parameter values is to look for a relationship between model parameter (MP) values and physical catchment descriptors (PCD). Given that a relation exists between the PCD and the MP, the great advantage of such an approach is that ungauged catchments can be analyzed as well with this sort of models.

The objective of the present research was to investigate on whether the set of PCD for the four medium size catchments depicted in Fig. 4a could be directly related to the NAM model parameters (NMP). The eventual existence of such relationships would allow using the NAM model for simulating runoff in ungauged catchments located within similar regions and would make it possible to address the physical basis of the model as relationships could be expected from physical reasoning. The study sites, having surface areas ranging between 300 and 600 km^2, were chosen due to their variability in terms of catchment properties, representing different topographic and geomorphologic conditions (Vázquez 2003). The methodology involved three distinct stages: (i) a set of PCD lumped to the catchment scale was assembled using GIS; (ii) the calibration of the NAM model was performed for every basin and (iii) the relationships between NMP and PCD were derived and evaluated.

4.1.2 *Physical catchment descriptors (PCD)*

Description of the physical features of catchments was thereby in general based on morphological indices, percentage coverage of soil types, percentage coverage of land use types and climatic indices. Land use was classified in five groups using the LANDSAT satellite imagery available in digital format from the CO-oRdination of INformation on the Environment (CORINE) project (EEA, 1996) via the European Environment Agency (EEA), but using only the first level of the classification. Soils were divided into five groups as well, according to the Belgian textural classification (Vázquez 2003). Mean annual rainfall was chosen as the main climate index. In this way a set of 27 descriptors was assembled.

4.1.3 *Model parameterisation*

Rainfall-runoff processes were analyzed on a daily basis. The Thiessen polygon method was used for the calculation of the lumped precipitation. Potential evapotranspiration (ET_p) was estimated as described in Feyen et al. (2000) and Vázquez & Feyen (2002a). Digital terrain models (DTM) were produced from point elevation data. Firstly a triangular irregular network (TIN) was produced and from it a grid (LATTICE) was created (Vázquez & Feyen 2007). Digital maps of the river network, the land use and the soil type were also used in the analysis.

4.1.4 *Model calibration*

As a result of data constraints, model calibration was performed using daily values of rainfall, evapotranspiration and runoff. Every study site was modelled as a single lumped unit so no channel routing calibration was performed. Calibration was based on a split-sample test (Refsgaard 1997). Two major steps were taken, namely (i) automatic and (ii) manual calibration. The automatic calibration routine was carried out as a first and fast parameter screening prior the manual (more detailed) calibration. It was based on a multi-objective optimization strategy in which four different objectives were optimized simultaneously. The routine uses the overall volume residual (agreement between the average simulated and observed runoff), the overall root mean square error (RMSE) that depicts the agreement of the shape of the hydrograph, the average RMSE of peak flow events and the RMSE of low flow events (DHI, 1999).

In the manual (stepwise) calibration, a trial and error parameter adjustment was made until satisfactory results were obtained. The first step in the stepwise-calibration was to adjust the overall water balance, by adjusting the balance between the evapotranspiration (ET) and the difference between net precipitation and the runoff. The second step was to adjust the hydrograph shape of the peak runoff events. Peak volume adjustment, controlling the overland flow runoff coefficient (CQ_{OF}) and the time constant in the runoff routing ($CK_{1,2}$) accomplished this, respectively. This step was mainly applied to winter peaks, as winter peaks are much larger and more representative than summer peaks. The third step was to adjust the baseflow, mainly using the baseflow time constant (CK_{BF}). If the baseflow recession changes to a slower recession after a certain time, a lower groundwater reservoir should be added, including calibration of the recharge to lower groundwater storage CQ_{low} and the time constant for routing lower baseflow CK_{low} (DHI, 1999).

Ten NMP out of the maximum set of 15 parameters were used since the irrigation and snowmelt components were not considered. Model calibration was performed using hydrological time series of 7 to 12 years. The EF_2 coefficient was considered for the stepwise calibration, as in earlier research (Feyen et al. 2000, Vázquez et al. 2002a).

4.1.5 *Derivation of NMP/PCD relationships*

In this study the multiple regression analysis technique was used for estimating the hydrologic behaviour from physiographic and climatic catchment attributes. Data were collated for a dependent variable and one or more independent or explanatory variables and statistical calculations were used to derive an equation relating variations in the dependent variable to values of the independent variables (Beven, 2001).

4.1.6 *Results and discussion*

A good agreement between observed and simulated discharges was reached. For calibration, EF_2 values higher than 0.76 were obtained, while for the evaluation period, the EF_2 values were higher than 0.73.

Of the 10 NMP, 9 were found to be statistically independent, as the analysis showed only a significant correlation between parameters T_{OF} and CQ_{OF} with a determination coefficient (r^2) large enough (0.9978) to support a linear relationship. Each NMP was analyzed not only from a statistical point of view but also from a physical one, attempting to give a hydrologic meaning to each of them. Only 12 PCD exhibited independence out of a total of 27 considered, since the remaining PCD were determined using the same catchment characteristic or a previously calculated PCD. The relationships between the retained 9 NMP and 12 PCD were studied applying a multiple linear

regression technique. However, the small number of catchments made it impossible to perform a multiple regression analysis with more than two PCD as independent variables in order to avoid leaving zero degrees of freedom. It was found high correlation coefficients between the NMP and the PCD. Only three relationships (U_{max}, $CK_{1,2}$ and C_{area}) were found statistically significant (at the 5% level when performing F statistics), and the drainage density (D) was found to be the most important PCD, being present in four relationships.

The authors recommend validating the derived relationships before using them in a predictive context, in particular because none of the relationships had a physical basis. Nevertheless, the study can be considered as a first approach to the estimation of the NMP of medium sized Belgian catchments based on PCD.

4.2 *Case 2: Identification of the appropriate modelling grid-size for a distributed model applied on a catchment scale and analysis of the role of grid-size on the modelling*

4.2.1 *Introduction*

The problem of scale is not new for hydrologists. Research is nowadays, as in the past directed towards the problems of sub-catchment variability, sub-grid variability, macroscale modelling and the relationship between the modelling scale and the scale of the individual processes that are included in the catchment response (Refsgaard et al. 1999, Vázquez et al. 2002a).

MIKE SHE applied to a catchment scale implies the assumption that process descriptions, originally tested on smaller scales, are valid also to the larger scale. Thus, it performs an upscaling operation using effective parameter values. This is a scaling issue directly related to sub-grid variability of parameters that needs to be addressed through model calibration. Furthermore, different effective parameter values may be expected when the grid-size is shifted from one scale to another. An additional scaling issue may be identified when realising that, in catchment hydrologic modelling the modelling grid blocks usually have dimensions of decametres or even hectometres whilst vertical discretisations vary from few centimetres to few metres.

In general, increasing the level of discretisation increases the accuracy of the simulation, but there should be a level beyond which the model performance cannot be improved (Wood et al., 1988). Moreover, the smaller the grid-size in which a catchment is being subdivided using distributed models, the larger is the volume of catchment information that will be considered in the analysis, which means a tremendous increase in the work involved in data collection and processing. Reduction of the grid-size also means considerable more computational time.

This study was conducted to inspect how the spatial variability of catchment properties can be included in the modelling process and how this modelling process is affected by the grid-resolution (discretisation) in terms of the effective parameter values and of the modelling performance. The study was applied on the Gete catchment (Fig. 2b).

4.2.2 *Model parameterisation*

The model parameterisation common to both study sites (i.e. the Gete and Nete catchments) is briefly described in the following. Additional details on the properties of the catchments and on the characteristics of the distributed models of both catchments are available for instance at (Vázquez et al. 2002a, Vázquez & Feyen 2003, Vázquez 2003, Vázquez, unpubl.).

The definition of the Land Coverage (LC) is based on data from the CORINE project. The processing of this data set (CORINE-89) involved the transformation of the geographical coordinates (ETRS-89 system) to Cartesian (map) coordinates projected according to the Lambert Conformal Conic system for Belgium (L-72), resizing the original map resolution ($250 \times 411\ m^2$) according to the modelling resolutions and clipping the areas of interest in agreement with the modelling spatial domains.

The profile definition of the river tributaries was based on interpolation/extrapolation of a few measured profiles. Drains were specified in the model set-up to improve the simulated hydrograph shape and to account for the small canals and ditches present on a scale smaller than the modelling resolution.

The spatial extent of the soil units and their vertical properties could be assessed considering two soil databases AARDEWERK-BIS and AARDEWERK-SISBIS (Van Orshoven et al. 1991). Parameters for describing the flow through the soil system were calculated with pedo-transfer functions (PTFs) for the van Genuchten and the Averjanov equation (van Genuchten 1980, Vogel et al. 2000, Feyen et al. 2000, Vázquez 2003). Despite the uncertainties associated with the PTFs and the soil databases, the soil parameters were not included in the calibration process mainly to avoid considering too many parameters during model calibration (several parameters for every one of the soil units), avoiding in this way incurring in model overparameterisation.

The input time step, during which no change in boundary conditions occurs (stress period), was taken as one day for both catchments, in recognition of the lack of more precise meteorological data. While the model simulation time step is generally much smaller, this would affect the simulation of some sub-daily processes of the hydrological cycle such as Hortonian overland flow. It is important therefore to evaluate the quality of the flow-component (model) predictions.

After the consideration of the composition of the layers and the pumping activity in the catchments (Feyen et al. 2000, Vázquez 2003), no-confining beds were identified in either catchment. Thus, all the upper most aquifers containing the water table were modelled as unconfined, since MIKE SHE uses the Specific yield (S_y) in the cells that contain the water table. In the cells below the water table, the Specific storage (S_s) is used. The complexity of the geologic system of both catchments indicated that a three-dimensional (3-D) groundwater model was necessary for simulating the flows and water tables for both catchments. Despite the acceptable knowledge on the conceptualisation of the lithostratigraphy of both catchments, it is likely that the groundwater component of the distributed models of both catchments involves the most significant uncertainties.

Crop actual evapotranspiration (ET_{act}) was modelled using the Kristensen & Jensen (1975) approach. For doing so, MIKE SHE requires crop potential evapotranspiration (ET_p) data that were estimated by means of the K_c-ET_o method that uses crop coefficients (K_c) and crop reference potential evapotranspiration (ET_o). ET_o time series were in turn estimated with the Food and Agriculture Organisation (FAO) Penman-method 24 (FAO-24). Locally-applicable parameter values were available for the various components of the FAO-24 method (Vázquez & Feyen 2004). The estimates produced by the FAO-24 method, using these locally-applicable parameter values, are similar to the estimates of the FAO-56 Penman-Monteith method (Allen et al. 1998, Vázquez & Feyen 2004), obtained using standard parameter values recommended by the FAO-56 report for the various components of the method. No locally-applicable parameter values were available for the components of the FAO-56 method. Thus, the estimates generated with the FAO-24 Penman method using locally-applicable parameters were assumed being adequate for the current modelling, given their similitude with regard to the estimates generated with the standard FAO-56 Penman-Monteith method.

In this specific study applied on the Gete catchment, the effective values for the parameters that MIKE SHE uses to estimate ET_{act} (Kristensen & Jensen 1975) were assessed from literature (DHI 1999). These parameter values were kept constant throughout the modelling analysis. The river network was assumed to run along the boundaries of the grid-squares whose resolution therefore, determines the detail of the river network in the model set-up.

4.2.3 *Model calibration and evaluation*

The calibration of the Gete model was based on a trial and error optimisation process (Fig. 5). The calibration window was chosen to span the period from 1 January 1985 until 31 December 1986. A one-year initialisation or warming up period was considered as a way of attenuating the effects of the assumed initial conditions on the transient simulation. Statistical analysis was limited to 1985–1986. The validation period was chosen as 1 January 1987–31 December 1988. Depending also on the availability of observed data, additional evaluation periods of variable length were considered in the scope of a Multi-Window (MW) test.

In the first phase, the model was calibrated against the overall discharge of the catchment. Next, the influence of some hydro-geologic parameters on the simulated discharge was examined. The analysis showed that the hydraulic conductivity of the first, Quaternarian clayey loam unit, and the third, Landeniaan clayey sand unit, influenced significantly the simulated discharge. When the

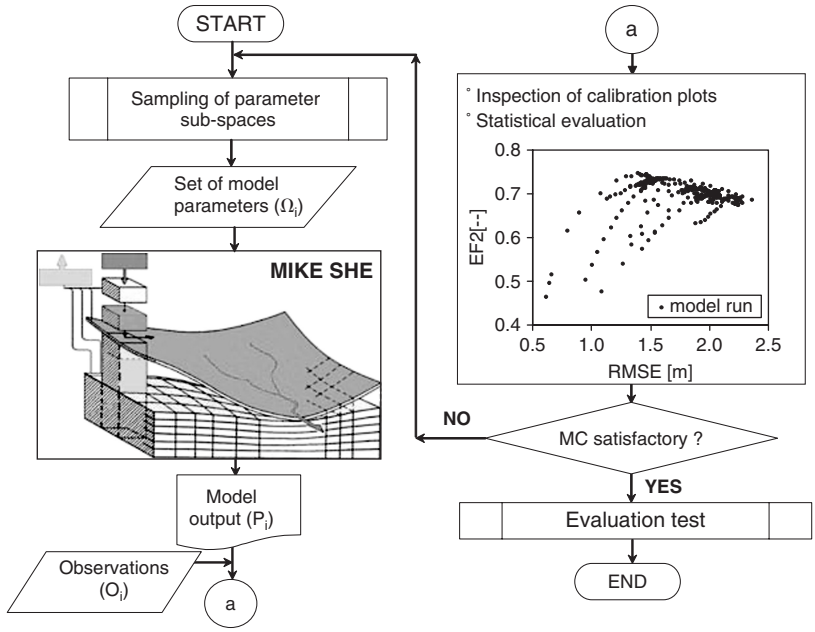


Figure 5. Flowchart of the calibration protocol (Vázquez 2003, after Vázquez & Feyen 2002a).

overall discharge was reasonably well simulated, the hydro-geologic parameters were optimised with the aim to achieve a better agreement between the simulated and observed piezometric levels in the observation wells considered in the calibration process.

In general terms, the model was calibrated and evaluated using a Split Sample (SS) procedure, against both daily catchment discharge measurements and observed water levels for 12 observation wells, with filters in different geologic layers. To investigate how well the calibrated model was able to simulate internal variables, a Multi-Site (MS) validation test was also performed for 2 internal discharge stations and 6 observation wells that were not considered during calibration (Feyen et al. 2000, Vázquez et al. 2002a).

The key scaling issue investigated in the present study deals with the effect of the grid-size on both the effective model parameters and on model performance. To support the analysis of this investigation two tests were implemented, namely a Multi-Resolution (MR) validation test and a Multi-Calibration (MC) test (Fig. 4b). In response to a pragmatic need for reasonable simulation times, as a starting step in the process, an initial model set-up was built-up with a 600 m resolution for the Gete catchment. First the 600-m grid-square model was calibrated (Feyen et al. 2000). This was then subjected to the MR evaluation test by using the effective parameters of the 600 m calibrated model in a 300-m and a 1200-m grid-square model.

In the MC test (Fig. 4b), the calibration and evaluation processes were kept as unique and standard as possible for every grid-size (i.e. 300 m, 600 m and 1200 m). Thus, the three models were subjected to identical calibration and evaluation processes for assessing the effect of grid-size on both the calibrated effective model parameters and the model performance (Fig. 4b) and determining the grid-size that yielded the best simulation results under (i) the given type and quality of input data, (ii) the MIKE SHE model structure and (iii) the lumped daily time step.

The drainage depth (z_{dr}) and the reciprocal time constant (T_{dr}), a sort of drainage coefficient, were calibrated. T_{dr} determines the velocity of the drainage and mostly influences the peak of the hydrograph; z_{dr} has more influence on the recession of the hydrograph. A maximum value of about 2 m below the ground level was considered during calibration. The feasible intervals for calibration of z_{dr} and T_{dr} for the models of both study sites were fixed on the basis of previous research and

the analysis of the shape of the hydrographs. The hydrogeologic parameters were also included in the calibration. The physical ranges considered in the calibration of the hydrogeologic parameters were derived from the literature considering comparable geologic formations.

4.2.4 *Results and discussion*

The calibration analysis showed that the simulated surface discharge and baseflow were significantly affected by the hydraulic conductivity of the Quaternarian and Landeniaan units. The hydrogeologic properties of the other units influenced mainly the baseflow. The MS test showed that for internal discharge stations (Fig. 2b) the model performance is inferior to the performance of the basin-wide station considered during calibration.

The MR test indicated that the model predictions for the analysed resolutions only differ marginally. The best evaluation results, in terms of river discharge, were obtained with a 600-m grid-resolution, independent of the stream-flow station. This, together with the exponential increase in computation time when reducing the grid-size, indicates that for the given level of data input and quality, the model type and structure, and the daily time step (i.e. stress period), a 600-m grid-resolution is most appropriate for the Gete catchment.

4.3 *Case 3: Identification of the appropriate ET_p data for the distributed modelling of a medium size catchment and analysis of the role of the ET_p data on the modelling*

4.3.1 *Introduction*

ET_{act} plays an important role in the hydrology of a catchment, together with other processes such as precipitation, runoff and infiltration. ET can be estimated directly from climatic measurements usually through energy balance methods, such as Penman type models that combine both energy and atmospheric vapour transport components. However, the effective quantitative estimation of ET still remains a difficult task to achieve, especially in a regional context, as it depends on the distribution of water supply to the evaporating surfaces (i.e. crops and soils).

Proper modelling of ET_{act} depends on the selection of an appropriate method for reducing ET_p and on the availability of realistic measurements or estimates of ET_p data. In MIKE SHE the estimation of ET_{act} is done on the basis of the Kristensen and Jensen model (Kristensen & Jensen 1975) that uses information on land use, crop characteristics, meteorological data and ET_p. The performance of many techniques for estimating ET_{act} or its components has been reported extensively in the literature (among others Wilson et al. 2001). However, there are few reports in the literature either about the performance of the MIKE SHE ET_{act} model, or about the influence of the ET_p data on the predictions of the model. An analysis of the Kristensen and Jensen model reveals that the ET_{act} components depend directly on the ET_p data. Therefore, it is important to investigate the relevance of ET_p on the model performance of the MIKE SHE code applied to a catchment scale. Thus, this study explores:

- the suitability of two Penman based estimating approaches, FAO-24 (Doorenbos & Pruitt, 1977) and FAO-56 (Allen et al. 1998) for deriving ET_p data to a point scale;
- the effect of the different point scale ET_p generation methods on the performance of the MIKE SHE code applied to a catchment scale, and;
- the dependency of the main model effective parameters on the ET_p estimation methods.

4.3.2 *ET_p generation methods*

To accomplish the investigation of the key issues of the present study, a MC test was implemented assembling different models as a function of the ET_p input data (c.f. Fig. 4b). The ET_p was estimated as a function of the reference evapotranspiration (ET_o) by means of a crop coefficient approach (c.f. section 4.2.2). Owing to the lack of more precise information, literature-derived mean K_c values were adopted (Doorenbos & Pruitt 1977, Allen et al. 1998).

Three methods for deriving ET_o estimates were used in modelling the Gete catchment: (A) Food and Agriculture Organisation (FAO) FAO-24 standard approach (Doorenbos & Pruitt, 1977);

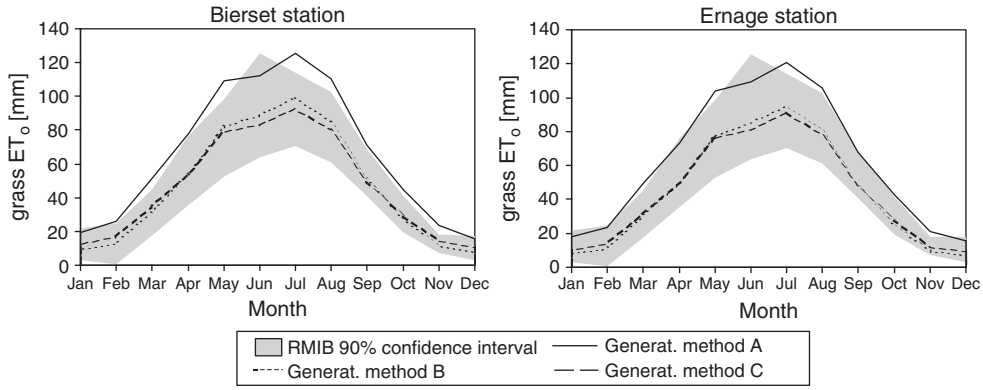


Figure 6. Mean ET_o estimates ($p[x \leq X] = 50\%$) for two meteorological stations used in the modelling compared to the RMIB 90% confidence intervals (ET_o guidelines) in the period 1967–1995 (Vázquez & Feyen 2003,2004).

(B) FAO-24 approach but used with coefficients, for both the wind and the Stefan-Boltzmann equations, differing from the standard formulation (Vázquez & Feyen 2003); and (C) FAO-56 standard approach (Allen et al. 1998). Daily data from the Royal Meteorological Institute of Belgium (RMIB) corresponding to two stations around the catchment were used in the period from the 1 January 1967 until the 31 December 1995 (10591 days). Thiessen polygons were utilised to extrapolate the point scale ET_p estimates for the referred stations to a catchment scale.

For assessing the performance of the estimation approaches, the point scale ET_o outputs of the generation methods were compared to local point scale ET_o guidelines (90% confidence intervals) derived from previous research (Fig. 6, Gellens-Meulenberghs & Gellens 1992). The generation method A significantly exceeded the ET_o guidelines (200 mm y^{-1} on average). The outputs of generation methods B and C showed a close agreement with the guidelines, suggesting that either of these data sets are more appropriate than data set A.

4.3.3 *Model parameterisation*

Model parameterisation is similar to what was already described in section 4.2.2. A 600-m modelling grid-size was used.

4.3.4 *Model calibration and evaluation*

Model calibration and evaluation were based on the SS, MC, MS and MW tests described in section 4.2.3. The parameters that were included in the calibration are z_{dr}, T_{dr}, the hydrogeologic parameters of the Quaternarian and Landeniaan layers and the parameters of the Kristensen & Jensen module for estimating ET_{act}. The calibration strategy is similar to the one described in section 4.2.3. The calibration ranges for the parameters of the Kristensen & Jensen module were assessed from literature (Kristensen & Jensen 1975, DHI 1999, Vázquez & Feyen 2003). The calibration and evaluation periods are the same as the ones considered in the previous section 3.4.

4.3.5 *Results and discussion*

The research revealed a significant dependency of the parameters related to the computation of ET_{act} on the different ET_p estimates. The respective dependency of the other inspected parameters is only marginal. This parameter sensitivity seemed to be caused mainly by the fact that MIKE SHE estimates ET_{act} as a function of the ET_p. In general, the ET_p generation methods seemed to have a greater impact on the streamflow model performance than on the model parameters. The dependency of the streamflow model performance on the ET_p generation methods seemed to be greater than the corresponding dependency in terms of predicted piezometric levels.

The effect of the ET_o estimates generated by methods B and C on the streamflow performance of the hydrological model seemed to be comparable. The best model performance was obtained by using the higher ET_o data generated by method A. This seemed to be a consequence of the interactions among the model parameters and the type and quality of input data in giving good fits to the observed time series. Furthermore, this suggest that higher ET_p estimates were needed to reflect that to a catchment scale some grid squares needed to have wetter characteristics that would produce evapotranspiration higher than predicted to a point scale (Thiessen polygons) by methods B and C. As a consequence, for methods B and C less water was released from the system in the form of vapour flux, which provoked an overestimation of the peak flows as revealed by the inspection of the predictions.

4.4 *Case 4: Incorporation of flow-component data into a catchment scale modelling in order to enhance the physical consistency of the predictions of the distributed model*

Although several multi-criteria protocols have been proposed in the literature to assess the performance of distributed models (Refsgaard 1997, Güntner et al. 1999), it has rarely been suggested an approach for evaluating the physical consistency of runoff predictions. This study suggests a multi-criteria evaluation protocol that includes an approach for evaluating the physical accuracy of runoff predictions during model calibration. This approach estimates flow-component data by partitioning the observed hydrograph and uses these data into model calibration. Furthermore, the protocol includes different statistical, analytical and visual criteria such as analysis of peak and low flows, cumulative volumes, extreme value statistics, performance statistics, etc. The assessment is applied on the modelling of the Nete catchment (Figs. 2a,c).

4.4.1 *Estimating flow-component data*

A flow-component hydrograph may be obtained by subtracting a time series of filtering flow from the total hydrograph (Willems 2000). This filtering time series can in turn be generated using the approach whereby the filtering flow-component, at every time step, evolves proportionally to the total discharge, as given by:

$$F_h(t) - a_{Ch}(t)F_h(t-1) = b_{Ch}(t)(Q(t) - c_{Ch}(t)Q(t-1)) \quad (2)$$

where $Q(t)$ = the ordinate of the total hydrograph (total signal) observed at time t [$L^3\ T^{-1}$]; $F_h(t)$ = the ordinate of the filtering signal (high frequency signal or filtering flow-component much quicker than the flow-component to be filtered) at time t [$L^3\ T^{-1}$]; $b_{Ch}(t)$ = the proportionality factor [–]; $a_{Ch}(t)$ and $c_{Ch}(t)$ = filter coefficients [–]. The filtered signal ($F_l(t)$) is obtained at time t after subtracting the filtering signal from the total flow, that is $F_l(t) = Q(t) - F_h\ (t)$.

The analysis revealed that the average flow-component contributions to the total discharge at the Varendonk station (Fig. 2c) are 65.2 % (Qbs), 14.3 % (Qov) and 20.5 % (Qin).

4.4.2 *Model parameterisation*

A 500-m resolution was considered. The effective values for the parameters that MIKE SHE uses to estimate ET_{act} were taken from a previous sensitivity analysis reported in section 3.5. These parameter values were kept constant throughout the modelling analysis.

4.4.3 *Model calibration and evaluation*

Model calibration and evaluation were based on the SS, MC, MS and MW tests described in section 4.2.3. To depict the advantages of the proposed multi-criteria protocol, the model was subjected to two calibration exercises. In the first exercise, the model was calibrated considering the observed basinwide streamflow (Q) as well as piezometric levels, but without making use of the estimated flow-component time series. In the second exercise, the suggested protocol was used considering this time the estimated (i.e. filtered) flow-component time series, namely overland (Qov), interflow (Qin) and baseflow (Qbs).

The main calibration and evaluation periods were chosen as 1 January 1986 – 31 December 1990 and 1 January 1991 – 31 December 1993. A one-year warming up period was run prior the calibration period. Calibration was carried out using a trial and error optimisation. The calibration strategy is the same as the one described in section 4.2.3.

The parameters that were included in the first calibration exercise are z_{dr}, T_{dr} and the hydrogeologic parameters of the geologic layers. For the second calibration exercise, after examining the physical consistency of the predictions of the first optimal model, two additional parameters were included in the calibration process, namely the leakage coefficient of the bed materials ($L_{Riv/GW}$) [T^{-1}] and the leakage coefficient of the ground surface materials ($L_{OV/GW}$) [T^{-1}]. These parameters are important for modelling the interaction between groundwater and surface water (both river and overland flow). This was done in an attempt to improve baseflow predictions.

4.4.4 *Results and discussion*

Besides total flow predictions (Q*), MIKE SHE can produce simulation results about the total inflows to the river system from overland (Qov*), drainage (Qdr*) and aquifers (Qbs*). The flow-component performance analysis was in general focused on qualitative aspects (i.e. visual inspection of calibration plots) because the filtered estimates cannot be used directly as observed data.

Calibration conditioned only on series of hydrometric data downstream the catchment and piezometric well levels in two well locations (Fig. 2c) produced physically inaccurate predictions (Fig. 7). Only after model refinement and the use in the calibration process of the estimated flow-component data, the physical consistency of the runoff predictions was enhanced significantly (Fig. 8). Hereby, based on the distributed modelling of the Nete catchment, it was shown that the application of the suggested protocol and in particular the use of the filtered flow-components in model calibration enhance the physical consistency of the model predictions, adding as such a considerable value to the calibration process.

4.5 *Case 5: Assessing the uncertainty attached to distributed hydrologic modelling*

4.5.1 *Introduction*

Hydrological models are usually applied under stresses and catchment conditions outside of the range of conditions for which the model was calibrated and initially evaluated. Currently available distributed models have limitations that are related mainly to (i) representation of hydrological processes at different scales; (ii) representation of boundary conditions and distributed patterns of model inputs; and (iii) identification of model parameters at different scales lacking both information and measuring techniques appropriate at different scales. In this context, the goal of attaining optimum parameter sets has been questioned, especially considering the nonlinearities and the parameter interactions in most hydrological models. Such difficulties also mean that it is effectively impossible to separate out the effects of all the different types of uncertainty in the modelling process (Fig. 4a).

It follows that the uncertainty associated with the data may arise from measuring or estimating uncertainties and from spatial/temporal scaling uncertainties, and that physically-based distributed catchment models will require some degree of calibration to determine effective parameter values at the modelling scale. Model structure uncertainty arises in that even the most complex distributed hydrological models are only approximate representations of the physical processes occurring in the catchment system.

In common applications only the uncertainty associated with the estimation of effective parameter values, can be minimised. However, the question is how to do so effectively for distributed models that generally have long run times, and where sets of effective parameter values that give acceptable fits to the calibration data might be scattered widely in the parameter space. In this context, the use of uncertainty estimation procedures based on Monte Carlo simulations, implying a large number of runs of a given model with different sets of parameter values sampled from the parameter space, is still rarely reported in literature for fully distributed models. In this study the performance of the fully PhB distributed MIKE SHE code, is explored in an application to the Gete catchment. The

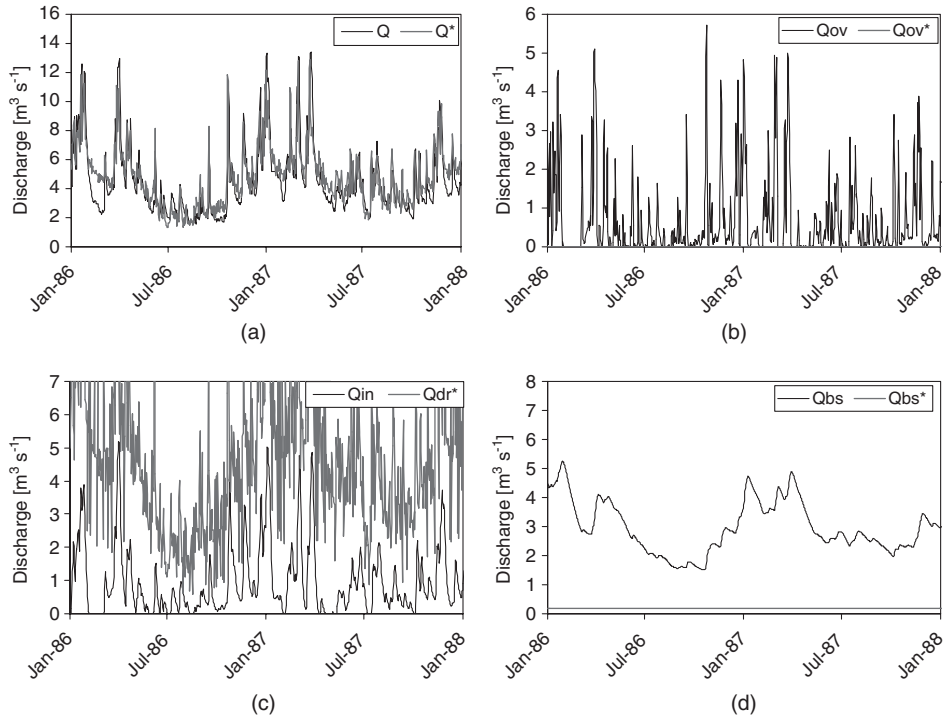


Figure 7. Predictions of the first optimal model in the calibration period.

uncertainty analysis has been carried out using the Generalised Likelihood Uncertainty Estimation (GLUE) methodology, based on the premise of equifinality of model structures and parameter sets.

4.5.2 *Model parameterisation*

The digital terrain model (DTM) was defined by processing available point elevation data by means of the Bilinear (Bi) interpolation method available in the MIKE SHE code (see section 3.6.2). The grid-size used in the analysis is 600 m, which has to be considered a coarse discretisation for describing accurately hillslope processes. However, the number of grid elements in the computational domain (1629) is larger than some previous catchment applications of the MIKE SHE code (Jain et al. 1992, Xevi et al. 1997, Christiaens & Feyen 2002) and other distributed modelling activities. This aspect in conjunction with the complex vertical description of the catchment (six geologic layers) led to a large computational time associated with every model run. As a consequence, a short six-months calibration period 1 March 1985 – 31 August 1985 was chosen. This period was preceded by a six-months spin-up period for attenuating the effects of the initial conditions that were kept the same for every one of the Monte Carlo simulations in the application of the GLUE methodology. A second period 1 September 1985 – 1 March 1986 was used for model evaluation. The simulations were made in a single run for each parameter set to ensure continuity of fluxes and internal state variables between the two periods.

The parameters that were included in the analysis are z_{dr}, T_{dr} and the hydrogeologic parameters of the Quaternarian and Landeniaan layers. These parameters were selected on the basis of prior analyses (Feyen et al. 2000, Vázquez et al 2002a, Vázquez & Feyen, 2003).

4.5.3 *The GLUE approach*

The fundamental principle of the philosophy supporting the GLUE approach is the rejection of the idea of an optimum parameter set in favour of the concept of "equifinality" or equivalence of

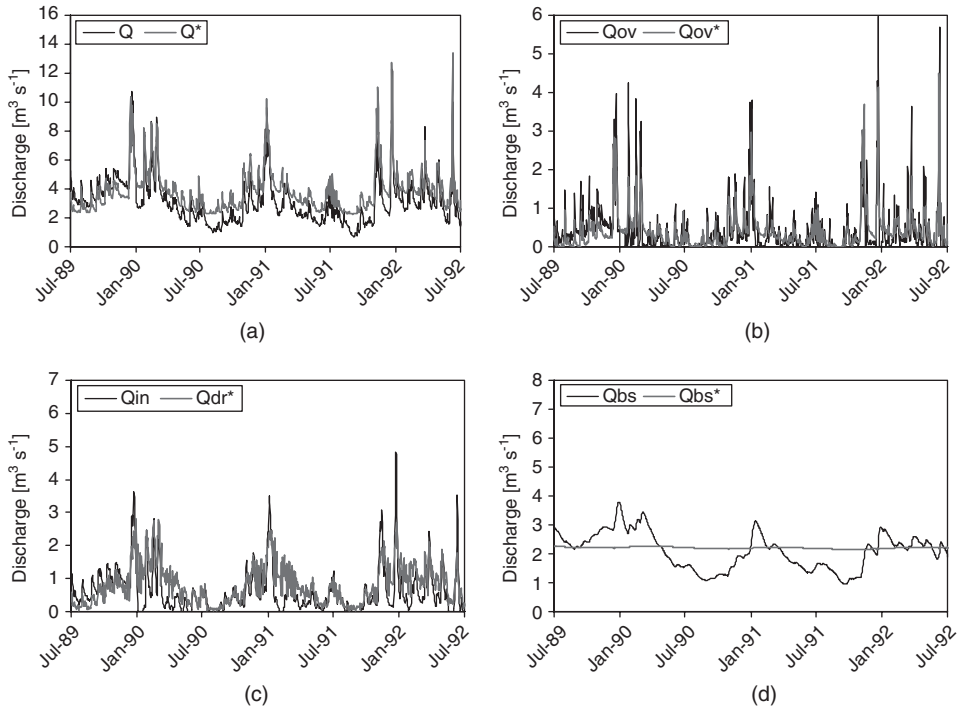


Figure 8. Predictions of the second optimal model in the evaluation period.

model structures and parameter sets due to the limitations in boundary conditions and observed data (Vázquez et al. 2002b). This principle implies that it is only possible to evaluate the relative performance of the models, associated with the multiple parameter sets, in terms of some likelihood measure.

Multiple simulations are run considering the multiple sets of parameter values, drawn according to their corresponding assumed prior probability distribution, in the framework of Monte Carlo simulations. The simulation results from the Monte Carlo realisations are then evaluated in terms of one or more performance measures. All the simulations (and their associated parameter sets) that exceed a specified threshold of performance are given a positive likelihood weight in prediction and are retained for further consideration. The parameter sets with likelihood weights below the threshold are considered to be non-behavioural and are given a prediction weight set to zero. If the likelihood weights of the behavioural sets are re-scaled such that the sum of all the likelihood weights equals 1, then the distribution functions of both the parameter values and predicted variables can be calculated. In this research the EF_2 index (Eq. 1) was used as a likelihood measure.

4.5.4 *Results and discussion*

The model had some problems to perform well as a continuous hydrograph simulator; it had especial difficulties to represent well the second main wet period and also the low flow during the months of September and October 1985 (Fig. 9). These seem to be the consequences of errors in boundary conditions and input data; model structure uncertainties and scale discrepancies between model structure and the modelling grid-size that possibly were magnified by the lumped daily time resolution included in the modelling. However, the results suggest that the model can predict streamflow within the limitations of the errors in the input variables.

The GLUE analysis showed that the horizontal and vertical components of the hydraulic conductivity of the Landeniaan layer have the most influence on the streamflow model performance in

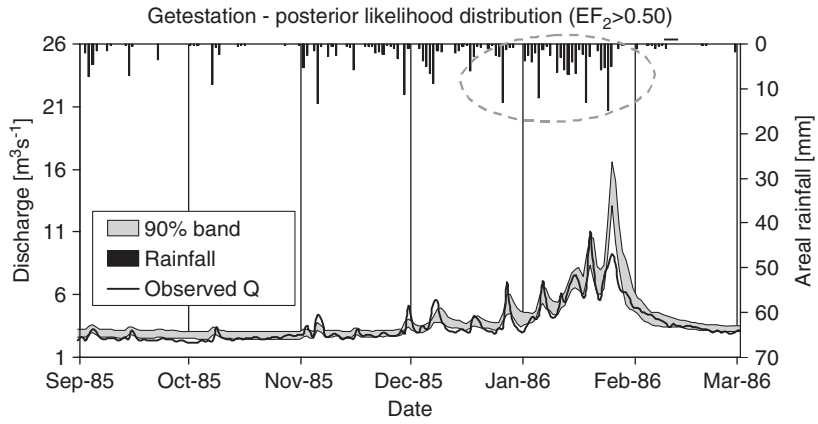


Figure 9. 90% streamflow uncertainty limits in the period [1 September 1985 – 1 March 1986] calculated using posterior likelihood distribution after conditioning based on observed streamflow in the period [1 September 1985 – 1 March 1986].

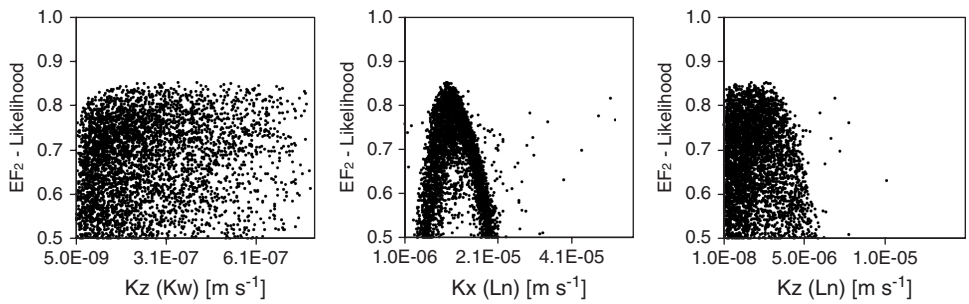


Figure 10. Scatter plots of behavioural parameter sets run through the period of analysis time series.

the application catchment (Fig. 10). The sensitivities of the other process parameters to the model performance are very low and there was no clear optimal parameter set evident in the sample of models run.

The prediction limits provided a measure of the total uncertainty in the model output (Fig. 9). The analysis depicted marked departures of the observed discharge from the predicted uncertainty bands, for some parts of the catchment-wide hydrograph, especially during the wet period (January and February 1986). The study revealed a considerable uncertainty attached to the simulation of both high flows and low flows (Fig. 9). Similarly, wide prediction intervals were obtained for the piezometric levels in relevant wells. In general, the GLUE analysis indicated that there is significant uncertainty associated with the present modelling resulting from several factors such as parameter and structure uncertainties, parameter interactions, errors in boundary conditions, data uncertainties and significant scale discrepancies between model structure and the modelling grid-size. Nevertheless, the results suggest that the model of the catchment predicts outputs within the limitations of the errors in the input variables and the discrepancies between the hydrologic code structure and the actual catchment processes.

5 GENERAL CONCLUSIONS

This study concentrated on the brief assessment of some topics that are relevant to the application of distributed models on a catchment scale. Specifically, the study explored on (i) how best to represent hydrological processes; (ii) how best to identify the parameters of distributed models at different scales; (iii) how best to incorporate new sources of data into the modelling process; and (iv) how to assess general uncertainty attached to the application of distributed hydrologic models on a catchment scale. These aspects were briefly addressed by means of 5 study cases.

Case study 1 explored on the feasibility of relating the parameters of a lumped conceptual model applied on a catchment scale to the physical properties of the modelled catchments. The study was based on the modelling of 4 catchments having different geomorphologic characteristics. The study showed that the low number of study sites contributed to obtain relationships without a physical sound meaning. Nevertheless, this study constituted the first approach to investigate on the feasibility of enhancing the way of parameterising conceptual models within a physical framework. Future work will involve the coordinated and conjunctive use of lumped conceptual models and PhB distributed models with the same purpose in mind.

Case 2 showed that grid-size plays an important role on both the effective values of the parameters as well as the predictions of distributed models. The study was based on the independent calibration of models differing only in the modelling grid-size, which enabled to select the best compromise between computing time and accuracy of predictions. Case study 3 identified ET_p as being also a very important modelling variable that plays a major role on the simulated water balance and as such on the model predictions. These cases in particular demonstrated the difficulty inherent in the process of assigning parameter values that are effective at the modelling scale.

Case study 4 showed the importance of the incorporation of data derived from non-conventional sources on enhancing the physical consistency of model predictions. The study revealed that by incorporating into model calibration flow-component data derived indirectly from observed basin-wide discharge the quality of runoff predictions to a catchment scale improved significantly. The methodology applied herein will allow the modeller to evaluate and calibrate the model in a more sound way. It will meet part of the problem of "over-parameterisation" and consequently the problem of "over-calibration" when detailed distributed hydrological models such as MIKE SHE are implemented for catchments with limited evaluation data.

Finally, Case study 5 showed that large uncertainties could be attached to the hydrological predictions as a result of sub-grid variability of parameters, data uncertainty, scale discrepancies (description of physical processes/data/modelling), errors in the definition of initial and boundary conditions, etc. It was found that the GLUE approach, despite the high number of required Monte Carlo simulations, is an attractive possibility for easily assessing model uncertainty.

Future work on distributed hydrological modelling is likely to continue on the subjects briefly covered in this study. In this context, questions that still remain to be answered in the future are for instance (Beven & Feyen 2002):

- How can process descriptions be improved without introducing more parameter identifiability problems?
- How can a model, with scale-dependent effective parameter values, be applied without calibration?
- How simple need a distributed model be, given the uncertainties in the modelling process?
- Are stochastic approaches needed either to assess the predictive uncertainty arising from model and parameter uncertainty or to search for the 'single realization' that is reality?
- Is there really an equifinality problem in the identification of distributed model structures, or will it be reduced or eliminated in future by improved observational techniques?
- Do we need flexible/incremental "à la carte" models for different modelling situations or environments?
- Is there an alternative approach to distributed modelling that could replace, rather than extend, conventional distributed modelling?

All these aspects will produce surely exciting research discussions in the future. In particular the last two aspects open a broad possibility of significantly improving distributed hydrologic modelling as suggested already by ongoing research. For instance, Branger (2007) successfully assembled some different modules that can be activated (manually or automatically) "à la carte" to undertake the modelling of processes that are more or less significant at different scales. This not only brings up the possibility of enhancing the physical representation of processes at different scales but also the feasibility of considering irregular computational geometries (Loudyi et al. 2007) differing from conventional regular geometries used for instance in MIKE SHE (Refsgaard and Storm 1995) and MODFLOW (McDonald & Harbaugh 1988), which in turn brings the possibility of incorporating into distributed modelling the concepts of Representative Elementary Area (REA, Wood et al. 1988) or the more elaborated and physically correct concept of Representative Elementary Watershed (REWs, Reggiani et al. 1998, Tian et al. 2006). Additional benefits from this is the potential of significantly reducing computational times since the number of computational elements would decrease notoriously and the ease of coupling land phase models with meteorological (i.e. "atmospheric") models that is currently difficult owing to scale constrains. Problems that still need to be resolved are related to numerical perturbations that might arise for certain geometric configurations of the computational elements (Loudyi et al. 2007).

ACKNOWLEDGEMENTS

This investigation was possible thanks to a research grant from the Katholieke Universiteit Leuven (project PDM/03/188) awarded to the first author. The completion of this article was achieved in the framework of the Research and Development contract of the first author funded by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA, Spain) and the Centro de Investigación y Tecnología Agroalimentaria de la Diputación General de Aragón (CITA-DGA, Spain).

REFERENCES

Allen, G.R., Pereira, L.S., Raes, D. & Martin, S. 1998. *Crop evapotranspiration-Guidelines for computing crop water requirements*. Rome: FAO.

Beven, K.J. 2001. Rainfall-Runoff Modelling: The Primer. Chichester: John Wiley and Sons.

Beven, K.J. & Feyen, J. 2002. The future of distributed modelling. *Hydrological Processes* 16: 169–172.

Beven, K.J. & Freer, J. 2001. A dynamic TOPMODEL. *Hydrological Processes* 15: 1993–2011.

Branger, F. 2007. Utilisation d'une plate-forme de modélisation environnementale pour représenter le rôle d'aménagements hydro-agricoles sur les flux d'eau et des pesticides. Application au bassin versant de la Fontaine du Theil (Ille-et-Vilaine). Doctoral dissertation. Lyon, France: Université Joseph Fourier.

Christiaens, K. & Feyen, J. 2002. Constraining soil hydraulic parameter and output uncertainty of the distributed hydrological MIKE SHE model using the GLUE framework. *Hydrological Processes* 16(2): 373–391.

Célleri, R., Timbe, L., Vázquez, R.F. & Feyen, J. 2003. Assessment of the relation between the NAM rainfall-runoff model parameters and the physical catchment properties. *HIP-VI UNESCO Technical Documents in Hydrology* 66: 9–16.

DHI 1999. *MIKE-SHE Water Movement User Manual Edition 1.2*. Denmark: Danish Hydraulic Institute.

Doorenbos, J. & Pruitt, W.O. 1977. *Crop water requirements*. Rome: FAO.

EEA 1996. *CD-ROM Natural Resources*. Copenhagen, Denmark: European Environment Agency.

Feyen, L., Vazquez, R.F., Christiaens, K., Sels O. & Feyen, J. 2000. Application of a distributed physically-based hydrological model to a medium size catchment. *Hydrology and Earth Systems Science* 4(1): 47–63.

Gellens-Meulenberghs, F. & Gellens, D. 1992. *L'evapotranspiration Potentielle en Belgique: variabilite spatiale et temporelle*. Brussels, Belgium: Belgian Royal Meteorological Institute (BRMI).

Güntner, A., Uhlembrook, S., Seibert, J. & Leibundgut, Ch. 1999. Multi-criterial validation of TOPMODEL in a mountainous catchment. *Hydrological Processes* 13: 1603–1620.

Havnø, K., Madsen, M.N., & Dørge, J. 1995. MIKE 11 A generalized modeling package. In V.J. Singh (ed.), *Computer Models of Watershed Hydrology*: 809–846. USA: Water resources publications.

Jain, S.K., Storm, B., Bathurst, J.C., Refsgaard, J.C. & Sing, R.D. 1992. Application of the SHE to catchments in India-Part 2: Field experiments and simulation studies on the Kolar subcatchment of the Narmada River. *Journal of Hydrology* 140: 25–47.

Kristensen, K.J. & Jensen, S.E. 1975. A model for estimating actual evapotranspiration from potential evapotranspiration. *Nordic Hydrology* 6: 170–188.

Legates, D.R. & McCabe, G.J. 1999. Evaluating the use of 'goodness-of-fit' measures in hydrological and hydroclimatic model validation. *Water Resources Research* 35(1): 233–241.

Loague, K. & Green, R.E. 1991. Statistical and graphical methods for evaluating solute transport models: overview and applications. *Journal of Hydrology* 7: 51–73.

Loudyi, D., Falconer, R.A. & Lin, B. 2007. Mathematical development and verification of a non-orthogonal finite volume model for groundwater flow applications. *Advances in Water Resources* 30: 29–42.

Madsen, H. 2003. Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives. *Advances in Water Resources* 26(2): 205–216.

McDonald, M.G. & Harbaugh, A.W. 1988. A modular three-dimensional finite-difference ground-water flow model. In USGS (ed.), *US Geological Survey (USGS) Techniques of Water-Resources Investigations*: 586 pp. USA: USGS.

Motovilov, Y.G., Gottschalk, L., Engeland, K. & Rodhe, A. 1999. Validation of a distributed hydrological model against spatial observations. *Agricultural Forest Meteorology* 98–99: 257–277.

Refsgaard, J.C. 1997. Parameterisation, calibration and validation of distributed hydrological models. *Journal of Hydrology* 198: 69–97.

Refsgaard, J.C. & Knudsen, J. 1996. Operational validation and intercomparison of different types of hydrological models. *Water Resources Research* 32: 2189–2202.

Refsgaard, J.C. & Storm, B. 1995. MIKE SHE. In V.P. Singh (ed.), *Computer Models of Watershed Hydrology*: 809–846. USA: Water Resources Publications.

Reggiani, P., Sivapalan, M. & Hassanizadeh, S.M. 1998. A unifying framework for watershed thermodynamics: balance equations for mass, momentum, energy and entropy, and the second law of thermodynamics. *Advances in Water Resources* 22(4): 367–598.

Saulnier, G.-M., Obled, C. & Beven, K. 1997. Analytical compensation between DTM grid resolution and effective values of saturated hydraulic conductivity within the TOPMODEL framework. *Hydrological Processes* 11: 1331–1346.

Tian, F., Hu, H., Lei, Z. & Sivapalan, M. 2006. Extension of the Representative Elementary Watershed approach for cold regions via explicit treatment of energy related processes. *Hydrology and Earth Systems Science* 10: 619–644.

van Genuchten, M.Th. 1980. A closed form equation for predicting the hydraulic conductivity in unsaturated soils. *Soil Sciences Society American Journal* 44: 892–898.

van Orshoven, J., Maes, J., Vereecken, H. & Feyen, J. 1991. A procedure for the statistical characterization of the units of the Belgian soil map. *Pedologie* XLI(3): 193–212.

Verhoest, N.E.C., Troch, P.A., Paniconi, C. & DeTroch, F.P. 1998. Mapping basin scale variability source areas from multitemporal remotely sensed observations of soil moisture behavior. *Water Resources Research* 34: 3235–3244.

Vogel, T., van Genuchten, M.Th. & Cislerova, M. 2000. Effect of the shape of the soil hydraulic functions near saturation on variably-saturated low predictions. *Advances in Water Resources* 24: 133–144.

Vázquez, R.F. 1997. Groundwater modelling of the Brusselean aquifer with GWVistas. MSc. Dissertation. Brussels, Belgium: Faculty of Engineering, Free University of Brussels (V.U. Brussel).

Vázquez, R.F. 2003. Assessment of the performance of physically based distributed codes simulating medium size hydrological systems. Doctoral dissertation. Leuven, Belgium: Faculty of Engineering, Katholieke Universiteit Leuven.

Vázquez, R.F. & Feyen, J. 2002. Assessment of the performance of a distributed code in relation to the ET_p estimates. *Water Resources Management* 16(4): 329–350.

Vázquez, R.F. & Feyen, J. 2003. Effect of potential evapotranspiration estimates on effective parameters and performance of the MIKE SHE-code applied to a medium-size catchment. *Journal of Hydrology* 270(4): 309–327.

Vázquez, R.F. & Feyen, J. 2004. Potential Evapotranspiration for the distributed modelling of Belgian basins. (ASCE) *Journal of Irrigation Drainage Engineering* 130(1): 1–8.

Vázquez, R.F. & Feyen, J. 2007. Assessment of the effects of DEM gridding on the predictions of basin runoff using MIKE SHE and a modelling resolution of 600 m. *Journal of Hydrology* 334: 73–87.

Vázquez, R.F., Feyen, L., Feyen, J. & Refsgaard, J.C. 2002a. Effect of grid-size on effective parameters and model performance of the MIKE SHE code applied to a medium sized catchment. *Hydrological Processes* 16(2): 355–372.

Vázquez R.F., Feyen, J., Beven, K. & Berlamont, J. 2002b. Preliminary results of an uncertainty analysis by using the MIKE SHE code. In Proceedings of the 5th International Conference on Hydro -Science & -Engineering, Warsaw: Warsaw University of Technology.

Vázquez, R.F., Willems, P. & Feyen, J. Unpubl. Improving the predictions of a MIKE SHE catchment-scale application by using a multi-criteria approach. *Hydrological Processes*.

Willems, P. 2000. Probabilistic immission modelling of receiving surface waters. Doctoral dissertation. Leuven, Belgium: Faculty of Engineering, Katholieke Universiteit Leuven.

Wilson, K.B., Hanson, P.J., Mulholland, P.J., Baldocchi, D.D. & Wullschlenger, S.D. 2001. A comparison of methods for determining forest evapotranspiration and its components: sap-flow, soil water budget, eddy covariance and catchment water balance. *Agricultural Forest Meteorology* 106: 153–168.

Wood, E.F., Sivapalan, M., Beven, K. & Band, L. 1988. Effects of Spatial Variability and Scale with Implications to Hydrologic Modelling. *Journal of Hydrology* 102: 29–47.

Xevi, E., Christiaens, K., Espino, A., Sewnandan, W., Mallants, D., Sorensen, H. & Feyen, J. 1997. Calibration, validation and sensitivity analysis of the MIKE-SHE model using the Neuenkirchen catchment as case study. *Water Resources Management* 11: 219–242.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Erosion and sediment transport processes in Mediterranean mountain basins

S. Beguería
Estación Experimental de Aula Dei – CSIC, Zaragoza, Spain

N. Lana-Renault, D. Regüés, E. Nadal-Romero, P. Serrano-Muela & J.M. García-Ruiz
Instituto Pirenaico de Ecología – CSIC, Zaragoza, Spain

ABSTRACT: Soil erosion and sediment production is a major environmental problem in the World, and especially in Mediterranean countries which are characterized by a harsh climate and a historical human pressure on the land. In a context of current and foreseen climate and land use change, our ability to predict the effects of these changes on soil erosion and on water quantity and quality become of great importance. A significant research effort has been made on the subject in the last two or three decades, based both in empirical observations from experimental field setups or in the use of simulation models. However, our understanding of some key processes is still poor, especially under a multi-scale perspective. This encourages the need for further empirical and modelling studies. In this paper, a review of the problem is made and some results are shown from studies in the Spanish Pyrenees. These results are followed by a discussion of several important research questions regarding the use of field experiments and simulation models.

1 INTRODUCTION

Soil erosion by water is currently considered one of the main environmental problems in sub-humid and semi-arid environments of the World, threatening the quality of life and the survival of millions of people. Erosion and desertification are responsible for large population movements in some countries, which cause political, social and economical instability at national and international scales (Puigdefábregas and Mendizábal, 1995). In the case of Mediterranean countries, soil erosion has been identified as a major problem (Poesen and Hooke, 1997), resulting in changes in soil characteristics, loss of productivity, reservoir siltation and changes in the quantity and quality of water resources (García-Ruiz and Valero-Garcés, 1998). Ultimately, soil erosion can cause a loss of biodiversity and welfare.

Obviously, most of these erosion-derived problems affect directly the areas where erosion occurs, but others, such as changes in fluvial channel dynamics, flood hazard and development of alluvial fans and deltas, can affect territories located far away. This is particularly the case of Mediterranean environments in which dramatic contrasts exist between the highlands and the lowlands, the former behaving as "islands" of humidity where most of water resources are generated, whereas most of the population and economic activities are located in the latter. Climate and land use change in these spatially restricted areas would disturb significantly the hydrological cycle, affecting water availability, flooding, and sediment transport in a much bigger area. For this reason, soil erosion can be considered a global problem, the causes and solutions of which being very complex and needing political and social implications, as well as deep changes in the perspective of land management.

Studies on soil erosion in Mediterranean countries underwent an important push during the 1980's, and many research groups and projects appeared particularly in Spain but also in

Israel, France and Italy and, to a lesser extent, in Greece, Turkey and the North of Africa. Nevertheless, some signs of stagnation are the more and more evident. The specialists are able to analyze in detail the processes, to deepen into the soil and rain drop characteristics, to improve the field and laboratory instruments, but our understanding of soil erosion as a global process is not improving: Our body of knowledge is basically the same than it was twenty years ago. Some errors from the past continue existing in the present (e.g., the use of erosion rates, the comparison of results from different methods or from different scales), and a holistic approach is increasingly needed to understand the spatial and temporal organization of soil erosion and sediment delivery problems. Time has shown the importance of both empirical and simulation approaches to the problem of erosion and sediment transport, but an effort needs to be made to link together the processes acting at various temporal and spatial scales. Only that would allow us to answer current questions such as the relative importance of climate and land use practices on soil erosion, or evaluate global change scenarios with the purpose of land planning.

2 EROSION AND SEDIMENT TRANSPORT AS A CHRONIC ENVIRONMENTAL PROBLEM IN THE MEDITERRANEAN

Soil erosion is a very old problem in the Mediterranean area. Geo-archaeological, palinological and sedimentological studies illustrate on the decrease in vegetation as human pressure increased. Geomorphic studies demonstrate that during the second half of the Holocene many hillslopes experienced accelerated erosion rates, and valleys were filled with sediment. This was the case, for example, in the Ebro Valley since the Neolithic (i.e., Peña et al., 1993). Erosion was so intense in the Ebro Valley in the last 4000 years that no in-situ remnants of pre-Bronze Age archaeological sites have been preserved (González-Sampériz and Sopena, 2002). In the Pyrenees, Montserrat (1992) studied a core of sediment in the Tramacastilla Lake which shows clear evidence of severe and generalized fires occurred 4000 and 900 years ago, followed immediately by increases in the sediment size and the rate of sedimentation—several orders of magnitude higher than before the fires. The development of the Ebro Delta since the Roman times is also a good example of the geological consequences of an increasing human pressure accompanied with deforestation, grazing and farming. The period of maximum deltaic progradation between the 15 and 19th centuries coincides with a general expansion of cereal crops and livestock in the Ebro basin (Maldonado, 1983). Nevertheless, in these examples and in many others in the Mediterranean it is still very difficult to discriminate the relative roles of human activities and climatic fluctuations. In a context of foreseen climate and land use changes, this is one of the main scientific problems we need to solve.

It is well known that soil erosion by water depends on the complex interaction of several factors, mainly climate, soil characteristics, bedrock, plant cover and land management. All these factors tend, in general, to encourage soil erosion in Mediterranean environments. Mediterranean climate is characterized by strong seasonal contrasts in temperature and precipitation, and by intense rainstorms in any season of the year, particularly in autumn and summer (White et al., 1997). Many soils tend to be fragile and easily erodible due to the poor organic matter content or to sealing and crusting processes caused by raindrop impact and aggregate breaking (Yaalon, 1997). This is especially evident over certain impervious bedrocks (marls, lutites), where regolith weathering is very rapid and intense as a consequence of freezing-thawing and wetting-drying processes (Regüés and Gallart, 2004; Nadal-Romero et al., 2007). Finally, land-use / plant cover changes have been defined as the main factor of desertification in Mediterranean environments (Kosmas et al., 1997), with higher influence on soil erosion than climate change, since they can occur in a few years or decades. For instance, the substitution of forests or dense scrubland by cereal fields represents: (i) an increase of rainfall that arrives directly to the soil, owing to the absence of interception; (ii) a direct impact of raindrops on the soil, thus favoring sealing, crusting and splash processes; (iii) an increase of overland flow and sheet wash processes; (iv) an expansion of runoff and sediment generating areas; (v) a change in the frequency and magnitude of floods, reducing the return period for the same flood event; and (vi) a rapid change in the hydrological functioning

of the soils (infiltration rate, water storage, timing of wetting-drying processes). Similar changes occur if the scrubland is open by overgrazing or by periodical burning.

The consequences of these changes are also well known: increase in soil stoniness, elimination of the upper soil horizons, development of rills and, in the worst cases, gullies and badlands. These processes occur in both relatively wet and dry environments (Central and Eastern Pyrenees, and Southeast Spain and the Central Ebro basin, respectively). Many alluvial fans have been considered as a secondary effect of increased soil erosion (Gómez-Villar et al., 2006) due to their recent origin, as well as aggradation of alluvial plains and the development of braided channels (Beguería et al., 2006). Finally, reservoir siltation is a very serious problem for water resources management, since it reduces the capacity for water storage and shortens the useful lifetime of the dams. Some reservoirs have a sediment trap capacity close to 100%, due to their large size in relation with the fluvial discharge (Almorox et al., 1994). The most important problems of reservoir siltation occur in semi-arid environments with occasional but very intense rainstorms, low plant cover density and presence of badlands in the basin, as it occurs in Southeast Spain. In general, badlands in marls and lutites are the main sediment sources (Beguería, 2005).

3 METHODOLOGICAL APPROACHES TO THE PROBLEM OF EROSION AND SEDIMENT TRANSPORT: MANY POINTS OF VIEW FOR A SINGLE, MULTI-SCALING PROBLEM

The processes of runoff generation, soil erosion and sediment delivery are characterized by a multi-scaling behavior, in which different process interplay at different spatial and temporal scales. Soil erosion and sedimentation have been studied by using a variety of methods and at several temporal and spatial scales. This is a huge problem when one tries to compare the information obtained from different sites. It is very important to understand that each method is related to certain research questions, which are necessarily different to those posed with other methods. Besides, in some cases, the use of certain methods is a consequence of the financial or personnel possibilities of a research group.

A rapid summary of methods or techniques used in soil erosion studies includes piquettes and micro-topographic profilers, rainfall simulation (less than 1 m^2, in most of cases), experimental plots (in general, lower than 100 m^2), experimental catchments (few tens to hundreds of hectares), estimation of sediment siltation in reservoirs (bathymetric approach), radioisotope inventories (particularly 137Cs) and mapping erosive features by aerial photo interpretation and remote sensing—in parallel with Geographic Information Systems. Each of these approaches faces the study of soil erosion at a given characteristic scale. It is well know that different erosion and sediment transport processes prevail at different scales, what adds to the difficulties of making general conclusions out of a broad range of methodologies: How is it possible to compare erosion rates obtained from rainfall simulation and those from experimental catchments? As a general rule, a progressive decrease in water and sediment yield is expected as we move from small to large study areas (Figure 1). Nevertheless, even this general principle is not necessarily true in some cases, as it has been shown in a characteristic catchment in the Spanish Central Pyrenees (Figure 2).

Furthermore, there is another important problem. Even if erosion rates should be obtained from the same scale and with the same methods, are we sure about their significance as average values? As the analysis of data obtained from experimental plots and catchments has shown, runoff and sediment transport rates exhibit large inter-annual fluctuations in the Mediterranean, as a consequence of the highly varying characteristics of precipitation and temperature. It is well known that most of soil erosion and transport tends to occur during a few, intense rainstorm events. For instance, in badlands of the Ebro Depression, Desir (2001) demonstrated that erosion was reduced by one order of magnitude when the study period was increased from only some months to nine years. Likewise, in the Izas experimental catchment (upper Gállego Valley, Central Pyrenees) one single rainstorm in October 1987 yielded as much sediment as the total amount recorded during the next decade. The conclusion is that information on erosion rates is strictly comparable only

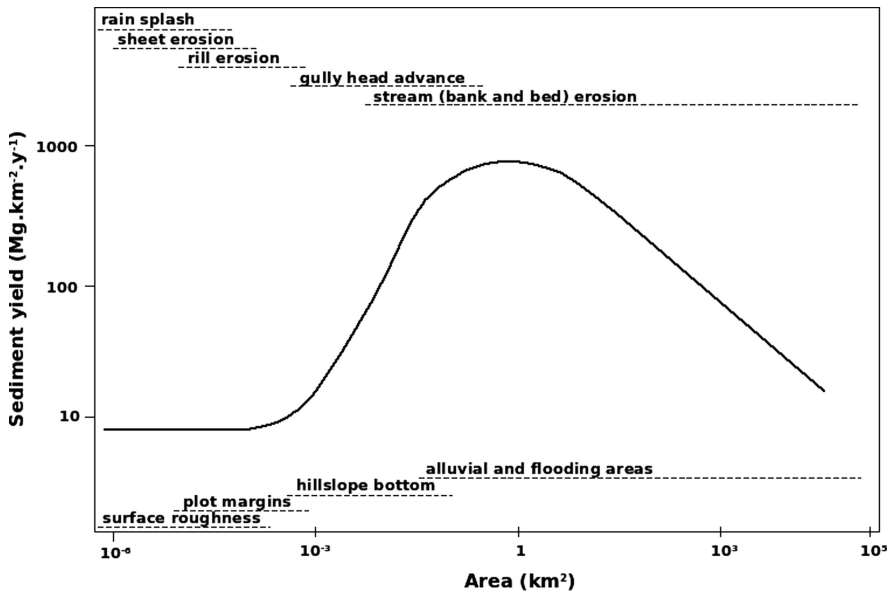


Figure 1. Variation of sediment yield with respect to the catchment size (conceptual sketch). Sediment yield at a given spatial scale depends on the interplay of several erosion processes, and different processes prevail at characteristic scales. Also, the role of the major sediment sinks (deposition areas) changes according to the size of the area of interest. The main erosion processes and sediment sinks are shown in the figure.

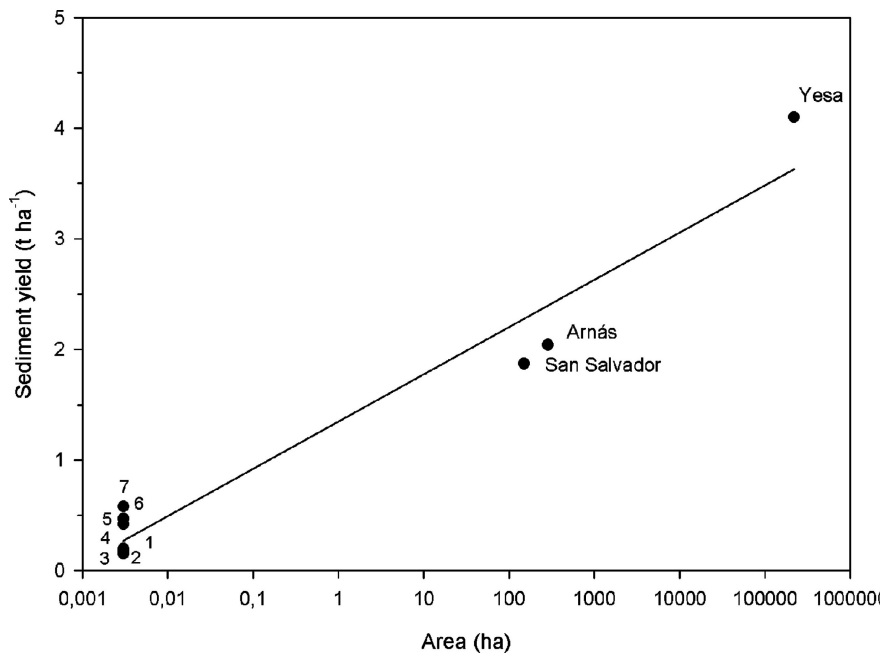


Figure 2. Relationship between catchment area and sediment delivery in the Central Spanish Pyrenees. 1–7: experimental plots in the Aísa Valley Experimental Station. In contradiction with the general law, results from the Central Spanish Pyrenees show a positive relationship between sediment delivery and the size of the catchment.

when the same study period is used, imposing a severe constraint to comparative studies gathering data from many different sources.

4 EROSION AND SEDIMENT TRANSPORT IN MEDITERRANEAN MOUNTAIN BASINS: MOST RELEVANT RESULTS FROM OBSERVATIONAL RESEARCH IN THE CENTRAL PYRENEES

Experimental catchments have been considered, in general, as the most adequate approach to study runoff generation, soil erosion and sediment transport (Walling, 1991). The catchments are usually chosen to be homogeneous in lithology, topography and plant cover and have a relatively small size, and all this factors favor the control of the main processes acting in the catchment. The catchments allow us to study the hydrological response during individual rainstorms and over long periods of time, the areas of active erosion, the connectivity between sediment sources and channels, and the role of different processes, such as the partitioning of precipitation between interception, infiltration and runoff due to plant cover. Two main research questions are the consequences on the hydrological and sediment behavior of changes in the climatic conditions and on the land use/cover. Mediterranean mountain areas are particularly sensitive to these changes due to: (i) a process of farmland abandonment after centuries of cultivation of south-facing slopes and the consequent conversion to scrubland and forests; and (ii) new scenarios of climate change, that affect southern Europe and, in general, all the mountains in the world, with a particular intensity (Beniston, 2003). For this reason, detailed studies in catchments with different plant covers and land use/cover are of paramount importance to compare the hydrological responses at various temporal scales (from daily events to long-term average values). This information is determinant to validate hydrological models and to forecast future trends on water resources and sediment transport under different environmental scenarios.

The Department of Geo-Environmental Processes and Global Change (Instituto Pirenaico de Ecología, CSIC) maintains a experimental station in the central Pyrenees composed by three monitored catchments. They are located in a middle mountain area, within a distance of less than 8 km, which ensures the same precipitation characteristics (Figure 3). Each catchment is characterized by different plant cover and lithology. The San Salvador catchment (92 ha) is a tributary of the Estarrún River, with flysch bedrock, composed of thin beds of alternating sandstones and marls. The outlet is located at 810 m, and the highest point at 1330 m. The catchment is covered by dense forest of *Pinus sylvestris*, *Fagus sylvatica* and *Quercus gr. faginea*. Soils are relatively deep and well preserved. The Arnás catchment (284 ha) is a west-east oriented tributary of the Lubierre River. The substratum is also Eocene flysch. The north-facing slope shows steeper slopes, a relatively open shrub cover and thin soils, with some sectors affected by sheet-wash erosion. The south-facing slope has gentler slopes and deep soils, and is densely covered by shrubs and patches of forest. The outlet is located at 900 m and the highest part at 1330 m. The whole catchment was cultivated until the middle of the 20th Century with cereal crops in non-terraced fields. Since then, the catchment was abandoned and affected by a process of natural plant colonization. The Araguás catchment (45 ha) is a tributary of the Lubierre River. The outlet is located at 780 m and the highest point is at 1105 m. Bedrock is Eocene marl in the lower part, and flysch in the upper zone. The main characteristic is the presence of a dense network of badlands in which small mudflows, gullying, rilling and sheet wash erosion prevail. The average annual precipitation is about 800 mm in the three catchments, though a remarkable inter-annual variability exists.

The three catchments are equipped with automatic weather and gauging stations which allow to monitor several climatic parameters as well as the discharge—by an ultra-sound sensor—and suspended sediment concentration—thanks to a turbidimeter and an automatic water-sampler. Besides, other instruments were installed: Seven and three piezometers respectively in Arnás and San Salvador to continuously record the height of the water table; at least two more pluviometers in each catchment to control the spatial variability of the rainfall; three plots to measure rainfall interception under forest cover (pine, beech-tree and oak) in San Salvador; and temperature sensors at different

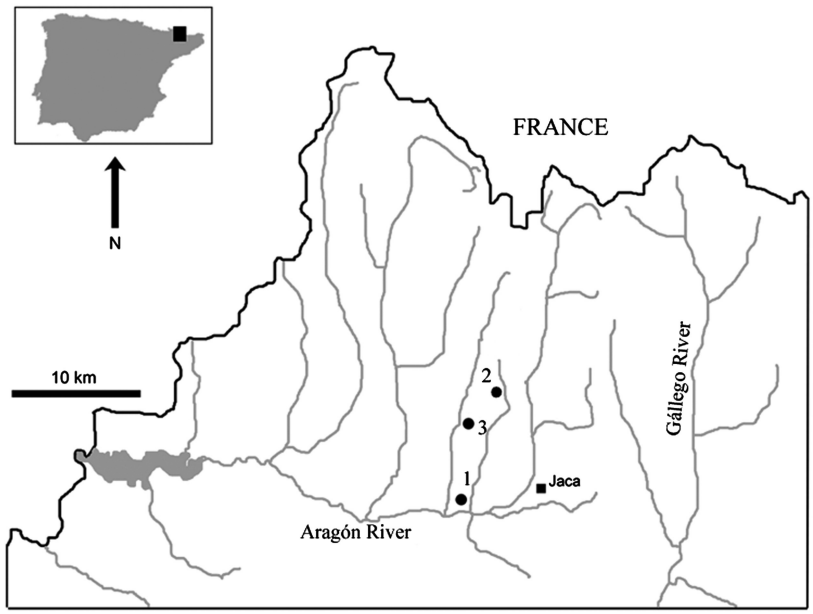


Figure 3. Location of the three experimental catchments within the Yesa reservoir catchment (upper Aragón River). Araguás (1), Arnás (2), San Salvador (3).

depths of the regolith in Araguás. A database was elaborated with series of precipitation, discharge and suspended sediment concentration, with a time resolution of five minutes.

The most relevant results obtained up to now are summarized as an example of the information supplied by the catchment-scale approach. The main question that this experimental setting tries to answer is how land cover influences runoff and flood generation, the seasonality of floods and the spatial and temporal variability of sediment transport. It is interesting to note that, given a similar precipitation and almost similar topographic characteristics for the three catchments, the differences in runoff and sediment transport can be attributed to the differences in plant cover.

A major hydrological difference refers to the runoff coefficient, for which a clear decrease is observed as plant cover density increases (Figure 4). The overall runoff coefficient is 69% in the badland catchment (Araguás), followed by the old cultivated catchment (Arnás, 25%) and the forest catchment (San Salvador, 11%). Evapotranspiration is, obviously, the main factor explaining these differences. Thus, for a six-month period in 2006, interception under pine, beech tree and oak was 15, 19 and 16% in average respectively. In the case of beech tree with shrubs, interception reached 27% of the rainfall.

The relationship between precipitation and discharge at event scale also shows big differences between the catchments. No good correlation was found between both variables in San Salvador catchment ($r^2 = 0.047$), in which the hydrological response shows an extreme variability scarcely related to the volume of the rainstorm. In Arnás catchment, the variability is also high, although in general the peak flow increases as the rainstorm event does ($r^2 = 0.50$). In Araguás the correlation is very high, particularly in the case of rainfalls over 10 mm ($r^2 = 0.84$), in such a manner that the characteristics of the rainstorm event explains a large proportion of the variability of the hydrological response.

The intensity of the response during the year also varies greatly between the catchments (Figure 5). Araguás reacts very fast to each rainstorm event, with flash floods characterized by steep rising and recessing limbs and narrow peak flows. Arnás also reacts with frequent floods, although the intensity of the response is lower particularly in autumn, when the soils in the catchment are still recovering from the exhaustion of the summer. In San Salvador there are few floods, in

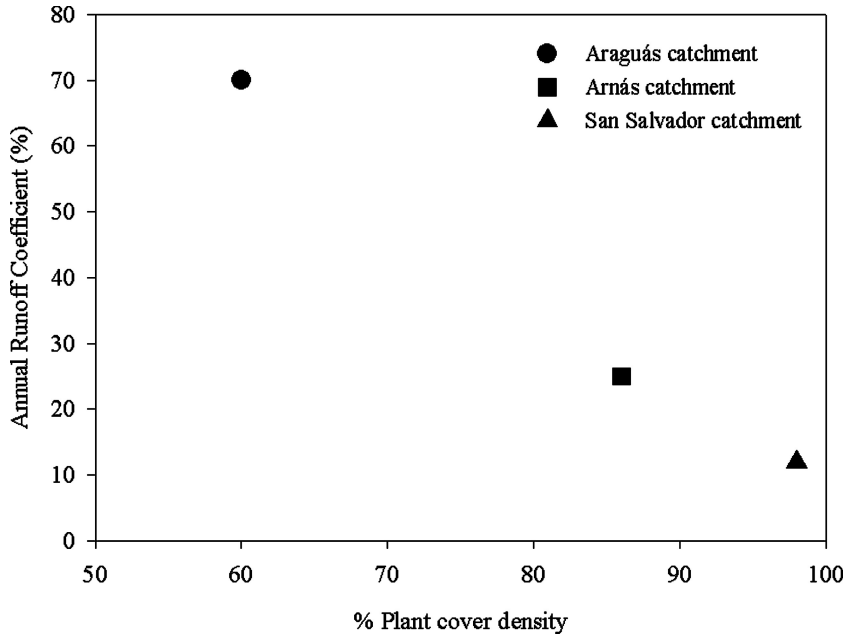


Figure 4. Relationship between plant cover density and runoff coefficient for the three experimental catchments, illustrating the increase in runoff production as the vegetation cover decreases.

general of low intensity, except in spring when they can be of similar importance than those of Arnás. In fact, during the 2005-06 hydrological year, which can be defined as "dry", Araguás recorded 44 floods, Arnás 12 and San Salvador only six. Remarkable differences in the timing of the floods are also evident. In Araguás floods are possible in any season of the year, even in summer and autumn; in Arnás floods tend to be relatively moderate during the autumn, except during very intense rainstorms, they are also of moderate size in winter, and very intense in spring, whereas no floods are recorded in summer; in San Salvador, the only season with intense floods is the spring, and only small and infrequent hydrological reactions are observed in autumn and winter.

This demonstrates that plant cover is able to change not only the magnitude and frequency of floods but also their seasonal distribution, and it is a key aspect to understand the hydrological functioning of the catchments. In the case of Araguás, the hydrological pattern is very simple: it is ready to produce a flood at any moment of the year given the almost total absence of soil and the impervious behavior of the marl. Thus, the occurrence of floods in Araguás is parallel to the occurrence of rainstorms. But, what is the difference between Arnás and San Salvador? Information on the groundwater table (Figure 6) shows that in San Salvador runoff in the main channel is only possible when the water table is close to the surface, and this tends to occur in spring. This implies that overland flow is caused by saturation excess flow, and it is the reason why the rest of the year runoff generation is exceptional (Serrano-Muela et al., in press). In Arnás, the water table has also a seasonal behavior related to the evolution of precipitation and temperature (Figure 7). After the summer drought, a recovery period is observed during the autumn in which the water table shows a constant rising; a stable period occurs from December to April, when the water table is close to the surface; and a recession period occurs in late spring and summer, when evapotranspiration exhausts progressively the water reserves. Nevertheless, floods, even of limited magnitude, are relatively frequent in autumn and winter. This is only possible taking into account the existence of relatively small sheet-wash erosion areas close to the channel, which yield overland flow with any rainfall intensity and volume. Then, the Arnás catchment produces small floods when the catchment is dry or in intermediate periods, and only small impervious areas contribute runoff, and large floods when it is wet and there is an expansion of the contributing areas. The hydrological results presented

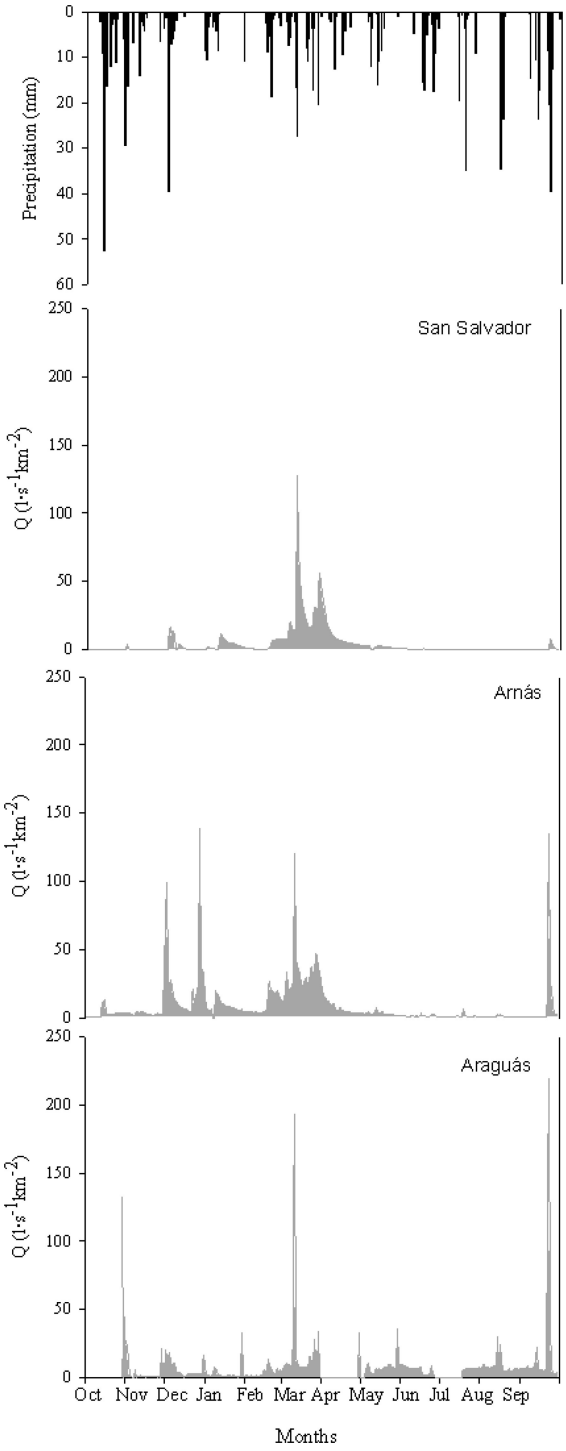


Figure 5. Hydrological response (water discharge) in the San Salvador, Arnás and Araguás catchments during one year. The input precipitation was roughly the same, so it is represented only once.

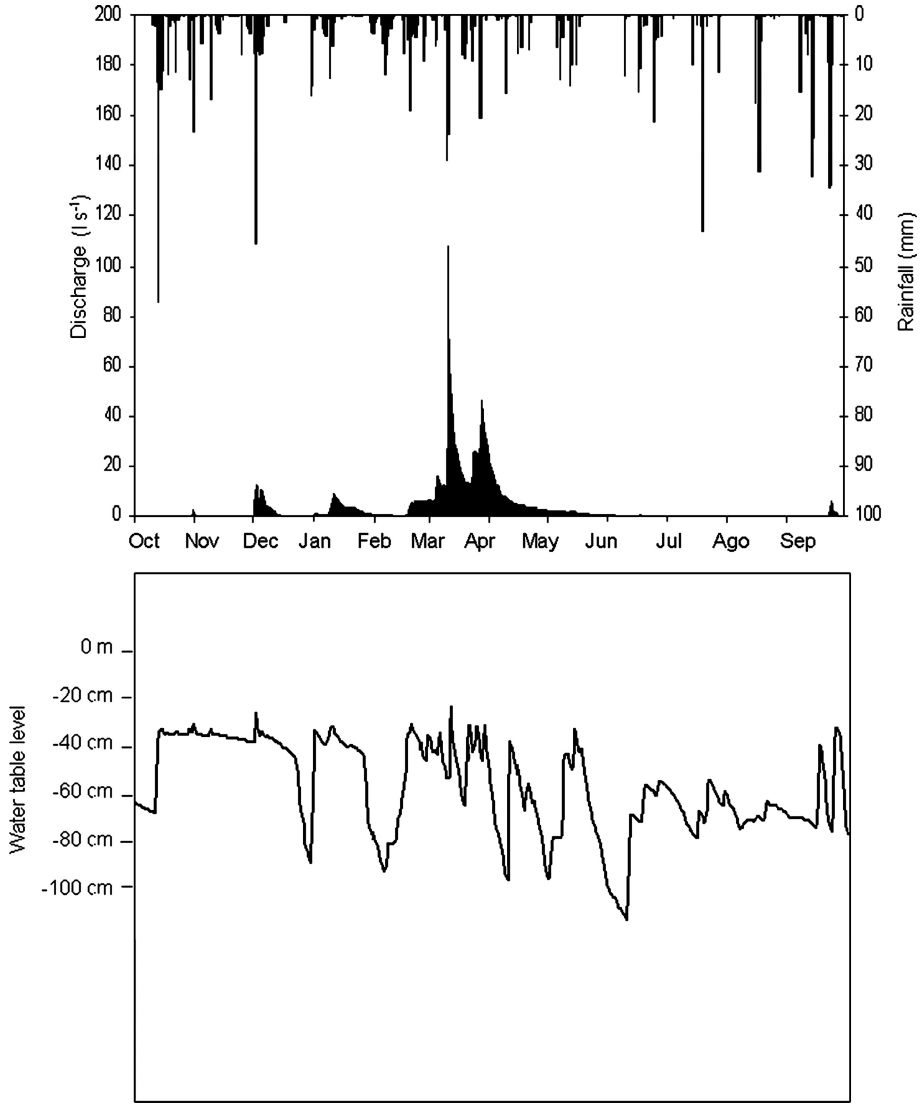


Figure 6. Temporal variation of water table depth and discharge during one year in the San Salvador catchment.

so far explain the differences in sediment transport. Major differences exist in the composition of the sediment exported from each catchment. In San Salvador most of the sediment exported is in the form of solutes (73.5%), because water circulates preferably within the soil, and only 26.5% is exported as suspended sediment. No bed load has been recorded. In Arnás, suspended sediment prevails (46% of total), followed by solutes (34%) and bed load (about 20%). This is an indicator that overland flow and soil erosion are active processes in Arnás, at least in some areas which have good connection with the channel or within the channel itself. In the case of Araguás, there are no records about the proportion of each type of sediment yet, though suspended sediment ought to be predominant since sediment sources are mostly bare marls in badlands. The maximum values of suspended sediment concentration recorded in each catchment are a good element for comparison: 1.9 g L-1 in San Salvador, 10 g L-1 in Arnás and more than 500 g L-1 in Araguás, with many floods recording more than 100 g L-1. With these values and those from runoff coefficients it is evident that the sediment yield per hectare is several orders of magnitude higher in Araguás than

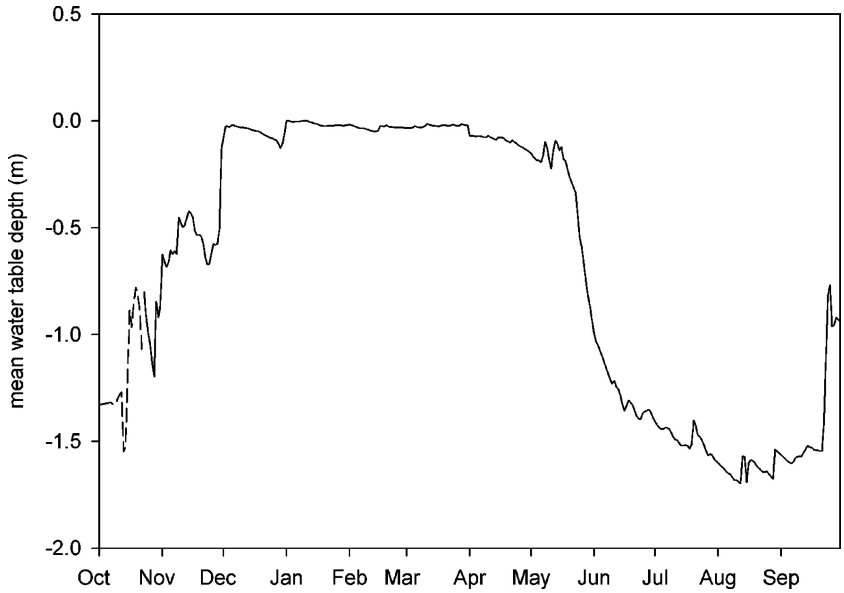


Figure 7. Temporal variation of the water table depth during one year in the Arnás catchment.

in Arnás and even more than in San Salvador. One more time, plant cover differences are the best explaining factor.

5 THE LINK BETWEEN EMPIRICAL STUDIES AND SIMULATION MODELS

Erosion and sediment transport simulation models have experienced vast improvements in the last two decades, in parallel with the increase in computing power. Simulation models are attractive not only to the basic scientists who seek a convenient tool to test their ideas about the interplay of different erosion processes, but also to the applied researchers and practitioners due to their predictive and scenario testing capabilities.

Catchment stakeholders and decision makers need spatially distributed information on erosion and deposition rates as well as estimations of the water and sediment discharge from individual catchments. This information is preferably required at an event temporal scale, which allows reflecting the spatial pattern and timing of erosion and sediment transport processes in response to particular events. Current computation capabilities make it possible, from event based models, to conduct long simulations and derive long term average erosion and sediment export values.

Even if we restrict ourselves to spatially distributed, event based models, many candidate models exist which differ in the processes they incorporate and the manner these processes are treated, their spatial and temporal scales, complexity, input needs, etc. This makes choosing a model a difficult task, which depends very much on what are the main questions the model user wants to get an answer to. This requires at least a basic background on the subject of computer modeling, and that is something which model users often lack; hence, it is not infrequent to see mistaken model choices which lead to unsound results.

A large portion of the sediment budgets at catchment scale are based on field determined erosion and deposition rates, or on the use of empirically derived relationships. The use of process-based models, on the contrary, is still a minority. This is due to a combination of factors, and some of which are discussed in the next paragraphs.

Some of the difficulties of building good process-based simulation models at catchment scale arise obviously from the natural complexity of the system being modeled. As it has been shown

before, several erosion and sediment transport processes interplay at different spatial and temporal domains, and with differences of several orders of magnitude. Although we have achieved good understanding of many erosion processes separately, an effort has still to be done to fully understand and incorporate into the models the interrelation between different processes. For example, there is a limited knowledge about the role of sediment sinks in the spatial connectivity of the erosion processes: in the plot domain, sediment trapping in small depressions which are usually finer than the resolution of digital terrain models—surface roughness; in the hillslope domain, sediment retention due to changes in the vegetal cover, in small concavities or in the bottom part of the slope; in the catchment domain, importance of stream erosion and deposition processes. In all different domains, there is a need to characterize these sediment storages, determining their importance and the mean residence time of the sediment.

Probably as a response to this natural complexity, models have grown very complex too. In most cases, simulation models are tested in small, controlled and data intensive experimental catchments, but they show big problems when applied to spatially larger, data scarce environments. Complex models can better represent a complex reality, but this often at the expense of becoming more prone to error propagation. Each new parameter in a model comes in association with some level of uncertainty, and this uncertainty is usually magnified in the model's results. In simple words, more complex does not always mean more accurate or less uncertain, these depend very much on the particular case of study.

This leads us to the problem of data availability. Process-based models often require large amounts of input data. Very often this data does not exist for a particular area, and need to be generated. In many cases this data is inferred from another auxiliary variable, such as the topography or the vegetation cover, resulting in an apparently good estimation in terms of spatial resolution, but with a high level of uncertainty. In other cases a field sampling campaign is undertaken, but in general only a few samples can be analyzed which allow for a general picture of the average values of the model's parameters at some characteristic areas, but do not give an adequate description of the spatial variability of the parameter. Spatial variability can be very large for some parameters which are of great importance to erosion models—such as the soil properties—and this variability needs to be accounted for in the model. Usually, the implicit assumptions made to provide a spatial representation of the input parameters are not recognized, resulting in an important underestimation of the uncertainty introduced in the simulations, and leading ultimately to frustrating results.

6 CONCLUSIONS

Compared to natural climatic variability, plant cover and land use explain a large part of the hydrological and erosive variability in Mediterranean mountain areas. However, we need to improve our understanding of the balance between the climatic and land cover drivers, since the quantity and quality of water resources in the Mediterranean region mainly depends on mountain areas. We know that abrupt plant cover changes occurred in short periods in the past. In the case of the Pyrenees strong population pressure, overgrazing, farming in steep slopes and frequent man-induced fires for centuries resulted in landsliding, flooding and active erosion processes which caused soil degradation in many hillslopes, and the development of braided channels in most of rivers (Beguería et al., 2006). However, throughout the 20th century a rapid process of depopulation and farmland abandonment occurred, and hence the old fields and grazed areas were progressively colonized by scrubland and forests. These changes in plant cover resulted in a shrinkage of sediment sources (García-Ruiz and Valero-Garcés, 1998), a reduction of flooding (López-Moreno et al., 2006) and a negative temporal trend in stream flow (Beguería et al., 2003). Being able to get operative estimations of the consequences of possible future climate and land use changes is of paramount importance everywhere and specially in the Mediterranean mountain context. This will help to plan the management of reservoirs and of different water uses (agriculture, industrial or domestic uses, environmental flow, etc).

The results obtained from the experimental catchments in the Pyrenees confirm all these trends and allow us to understand the leading processes in catchments of different characteristics, such as the seasonal fluctuations of water table, the rate of regolith weathering, the rainfall interception under forest and scrubland cover, the location of runoff and sediment sources and the temporal variability of contributing areas. This is basic information useful to model makers as well as to basic scientists, and the effort should be maintained to include the new problems that arise from the research, such as the role of the sediment sinks in the spatial connectivity of the erosion processes.

Research on simulation models should be maintained at the same level, too. The experience won from observational studies and from using complex partial models should be used to determine which are the main processes leading erosion and sediment transport at the different scales. This knowledge should be used to build a new generation of models, characterized by their robustness and their ability to work in data scarce environments. Parameter uncertainty and error propagation is another feature that must be included explicitly in the models, as a way to make the model users aware of its importance. In any case, the usefulness of new simulation models will depend strongly on their ability to integrate processes operating at different spatial and temporal scales, from the event/plot to the long term/catchment.

REFERENCES

Almorox, J., De Antonio, R., Saa, A., Díaz, M.C. & Gascó, J.M. 1994. *Métodos de estima de la erosión hídrica*. Madrid: Editorial Agrícola Española.

Beguería, S. 2005. Erosión y fuentes de sedimento en la cuenca del embalse de Yesa (Pirineo Occidental): Ensayo de una metodología basada en teledetección y análisis SIG. Zaragoza: Instituto Pirenaico de Ecología.

Beguería, S., López-Moreno, J.I., Gómez-Villar, A., Rubio, V., Lana-Renault, N. & García-Ruiz, J.M. 2006. Fluvial adjustments to soil erosion and plant cover changes in the Central Spanish Pyrenees. *Geografiska Annaler*, 88A(3): 177–186.

Beniston, M. 2003. Climatic change in mountain regions: a review of possible impacts. *Climatic Change*, 59: 5–31.

Desir, G. 2001. *Erosión hídrica de terrenos yesíferos en el sector central de la Depresión del Ebro.* Consejo de Protección de la Naturaleza en Aragón, 326 pp. Zaragoza.

García-Ruiz, J.M., Lana-Renault, N., Beguería, S., Valero-Garcés, B., Lasanta, T., Arnáez, J., López-Moreno, J.I., Regüés, D. & Martí-Bono, C. 2004. Temporal and spatial interactions of slope and catchment processes in the Central Spanish Pyrenees. *IAHS Publ.*, 288: 21–28.

García-Ruiz, J.M. & Valero-Garcés, B. 1998. Historical geomorphic processes and human activities in the Central Spanish Pyrenees. *Mountain Research and Development*, 18(4): 309–320.

Gómez-Villar, A., Álvarez-Martínez, J. & García-Ruiz, J.M. 2006. Factors influencing the presence or absence of tributary-junction fans in the Iberian Range, Spain. *Geomorphology*, 81: 252–264.

González Sampériz, P. & Sopena, M.C. 2002. Recent Holocene Palaeoenvironmental evolution in the Central Ebro Basin (NE Spain). *Quaternary International*, 93–94: 177–190.

Kosmas, C., Danalatos, N., Cammeraat, L.H., Chabart, M., Diamantopoulos, J., Farand, R., Gutiérrez, L., Jacob, A., Marques, H., Martínez-Fernández, J., Mizara, A., Moustakas, N., Nicolau, J.M., Oliveros, C., Pinna, G., Puddu, R., Puidefábregas, J., Roxo, M., Simao, A., Stamou, G., Tomasi, N., Usai, D. & Vacca, A. 1997. The effect of land use and soil erosion rates under Mediterranean conditions. *Catena*, 29: 45–59.

López-Moreno, J.I., Beguería, S. & García-Ruiz, J.M. 2006. Trends in high flows in the Central Spanish Pyrenees: response to climatic factors or to land-use change? *Hydrological Sciences Journal*, 51 (6): 1039–1050.

Maladonado, A. 1983. Dinámica sedimentaria y evolución litoral reciente del delta del Ebro. In M. Mariño (edr.), *Sistema integrado del Ebro*, Madrid, pp. 33–160.

Montserrat, J. 1992. *Evolución glaciar y postglaciar del clima y la vegetación en la vertiente sur del Pirineo: Estudio palinológico.* Zaragoza: Instituto Pirenaico de Ecología.

Nadal-Romero, E., Regüés, D., Martí-Bono, C. and Serrano-Muela, P. 2007. Badland dynamics in the Central Pyrenees: temporal and spatial patterns of weathering processes. *Earth Surface Processes and Landforms*, 32 (6): 888–904.

Peña, J.L., Echeverría, M.T., Petit-Maire, N. & Lafont, R. 1993. Cronología e interpretación de las acumulaciones holocenas de la val de las Lenas (Depresión del Ebro, Zaragoza). *Geographicalia*, 30: 321–332.

Poesen, J.W.A. & Hooke, J.M. 1997. Erosion, flooding and channel management in Mediterranean environments of southern Europe. *Progress in Physical Geography*, 21 (2):157–199.

Puigdefábregas, J. & Mendizábal, T. 1995. *Desertification and migrations.* Logroño: Geoforma Ediciones.

Regüés, D. & Gallart, F. 2004. Seasonal patterns of runoff and erosion responses to simulated rainfall in a badland area in Mediterranean mountain conditions (Vallcebre, Southeastern Pyrenees). *Earth Surface Processes and Landforms*, 29: 755–767.

Serrano-Muela, M.P., Lana-Renault, N., Nadal-Romero, E., Regüés, D., Latron, J., Martí-Bono, C. & García-Ruiz, J.M. In press. Forests and water resources in Mediterranean mountains: The case of the Spanish Pyrenees. *Mountain Research and Development*.

Walling, D.E. 1991. Drainage basin studies. In O. Slaymaker (ed), *Field experiments and measurement programs in Geomorphology*, Rotterdam: Balkema.

White, S., García-Ruiz, J.M., Martí-Bono, C., Valero, B., Errea, M.P. & Gómez-Villar, A. 1997. The 1996 Biescas campsite disaster in the Central Spanish Pyrenees, and its temporal and spatial context. *Hydrological Processes*, 11: 1797–1812.

Yaalon, D.H. 1997. Soils in the Mediterranean region: what makes them different? *Catena*, 28:157–169.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Fundamentals of mathematical modelling of morphodynamic processes. Application to geomorphic flows

R.M.L. Ferreira
Department of Civil Engineering & Architecture, Instituto Superior Técnico, TULisbon, Lisbon, Portugal

ABSTRACT: Computer aided mathematical modelling of fluvial processes contributed to the integration of basic phenomenological knowledge into practical engineering works, thus alleviating long standing tensions between engineering practitioners and river scientists. In this text, it is upheld the principle that any advance in computational fluvial hydraulics is closely linked to a progress in the characterization of fluvial phenomena. Bearing this principle in mind, the objective of this work is to present the development of a computational tool suitable to model highly unsteady geomorphic flows within an engineering framework. In order to fulfil this objective, a two-dimensional vertical conceptual model, based on Chapman-Enskog theory for dense gases, is presented and the corresponding flow is numerically simulated. A one-dimensional conceptual model is then derived and its equations analysed. Numerical simulations of geomorphic flows are presented. A real-case simulation example, the 1996 flood in River Ha! Ha!, is presented. It is shown that the model is able to find non-oscillatory solutions even when subjected to important mass and momentum inputs.

1 INTRODUCTION

Fluvial Hydraulics is a discipline geared toward the study of the mechanics of open-channel flows with mobile boundaries including the joint study of hydrodynamics and the processes of entrainment, erosion, deposition and transport of sediment. The range of problems studied within the framework of Fluvial Hydraulics is wide, comprising issues such as the design of stable channels or the morphologic evolution of natural streams. Hazard mitigation has expanded the field; recent research and engineering activity have been developed in problems like avalanche and torrential flows, in particular dam-break flows.

Still a young discipline, Fluvial Hydraulics encompassed two different attitudes in what concerns the treatment of its empirical basis: that of an applied science and that of an engineering practice. Fluvial Hydraulics bore visible, until recently, unresolved tensions between its applied science and engineering conceptions. In 1956, Thomas Blench, a leading river engineer, criticised the research program of H. A. Einstein, by claiming that it presented findings "from laboratory flumes with trifling flows" and suggested that such findings were of use in the same laboratory flows (Ettema & Mutel 2004).

Engineering practice and scientific methodology have since traveled the path of reconciliation. The most significative contribution for the integration of basic phenomenological knowledge into practical engineering works is computer aided mathematical modelling of fluvial processes. De Vries (1965) work in pioneering in this sense. It represents a first attempt to model unsteady open-channel flows with mobile beds composed of well sorted sediment. The conservation equations comprehend the Saint-Venant equations for clear water and a sediment mass conservation equation, or Exner equation. The system is closed by equations that specify flow resistance and

equilibrium sediment transport rate. De Vries' model avoided a number of conceptual difficulties: i) vertical fluxes of sediment between the flow and the bed are disregarded; it is thus a capacity or equilibrium model; ii) the dynamic role of the transport layer is neglected; the flux of sediment is an abstract reality and not allocated to any particular region of the flow; and iii) equation of conservation of momentum is in the form of an energy equation; although it can be written in divergent form, *i.e.* $\partial_t(u) + \partial_x(f(u,h,Y_b)) = g(u,h)$, it is not applicable to transcritical flows where shocks or bores are susceptible to be in the solution.

Computational modelling of the morphological evolution of rivers has benefited a considerable investment in research since de Vries' pioneering work. For a literature review *cf.* Yalin (1977) or Graf & Altinakar (1998). There are also abundant examples of investigation on the system of conservation equations that describes the morphologic evolution of open-channel flows with mobile boundaries. The number of equations and dependent variables as well as the relevant terms to be included in the systems of conservation laws have been studied by Sloff (1993) Morris & Williams (1996), Sieben (1996), Capart & Young (1998), Cao *et al*. (2002), Fraccarollo & Capart (2002), Ferreira (2005) or Spinewine (2005). The mathematical structure of the equations, with emphasis on its well-posedness and on its regular/singular perturbed nature, was studied by Lyn (1987) or Ferreira (1999). The growing interest on discontinuous flows, stimulated by the interest on dam-break flows, prompted the study of the wave structure (*sensu* Lax 1957) the system of equations. A theoretical solution for the Riemann problem was devised by Fraccarollo & Capart (2002) and, for a different set of equations, by Ferreira (2005).

Improvements on numerical discretization techniques have also been achieved. Preissmann's (1961) scheme was the "work-horse" of numerical simulation from the 1960s to early the 1990s. Codes based initially on uncoupled solutions (Cunge *et al.* 1980) gave way to coupled solution procedures (Correia *et al.* 1992, Ferreira & Leal 1998, Cao *et al*. 2004). Discretization techniques based on flux difference and flux vector splitting schemes, upwind schemes and high-order non-oscillatory Total Variation Diminishing (TVD) schemes (for a review of existing methods, essentially applicable to the compressible Euler equations but also to the shallow water equations, see Hirsh 1990, LeVeque 1990 or Toro 1997) have been, since early 1990s, tried as alternatives for the one-dimensional and two-dimensional shallow-water/Exner equations. Contributions by García-Navarro & Savirón (1995), Bermudez & Vasquez (1994) or Nujic (1995) in the domain of the shallow water equations were inspiring but required further work when applied to the stream morphology problems. In particular, TVD schemes were successfully applied to the set of shallow-water and morphodynamic equations by Cao *et al*. (2004), based on a WAF flux discretization (Toro 1997), and by Ferreira & Leal (1998) or Leal *et al*. (2003) based on MacCormack's (1969) scheme.

Finite volume methods, based on Godunov's method, have been extensively used in the past decade. The first studies were concerned on clear-water shallow-water equations (*e.g*. Alcrudo & García-Navarro 1993, Fraccarollo & Toro 1995). The introduction of a contact wave in the HLL Riemann solver (HLLC Riemann solver, Toro *et al.* 1994) and the use of Roe's (1981) approximate solvers have allowed for substantive progresses in the modelling of free-surface unsteady flows over mobile beds (Callefi & Valiani 2002, Fraccarolo *et al*. 2003, Spinewine 2005, Ferreira *et al*. 2006a). Special attention is still due to the non-conservative fluxes and to the momentum source terms (Brufau *et al*. 2000, Fraccarolo *et al*. 2003).

In spite of the conceptual advances on the phenomenology of sediment transport, on the mathematical structure of the systems of equations and on discretization techniques, it cannot be said that an autonomous discipline devoted to the study of computational issues and, in particular, numerical techniques, inherent to the modelling of fluvial systems has already matured. As a matter of fact, it is often observed that progresses on the phenomenological characterization of physical systems bring about new conceptual models with different structures of the conservation equations which, concomitantly, require new computational approaches. In other words, computational advances are induced by phenomenological progress.

Recent examples comprehend the research work undertaken in EU funded projects CADAM and IMPACT (Zech & Soares-Frazão 2007) envisaging the characterization of dam-break flows in mobile bed streams. As a result of these concerted actions, a better characterization of the involved

morphological and hydrodynamic processes was obtained through laboratory work (Capart & Young 1998, Leal *et al*. 2002, Spinewine 2003), theoretical work and numerical simulation efforts (Zech *et al*. 2006).

In this text, it is upheld the principle that any advance in computational fluvial hydraulics is closely linked to a progress in the characterization of fluvial phenomena. Bearing this principle in mind, the general objective of this work is to present the development of a computational tool suitable to model highly unsteady geomorphic flows, *i.e*. flows associated with severe morphological impacts, within an engineering framework. To meet the last premise, a compromise between computational time and accuracy in the description of the morphological changes must be achieved. In this text, a continuum one-dimensional framework was preferred, although some care is placed in the description of vertical exchanges of matter and on two-dimensional effects such as bank erosion and lateral afflux of mass. The particular objectives of the work include: i) the derivation of the conservation and closure equations (section 2); ii) the mathematical analysis of the system of conservation laws (section 3); the description of applicable discretization schemes and the presentation of examples of application (section 4). As a real-case simulation example, it was chosen the well-documented, both in the morphologic and hydrodynamic realms, 1996 flood in River Ha! Ha!, a composition of hydrologic and dam-break floods.

2 CONCEPTUAL MODELS

2.1 *A conceptual model for the flow of dense mixtures of fluid and granular matter*

Figure 1 shows the structure of flow of a dense mixture of fluid and well sorted sediment. The diagram on the left hand side shows the distribution of the shear stresses over the flow depth. The nature of the production mechanisms is identified in the diagram. An idealisation of the mean (time-averaged) velocity profile is shown in the right hand side of figure 1.

Computational layer [1] includes the uppermost flow layer (**A** in figure 1) and the transitional layer **B**. In layer **A**, turbulent stresses are dominant. Layer **B** is characterized by relatively low solid fractions and streaming stresses dominate over collisional. Streaming stresses are originated by collisions but not at the moment of the collision; they represent the momentum transported by the erratic particle motion between collisions. In this layer, Reynolds stresses become important.

Computational layer [2] is identifiable with the transport layer, composed of layers **D**, where frictional stresses are predominant, and **C**, where collisional stresses dominate and where fluid

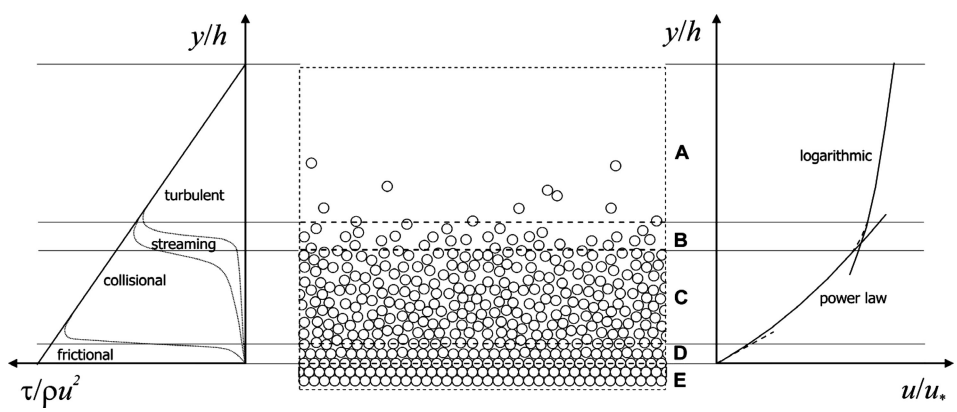


Figure 1. Idealised structure of the physical system as a layered granular/fluid medium. The relevant flow regions are: **A** – clear water/suspended sediment sub-layer; **B** – pick-up sub-layer; **C** – collisional sub-layer; **D** – frictional sub-layer and **E** – stream bed, composed of immobile grains (adapted from Ferreira 2005, p. 230).

viscosity plays a relevant role (details in Ferreira 2005, §3). The transport layer is characterized by high solid fractions (its depth-average value is typically between 0.2 and 0.5). In both these layers the sediment load is ultimately supported by the bed. This implies that the contribution of lift forces, whether due to turbulence or to the mean velocity gradient, is disregarded as negligible. In the frictional layer, the submerged weight of the sediment grains is equilibrated by a reaction force in the bed, solicited by quasi-permanent particle contacts with the bed.

Computational layer [3] corresponds to the bed. The bed surface is the plane that connects the centres of gravity of the uppermost layer of the pack of grains with no longitudinal velocity (region **E** in figure 1). The movement of the grains in the bed is reduced to its state of agitation.

The physical system depicted in figure 1 has a conceptual counterpart, the set of conservation and constitutive equations that describe the flow. Computational layer [1] is described by a fluid mass conservation equation and by the Reynolds-averaged Navier-Stokes equations (Shlichting 2000), which, for an incompressible fluid and neglecting suspended sediment, can be written:

$$u_{i,i}^{(w)} = 0 \tag{1}$$

$$\rho^{(w)} \mathrm{D}_t(u_j^{(w)}) = P_{ij,i} - T_{ij,i}^{(t)} + \rho^{(w)} g_j \tag{2}$$

where $\rho^{(w)}$ is the density of the sediment grains, $u_i^{(w)}$ stands for the velocity field, P_{ij} is the pressure tensor, $P_{ij} = P\delta_{ij}$, where P is the isotropic pressure, $T_{ij}^{(t)}$ is the turbulent stress tensor (viscous stresses are negligible) and g_j is the acceleration of gravity.

As for the characterization of the transport layer depicted in figure 1 (computational layer [2]), the most consistent theoretical body applicable is that drawn from the analogy to the kinetic theory of gases (Chapman & Cowling 1970), built by works such as Lun *et al.* (1984), Jenkins & Richman (1985, 1988) or Armanini *et al*. (2004) and herein designated by Chapman-Enskog theory for dense gases. If the flowing particles are smooth, round and only slightly inelastic, *i.e*. the amount of energy dissipated is small, the Chapman-Enskog theory for dense gases (Chapman & Cowling 1970, §16) requires only small modifications in order to provide conservation laws for the mass, momentum and fluctuating energy of the granular material and constitutive equations for the stress tensor, diffusion coefficients and dissipation rate. Kinetic theories became a new paradigm for the conceptualization of granular flows but this approach is not exempted from theoretical difficulties: the inclusion of frictional effects and viscous fluid-grain interactions is, at present, done by patching in *ad hoc* theories (details in Ferreira 2005, p. 249).

Within the framework of the Chapman-Enskog theory, it is possible to derive the equations of conservation of mass, momentum and energy (associated to the fluctuating motion) for the granular phase (details in Ferreira 2005, pp. 231–250) in computational layer [2]. These can be written, respectively, as

$$\rho^{(g)} \mathrm{D}_t(\nu) + \rho^{(g)} \nu u_{i,i}^{(g)} = 0 \tag{3}$$

$$\rho^{(g)} \nu \mathrm{D}_t(u_j^{(g)}) = T_{ij,i}^{(g)} + \rho^{(g)} \nu g_j + f_j^{(gw)} \tag{4}$$

$$\tfrac{3}{2} \rho^{(g)} \nu \mathrm{D}_t(\Theta) = -\Phi_{i,i} + T_{ij}^{(g)} u_{j,i}^{(g)} - \gamma^{(gw)} \tag{5}$$

where $\rho^{(g)}$ is the density of the sediment grains, ν is the solid fraction (concentration at a specific point in the flow), $u_i^{(g)}$ stands for the velocity field of the granular phase, $T_{ij}^{(g)}$ is the granular stress tensor, g_j is the acceleration of gravity, $f_j^{(gw)}$ is the force per unit volume encompassing the interaction (essentially of viscous nature) of the fluid and granular phases and buoyancy, $\Theta = \frac{1}{3}\langle c_i' c_i' \rangle$ is the granular temperature, c_i' is the fluctuating component of the particle velocity, Φ_i is the flux of fluctuating energy and $\gamma^{(gw)}$ is the rate of dissipation (due to inelastic collisions and viscous damping) of the fluctuating energy. In equations (3) to (5) the operator $\mathrm{D}_t(\cdot)$ stands for the material derivative for which the convective operator is relative to the mean low, Einstein's notation is used for space derivatives (for instance, $u_{i,i}^{(g)}$ is the divergence of the velocity field) and the bracket operator $\langle \cdot \rangle$

stands for point-wise time or ensemble average (ergodicity is assumed). The equation of conservation of fluctuating energy, equation (5), reveals that, unlike thermodynamic systems where every given constant temperature can be made to correspond to an equilibrium state, a granular system can maintain a steady state of agitation, characterized by a given granular temperature, if and only if the rate of production equals the diffusive flux and the dissipation, *i.e.* if $T_{ij}^{(g)} u_{j,i}^{(g)} = \Phi_{i,i} + \gamma^{(gw)}$.

Conservation equations are also needed for the water phase, in order to complete the description of computational layer [2]. These can be obtained from a control volume analysis within the continuum hypotheses:

$$-\rho^{(w)} \mathrm{D}_t(\nu) + (1-\nu)\rho^{(w)} u_{i,i}^{(w)} = 0 \tag{6}$$

$$(1-\nu)\rho^{(w)} \mathrm{D}_t(u_j^{(w)}) = T_{ji,i}^{(w)} + (1-\nu)\rho^{(w)} g_j - f_{Dj}^{(gw)} \tag{7}$$

Constitutive equations are required for the stress tensors, for the flux of fluctuating energy, for the dissipation of fluctuating energy, computation and for the interaction force. An example of computation is given in the next section.

2.2 *Steady flow description of the transport layer: numerical solution*

A detailed characterization of the two-dimensional (vertical) flow in the transport layer is obtained by solving numerically the following set of ODEs

$$\frac{\mathrm{d}\mathbf{Y}}{\mathrm{d}y} = \mathbf{M}(\mathbf{Y}, y) \tag{8}$$

where $\mathbf{Y} = [\ T^{(g)} \quad P^{(g)} \quad u_x^{(g)} \quad \Theta \quad \Phi \quad \nu \quad T^{(w)} \quad P^{(w)} \quad u_x^{(w)}\ ]^{\mathrm{T}}$ and

$$\mathbf{M}(\mathbf{Y}, y) = \begin{bmatrix} -\rho^{(g)}\nu g \sin(\beta) - f_D \\ -\rho^{(w)}(s-1)\nu g \\ \frac{5\pi^{\frac{1}{2}}}{8} \frac{T^{(g)}}{\nu\rho^{(g)}\vartheta_1\vartheta_3 d_s} \frac{1}{\Theta^{\frac{1}{2}}} \\ -\frac{\pi^{\frac{1}{2}}}{4} \frac{1}{\nu\rho^{(g)}\vartheta_1\vartheta_4} \frac{\Phi}{d_s\Theta^{\frac{1}{2}}} \\ \frac{5\pi^{\frac{1}{2}}}{8} \frac{\left(T^{(g)}\right)^2}{\nu\rho^{(g)}\vartheta_1\vartheta_3} \frac{1}{d_s\Theta^{\frac{1}{2}}} - \frac{24}{\pi^{\frac{1}{2}}}(1-e^{(gw)})\rho^{(g)}\nu\vartheta_1 \frac{\Theta^{\frac{3}{2}}}{d_s} \\ \frac{\pi^{\frac{1}{2}}}{4} \frac{1+\vartheta_1}{\rho^{(g)}\vartheta_1\vartheta_4\left(1+8\vartheta_1+4\nu^2\frac{\mathrm{d}g_0}{\mathrm{d}\nu}\right)} \frac{\Phi}{d_s\Theta^{\frac{3}{2}}} - \frac{s-1}{s} \frac{g\nu}{\left(1+8\vartheta_1+4\nu^2\frac{\mathrm{d}g_0}{\mathrm{d}\nu}\right)\Theta} \\ -\rho^{(w)}(1-\nu)g\sin(\beta) + f_D \\ -\rho^{(w)}g \\ \left(\frac{T^{(w)}}{\rho^{(w)}}\right)^{\frac{1}{2}} \frac{1}{\kappa y(1-\nu)^{\frac{1}{2}} f(\nu)} \end{bmatrix}$$

where β is the inclination of the channel, $s = \rho^{(g)}/\rho^{(w)}$ is the specific gravity of the grains, f_D is the drag force per unit volume (details *in* Ferreira 2005, p. 249), d_s is characteristic diameter of the grains, ϑ_1, ϑ_2, ϑ_3, ϑ_4, g_0 and $f(\nu)$ are functions of the solid fraction (details *in* Ferreira 2005, p. 246), $e^{(gw)}$ is the immersed restitution coefficient (Ferreira 2005, p. 248), κ is the Von Kármán constant and y is the vertical co-ordinate. It should be noted that the granular normal stresses are isotropic, hence reduced to the granular pressure $P^{(g)} = \nu\rho^{(g)}(1+4\vartheta_1)\Theta$ and the mixture behaves as an incompressible fluid. This is a direct consequence of the dense limit approximation (see details *in* Jenkins & Richman 1988, Ferreira 2005, pp. 231–238).

System (8) must be solved simultaneously, subjected to the following boundary conditions. The granular pressure at the bed must be equal to the submerged weight, per unit volume, of the granular material in the transport layer, i.e.

$$P^{(g)}(y = Y_b) \equiv N = g\rho^{(w)}(s-1)\int_0^{h_c} \nu \, \mathrm{d}\xi \tag{9}$$

where h_c is the thickness of the transport layer. It was affirmed that the boundary between the frictional layer and the bed is a Coulomb yield surface. In that case

$$T^{(g)}(y = Y_b) = N \tan(\varphi_b) \tag{10}$$

where $\tan(\varphi_b) = 0.4$ is the tangent of the internal friction angle at the bed (details *in* Ferreira 2005, p. 253). Knowing the granular pressure at the bed, the granular temperature can be obtained by the equation of state

$$\Theta(y = Y_b) = N/(\nu_0 \rho^{(g)}(1 + 4\vartheta_{1_0})) \tag{11}$$

where $\nu_0 = 1 - p$ is the solid fraction of the bed and p the bed porosity. No slip is assumed for both fluid and granular velocities, thus $u_x^{(w)}(y = Y_b) = u_x^{(g)}(y = Y_b) = 0$. The boundary condition for the flux of fluctuating energy can be derived in a manner similar to that of Jenkins & Askari (1991). Ferreira (2005), p. 255, obtained

$$\Phi(y = y_b) = -\frac{2}{\pi^{\frac{1}{2}}}(\Theta(y = y_b))^{\frac{1}{2}} N \sqrt{\vartheta_{4_0}\left(3(1 - e^{(gw)}) - \tfrac{5\pi}{4\vartheta_{3_0}} \tan^2(\varphi_b)\right)} \tag{12}$$

The flux of fluctuating energy requires another boundary condition at the top of the transport layer expressing particle rarefaction $\Phi(y = h_c) = 0$. Since this condition introduces another unknown, h_c, another condition must be imposed: the top of the transport layer is defined as the elevation for which the solid fraction is $\nu = 0.001$. As for the fluid bed shear stress, it is hypothesized that $T^{(w)}(y = Y_b) = 0.05\, g\rho^{(w)}(s - 1)\, d_s$. The fluid pressure over the transport layer is, assuming that there is no pore pressure build up, hydrostatic: $P^{(w)}(y = Y_b) = g\rho^{(w)}h$, where h is the full flow depth. It should be highlighted that this problem is over-conditioned, there are more boundary conditions than unknowns, and, hence, is ill-posed.

The numerical solution of the system of equations (11) and corresponding boundary conditions, can be obtained with a shooting method and a four-point Runge-Kutta discretisation. The shooting method can be described as follows: i) the value of the normal stress at the bed is imposed tentatively and the system (8) is solved for an arbitrarily large y; ii) the value of h_c is determined from the condition $\nu(y = Y_b + h_c) = 0.001$; iii) the normal stress is computed by equation (9) and compared with the tentative value; iv) the process is repeated until the normal stresses, computed and imposed, are approximately equal. This algorithm does not ensure vanishing flux of fluctuating energy at the top of the transport layer. This is a consequence of the ill-posed nature of the problem. An extra degree of freedom must be imposed; it was chosen to adjust the immersed coefficient of restitution. Every instance of the adjustment of this coefficient requires the application of the algorithm explained above.

The system of ODEs and respective boundary conditions was solved for the experimental conditions of Sumer *et al.* (1996). The results for plastic pellets with $d_s = 0.003$ m, $s = 1.27$ and dry coefficient of restitution $e = 0.825$ are shown in figure 2.

It was found that the immersed restitution coefficient is $e^{(gw)} = e - 0.005\theta$ where $\theta = T(y = Y_b)/(\rho^{(w)}g(s - 1)d_s)$ is the Shields parameter. The numerical results comprise profiles of the mean velocity of the granular constituent, of the solid fraction, of the granular temperature, of the flux of fluctuating energy, of the shear granular stresses and of the productive, dissipative and diffusive terms of the equation of conservation of the fluctuating energy.

In figure 2, it is observed that the solid fraction increases almost linearly with the flow depth. The shear stress is not constant in the transport layer because its work must balance that of the force

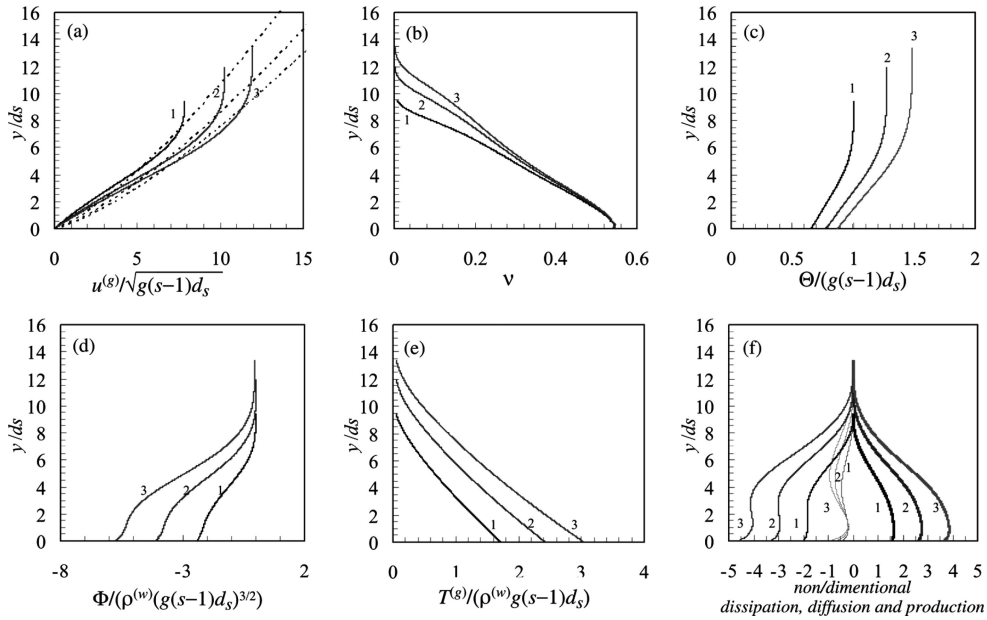


Figure 2. Computed profiles of the relevant non-dimensional quantities in the transport layer. (a) Velocity of the granular constituent. (b) solid fraction. (c) granular temperature. (d) flux of fluctuating energy. (e) granular shear and normal stresses. (f) dissipation, diffusion and production of fluctuating energy; production is considered positive while diffusion and dissipation are negative. Simulations 1, 2 and 3 correspond to, respectively, $\theta = 1.74$, $\theta \approx 2.49$ and $\theta \approx 3.07$.

of gravity in the direction of the flow. The modulus of the flux of the fluctuating kinetic energy increases towards the bed. This means that fluctuating energy is constantly being extracted from the mean flow and directed towards the bottom. As a consequence, the frictional sub-layer cannot increase indefinitely with increasing values of the Shields parameter and consequent increasing normal stresses.

The flux of fluctuating energy is, as required, zero at a large distance form the bed, as the system becomes sufficiently rarefied. The derivative of the flux, the diffusion of the fluctuating energy is plotted in figure 2(f) along with the production and dissipation. It is clear that the diffusion term is much smaller than the production and the dissipation terms, which justify its elimination from some theoretical formulations (see Armanini *et al.* 2004). An increase in the Shields parameter represents an increase in the shear rate. Thus, the absolute value of all quantities in figure 2 increase with the Shields parameter since they are all increasing functions of the shear rate.

2.3 *Shallow-flow conservation equations for the flow of dense mixtures of fluid and granular matter*

Having characterised the physical system, a tractable conceptual model for unsteady flows of fluid-granular mixtures is now devised. Vast applicability, low computational cost and conceptual simplicity are the main reasons that justify the option for a one-dimensional framework. In order to derive the conservation equations the following technique is be followed: i) the two-dimensional equations of conservation of each constituent are summed (equations 3 and 6 and 4 and 7); ii) cinematic and dynamic boundary conditions are applied and iii) the continuum hypothesis are employed and the resulting equations are integrated.

Implicit in the derivation process are a) the shallow-flow hypotheses, namely the hydrostatic distribution of pressure consequence of the small curvature of the streamlines, b) the hypothesis

that the sediment mixture is characterized by a single characteristic diameter and c) the hypothesis, corroborated by the numerical simulation carried out in the last section, that the relative velocity $u_r = u_x^{(w)} - u_x^{(g)}$ is small. In fact it was verified that $\langle u_r / u_r u_x^{(g)} \rangle = \mathrm{Ord}(10^{-1})$.

Conservation equations for fluid mass and momentum of computational layer [1], obtained by the integration of equations (1) and (2), are, respectively

$$\partial_t(h_w) + \partial_x(u_w h_w) - \phi_{2,1}^{net} = 0 \tag{13}$$

$$\partial_t(\rho^{(w)} u_w h_w) + \partial_x(\rho^{(w)} u_w^2 h_w) + \tfrac{1}{2} g \partial_x(\rho^{(w)} h_w^2) = \rho^{(w)} u_I \phi_{2,1}^{net} - g\rho^{(w)} h_w \partial_x(Y_b) - \tau_{2,1} \tag{14}$$

where $h_w = h - h_c$ is the thickness of layer [1], h is the water depth, $u_w = (uh - u_c h_c)/(h - h_c)$ is the velocity of layer [1], u_b is the velocity of layer [2], u is the depth-averaged flow velocity, Y_b is the bed elevation, and $\phi_{2,1}^{net}$, u_I and $\tau_{2,1}$ are, respectively, the net rate of mass exchange, a competent velocity and the shear stress between layers [1] and [2].

The conservation of fluid mass and momentum for computational layer [2] and the conservation of sediment mass for the granular constituent are, respectively

$$\partial_t (h_c) + \partial_x (u_c h_c) + \phi_{2,1}^{net} - \phi_{3,2}^{net} = 0 \tag{15}$$

$$\begin{aligned} &\partial_t (\rho_c u_c h_c) + \partial_x \left(\rho_c u_c^2 h_c\right) + \tfrac{1}{2} g \partial_x (2\rho^{(w)} h_w h_c + \rho_c h_c^2) \\ &\quad = -\rho^{(w)} u_I \phi_{2,1}^{net} - g\rho_c h_c \partial_x (Y_b) + \tau_{2,1} - \tau_{3,2} \end{aligned} \tag{16}$$

$$\partial_t (C_c h_c) + \partial_x (C_c u_c h_c) - \phi_{s\,3,2}^{net} = 0 \tag{17}$$

where C_c and $\rho_c = \rho^{(w)} (1 + (s - 1) C_c)$ are, respectively, the flux-averaged sediment concentration and the average density of layer [2], $\phi_{s\,3,2}^{net}$ is the net rate of mass exchanged between the transport layer and the bed. The global bed mass conservation equation is

$$\partial_t (Y_b) + \phi_{s3,2}^{net} = 0 \tag{18}$$

from which one deduces that $\phi_{3,2}^{net} = \phi_{s3,2}^{net}/(1 - p)$.

The integration of the equations (4) and (7) should be performed with some care. If the limits of integration are Y_b and $Y_s = Y_b + h_b$ the application of the Leibnitz rule renders

$$\int_{y=Y_b}^{y=Y_s} \partial_x \left(T_{yy}\right) \mathrm{d}y = \left| \int_{\mu=0}^{\mu=h_c} \partial_x \left(T_{yy}\right) \mathrm{d}\mu \right| + \partial_x (Y_b) \left(\rho_c h_c + \rho^{(w)} h_w\right) - \partial_x (Y_s)\, \rho^{(w)} h_w$$

The terms $\partial_x (Y_b) \left(\rho_c h_c + \rho^{(w)} h_w\right) - \partial_x (Y_s)\, \rho^{(w)} h_w$ account for pressure variations between two consecutive sections due to the variation of the channel profile at constant water level. They are related to the gravitic force in the sense that hydrostatic pressure variations are necessarily due to existence of a gravitic field. The true gravitic force is a mass force expressed by $g\rho_c h_c \sin(\beta)$ and by $g\rho_s h_s \sin(\beta)$. Since $\sin(\beta) = -\partial_x(Y_b)$ it appears that there is a duplication of the effect of the gravity in the equations of conservation of momentum. Works like Sieben (1999), on river flow, explicitly consider both terms. While this position can be sustained, it must be highlighted that if the $g\partial_x (Y_b) (\rho_c h_c + \rho_s h_s)$ is considered in equation (16), the convergence between unsteady flow and to steady uniform flow is precluded. In fact, the time and space derivatives of the velocity and of the water depth in equations (14) and (16) are eliminated it is obtained $g\rho_c h_c \sin(\beta) = \tau_{cb} - g\partial_x (Y_b) \left(\rho^{(w)} h_s + \rho_c h_c\right)$ instead of the correct result $g\rho_c h_c \sin(\beta) = \tau_b$. It is thus recommended that the integration of the gradient of the normal stresses should be performed with a change of variables to preserve the fact that the gradient of the pressure in a given section depends on the slope of the layer boundaries only inasmuch as these are included in the gradient of the water column.

2.4 *Closure equations for sheet-flow*

Equations for h_c, u_c, τ_b and $\phi^{net}_{s\,3,2}$ can and should be drawn from the same theoretical framework that produced the conservation equations. All four quantities are intimately interrelated. If the velocity and the shear stress are continuous across the boundary of layers [1] and [2] no closure equations are required for the shear stress at boundary and the vertical flux of momentum across that boundary. This issue will be further addressed in section 4.

The existence of a frictional layer across which the shear stress may vary allowed Ferreira (2005) p. 279 to determine the mass flux between the bed and the transport layer and the depth-averaged sediment concentration in the latter. The integration of the equation of conservation of momentum in the vertical direction over the frictional layer renders

$$\partial_t\,(Y_b) = \phi^{net}_{3,2} = \frac{g\rho^{(w)}(s-1)\tan(\varphi_b)}{u_c\,(\rho_c u_x)|_{z=Y_f}}\left(q_{sc} - q^*_{sc}\right) \tag{19}$$

where Y_f is the elevation of boundary between the frictional and collisional layers, $q_{sc} = C_c u_c h_c$ is the volumetric bedload discharge and $q^*_{sc} = C^*_c u_c h_c$ is the equilibrium volumetric bedload discharge. The equilibrium concentration is derived from the following considerations: i) by definition, one has $T(y=Y_b)=\tau_b \equiv \rho^{(w)} C_f\,u^2$, where $C_f = (u_*/u)^2$ is the friction factor of the overall flow; ii) in a steady, equilibrium flow one has $T(y=Y_b)=g\rho^{(w)}(s-1)C^*_c h_c \tan(\varphi_b)$, where C^*_b is the equilibrium concentration; hence, the equilibrium concentration is

$$C^*_b = \frac{C_f\,u^2}{g(s-1)\tan(\varphi_b)\,h_b} \tag{20}$$

The adaptation length implicit in (19) is dependent on the value of the longitudinal velocity at the top of the frictional layer. An estimate for this value can be obtained by considering the velocity profile in the transport layer. A good fit to the results in figure 2(a) is the equation

$$\frac{u(y)}{u_*} = 2.1\left(\frac{Z}{\theta}\right)^{\frac{3}{4}}\left(\frac{y}{h}\right)^{\frac{3}{4}}$$

adapted from Sumer *et al*. (1996), where $Z = h/d_s$. The resulting equation for the depth-averaged velocity is

$$u_b = u_*\tfrac{4}{7}2.1\left(Z/\theta\right)^{\frac{3}{4}}\left(h_b/h\right)^{\frac{3}{4}} \tag{21}$$

Depth-averaging the equation of conservation of the fluctuating energy, Ferreira (2005), pp. 278–287, obtained and algebraic relation for the thickness of the transport layer, graphically expressed in figure 3. In this figure, the theoretical lines are displayed along with the experimental results of Sumer *et al*. (1996). A sensitivity analysis to the values of the restitution coefficient is performed in figure 3a). Figure 3b) shows a different kind of sensitivity analysis: several parameters are varied, corresponding to the geometrical and mechanical characteristics of the sediment employed by Sumer *et al*. (1996).

The increase of the values of the coefficient of restitution results in an increase in the thickness of the transport layer. The most interesting result expressed in figure 3 is the invariance of h_c/d_s as a response to the variation of the type of sediment. If the proper values of the restitution coefficient and the internal friction angle at the bed are adopted, the non-dimensional thickness of the transport load layer is virtually independent from the type of sediment. The thickness of the transport layer can be approximated by

$$\frac{h_c}{d_s} = 1.7 + 5.5\theta \tag{22}$$

The results of Sumer *et al*. (1996) also allow for the computation of the friction factor. It was found that the bed shear stress can be adequately described by $\tau_b \equiv \rho^{(w)} C_f\,u^2$ provided that the friction coefficient is $C_f = C_f\,(d_s, h, w_s, u)$.

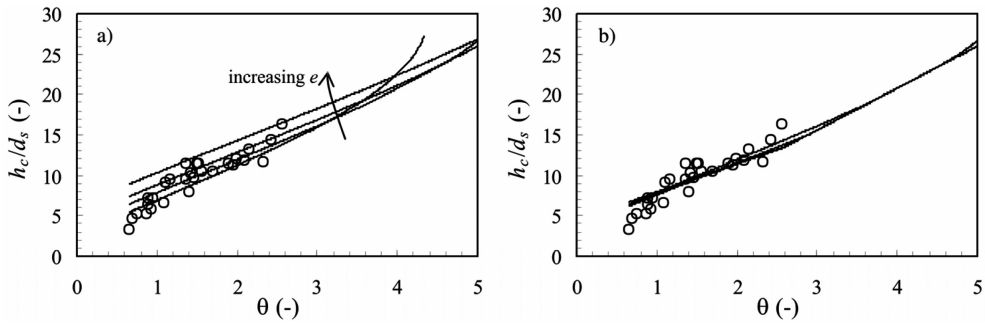


Figure 3. Thickness of the transport layer as a function of the Shields parameter. Sensitivity analysis to (a) the restitution coefficient; $e = 0.5$, 0.8, 0.9 and 0.95; $\tan(\varphi_b) = 0.35$; $\nu_0 = 0.65$; $q = 1.0$ m^2/s and (b) the characteristics of the sediment grains. The four types of granular material used by Sumer *et al*. (1996), were plastic ($d_s = 0.003$ m, $s = 1.27$, $e = 0.75$ and $d_s = 0.0026$ m, $s = 1.14$, $e = 0.75$), acrylic ($d_s = 0.0006$ m, $s = 1.13$, $e = 0.75$) and sand ($d_s = 0.00013$ m, $s = 2.67$, $e = 0.8$); $\nu_0 = 0.6$; $\tan(\varphi_b) = 0.4$; $q = 1.0$ m^2/s. Circles (o) stand for Sumer's *et al*. (1996) data.

3 ANALYSIS OF THE EQUATIONS

3.1 *Hyperbolicity and eigenstructure*

Once the conceptual model is derived, its mathematical properties must be investigated. In particular, the nature of the system of PDEs must be characterized with respect to linearity and hyperbolicity and also in what concerns its eigenstructure.

Experimental and numerical evidence can be drawn to support the assumption that velocity and shear stresses are continuous across the boundary of layers [1] and [2] (see section 2.2 and details in Ferreira 2005, §3). Equations (13) and (15) (mass) and (14) and (16) momentum can be added and the system of conservation laws becomes

$$\partial_t (h) + \partial_x (uh) = \phi^{net}_{3,2} \tag{23}$$

$$\partial_t (\rho_m uh) + \partial_x \left(\rho^{(w)} u_w^2 h_w + \rho_c u_c^2 h_c\right) + \tfrac{1}{2} g \partial_x \left(\rho^{(w)} h_w^2 + 2\rho^{(w)} h_w h_c + \rho_c h_c^2\right) + g\left(\rho^{(w)} h_w + \rho_c h_c\right) \partial_x (Y_b) = -\tau_b \tag{24}$$

$$\partial_t (C_c h_c) + \partial_x (C_c u_c h_c) = \phi^{net}_{s\,3,2} \tag{25}$$

$$\partial_t (Y_b) = \phi^{net}_{s3,2} \tag{26}$$

where $\rho_m = \rho^{(w)} (1 + (s - 1)C)$ is the layer-averaged flow density and $C = C_c u_c h_c / uh$ is the layer-averaged sediment concentration.

If space-lag and time-lag effects are small, *i.e*. if, given an imposed imbalance between sediment load and transport capacity, the flow is able to adjust the sediment load to the capacity value within a length of the order of magnitude of the flow depth, an equilibrium formulation can be used. The conservation equations become

$$\partial_t (Y) + \partial_x (uh) = 0 \tag{27}$$

$$\partial_t (\rho_m uh) + \partial_x \left(\rho^{(w)} u_w^2 h_w + \rho_c u_c^2 h_c\right) + \tfrac{1}{2} g \partial_x \left(\rho^{(w)} h_w^2 + 2\rho^{(w)} h_w h_c + \rho_c h_c^2\right) + g\left(\rho^{(w)} h_w + \rho_c h_c\right) \partial_x (Y_b) = -\tau_b \tag{28}$$

$$(1-p)\partial_t(Y_b) + \partial_t(C_c h_c) + \partial_x(C_c u_c h_c) = 0 \tag{29}$$

where $H_b = Y_b(1-p) + C_c h_c$. Equations (23) to (25) (if the discretization scheme is explicit, equation 26 may be solved independently from the remaining) and (27) to (29) are autonomous quasi-linear systems of PDEs and may be written in the form

$$\mathbf{A}\,\partial_t(\mathbf{V}) + \mathbf{B}\partial_x(\mathbf{V}) = \mathbf{G}(\mathbf{V}) \tag{30}$$

where $\mathbf{V} = [h \quad u \quad C_c]^T$ (equations 23 to 25) or $\mathbf{V} = [h \quad u \quad Y_b]^T$ (equations 27 to 29) is the vector of primitive variables, $\mathbf{A} = \partial_{\mathbf{V}}(\mathbf{U})$, $\mathbf{B} = \partial_{\mathbf{V}}(\mathbf{F}) + \mathbf{A}^*$, $\mathbf{U} = [h \quad \rho_m uh \quad h_c C_c]^T$ (equations 23 to 25) or $\mathbf{U} = [Y \quad \rho_m uh \quad H_b]^T$ (equations 27 to 29) is the vector of conservative variables, $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ is the conservative flux, defined as

$$\mathbf{F} = \begin{bmatrix} uh \\ \rho_c u_c^2 h_c + \rho^{(w)} u_w^2 h_w + \frac{1}{2} g\left(\rho^{(w)} h_w^2 + 2\rho^{(w)} h_w h_c + \rho_c h_c^2\right) \\ C_c u_c h_c \end{bmatrix}$$

$\mathbf{G} : \mathbb{R}^3 \to \mathbb{R}^3$ is the vector of source terms, defined as

$$\mathbf{G}(\mathbf{V}) = \left[\begin{array}{ccc} \phi_{3,2}^{net} & -\tau_b & \phi_{s3,2}^{net} \end{array}\right]^{\mathrm{T}} \text{ or } \mathbf{G}(\mathbf{V}) = \left[\begin{array}{ccc} 0 & -\tau_b & 0 \end{array}\right]^{\mathrm{T}}$$

and $\mathbf{A}^*$ is such that $\mathbf{A}^*_{23} = g\left(\rho^{(w)} h_w + \rho_c h_c\right)$ and $\mathbf{A}^*_{ij} = 0$ for all other values of i and j.

The non-conservative flux $f_b = g\left(\rho^{(w)} h_w + \rho_c h_c\right)$ may raise some concerns about the overall conservative nature of the system of conservation laws. If the system is conservative, then

$$\int_{-\infty}^{-\infty} \mathbf{U}(\xi, t)\,\mathrm{d}\xi = \mathbf{m}$$

where $\mathbf{m}$ is independent of t. In what concerns the non-conservative flux, this requires that

$$\int_0^t \int_{-\infty}^{+\infty} f_b \partial_\xi(Y_b)\,\mathrm{d}\xi\mathrm{d}\tau = 0 \tag{31}$$

Ferreira (2005), p. 346, proves a necessary condition for (31) for the case of system (39) to (41). It can be further demonstrated that the system is conservative if $(h,u) : \mathbb{R} \times]0, +\infty[\to \mathbb{R}^2$ is piecewise continuous with compact support, $f_b(h(x,t), u(x,t))$ is smooth and of bounded variation, $Y_b(x,t)$ is Lipschitz continuous a.e. with compact support and if Y_b is independent of (u,h). It is easy to verify that $f_b(h(x,t), u(x,t))$ is indeed smooth; since continuity of Y_b, u or h is not required, it is concluded that the system is conservative.

In what concerns hyperbolicity, the nature of the conservation systems is investigated through the characteristic polynomial $|\mathbf{B} - \lambda\mathbf{A}| = 0$, where the operator $|\cdot|$ stands for the determinant operator. The characteristic polynomial can be written as

$$\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0 \tag{32}$$

where the coefficients a_1, a_2 and a_3 depend on u, h and, for system (23) to (25), on C_c, and on its derivatives and are required to be $C^1(\mathbb{R})$. Expressions for these coefficients are shown in Ferreira (2005), p. 348. For both systems, the characteristic polynomial (32) possesses three real and different roots, such that $\lambda^{(1)} > \lambda^{(2)} > \lambda^{(3)}$, called the characteristics of the system. Consequently, there are three independent eigenvalues associated to each of the $\lambda^{(k)}$ eigenvalues which form the base of eigenspace whose dimension is equal to the rank of $\mathbf{B} - \lambda\mathbf{A}$. It is thus concluded that the systems are strictly hyperbolic (*cf. e.g.*, LeVeque 1990). The characteristics can be plotted as a function of the Froude number.

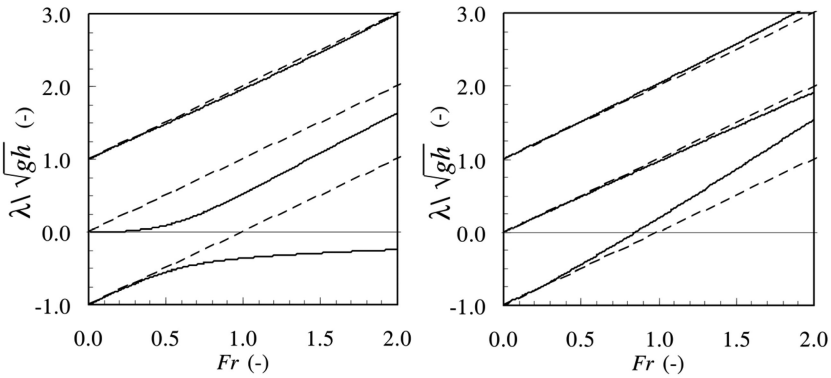


Figure 4. Eigenvalues of systems (27) to (29) (left) and (23) to (25) (right) as a function of the Froude number. Variation of velocity and concentration with the water depth obtained for the flow across a centered rarefaction wave. Dashed lines (- - -) stand for $Fr+1$ (higher), Fr (middle) and $Fr-1$ (lower). Calculations for $s-1=0.5$.

Figure 4 (right) shows the characteristics corresponding to system (23) to (25), for low sediment concentrations; Figure 4 (left) shows the characteristics corresponding to system (27) to (29). In the latter case, the characteristics are such that $\lambda^{(1)}>\lambda^{(2)}>0$ and that $\lambda^{(3)}$ is always negative. The magnitude of the characteristics depends on the mobility of the sediment. Increasing the mobility of the sediment it is observed that characteristic fields $\lambda^{(2)}$ and $\lambda^{(3)}$ are the most affected. The characteristics of system of equations (23) to (25), seen in figure 4(right), are such that $\lambda^{(1)}>\lambda^{(2)}>\lambda^{(3)}$ but $\lambda^{(3)}$ can be positive or negative, depending on the value of the Froude number and the sediment concentration. In case, $\lambda^{(2)}/(gh)^{0.5}$ is approximately equal to Fr; in fact, this result should be interpreted with care: the velocity of propagation of information relative to bedload corresponding to the homogeneous problem is indeed $\lambda^{(2)}$; however, source terms can dramatically modify the actual velocity of propagation.

3.2 *A theoretical solution for the Riemann problem*

A solution for the Riemann problem

$$\partial_t(\mathbf{U})+\partial_x(\mathbf{F}(\mathbf{U}))=\mathbf{0} \tag{33}$$

$$\mathbf{U}_0\equiv\mathbf{U}(x,0)=\begin{cases}\mathbf{U}_{\mathrm{L}}, & x\le x_0\\ \mathbf{U}_{\mathrm{R}}, & x>x_0\end{cases} \tag{34}$$

can be devised for system of equations (27) to (29) (Fraccarollo & Capart 2002, Ferreira 2005, §5). The solution is not unique since it depends on the type of linearization for the non-conservative flux, f_b, across discontinuities. Ferreira (2005), p. 368, following LeFloch (1996) proposed $\langle f_b\rangle=\frac{1}{2}\left\{f_b\left(h^+,u^+;\cdot\right)+f_b\left(h^-,u^-;\cdot\right)\right\}$.

From the theorem of Lax (1957) the solution is composed of three waves (because the dimension of the eigenspace is three) that can be rarefaction waves or shocks, if all characteristic spaces are genuinely non-linear, or contact waves, if the characteristic spaces are linearly degenerate. Hence, before attempting to calculate the solution, it is necessary to know the type of non-linearity of the characteristic spaces. A characteristic field is genuinely non-linear if

$$\partial_{\mathbf{V}}\left(\lambda^{(k)}\right)\cdot\mathbf{r}^{(k)}\neq 0\quad\forall_{\mathbf{V}\in\Delta_{hu}} \tag{35}$$

The left-hand side of (35) is the directional derivative of the characteristic in the direction of the correspondent $\mathbf{r}$, a right eigenvector. It should be recalled that $\mathbf{r}^{(k)}=\mathrm{d}\mathbf{V}/\mathrm{d}\mathbf{V}\mathrm{d}W^{(k)}$, where $W^{(k)}$ is

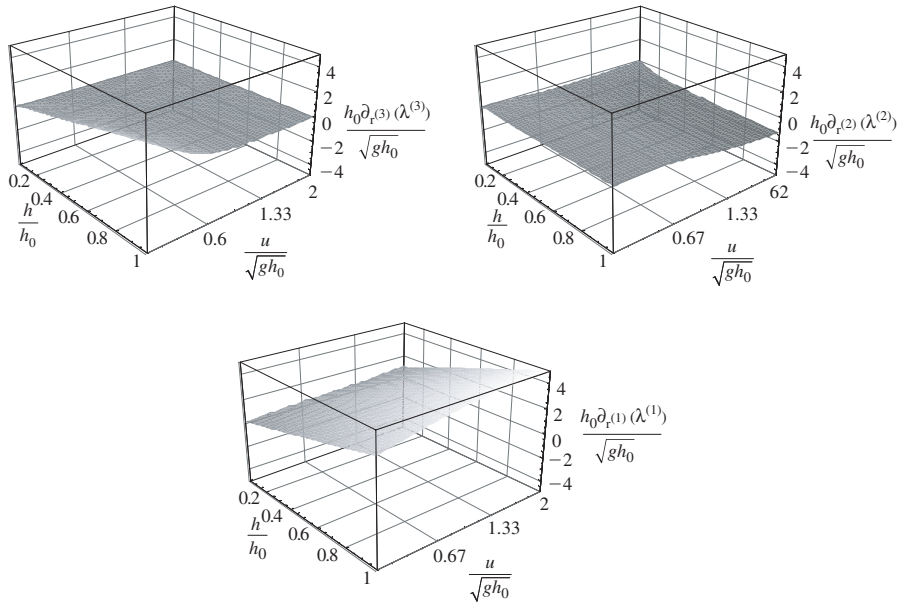


Figure 5. Surface and density plots of $\mathrm{E}^{(k)} = h_0\partial_{\mathbf{V}}(\lambda^{(k)})\cdot\mathbf{r}^{(k)}/\sqrt{gh_0}$: (a) $k=1$, (b) $k=2$, (c) $k=3$, thick line in the density plot of stands for $\mathrm{E}^{(3)}=0$. Computations performed with $C_{f0}=0.1$, $\tan(\varphi_b)=0.5$, $s=2.65$ and $d_s=0.003$ m. Scale:

the characteristic variable corresponding to the $\lambda^{(k)}$-field, such that $\partial_t\left(W^{(k)}\right)+\lambda^{(k)}\partial_t\left(W^{(k)}\right)=0$. Hence, condition (35) ensures that across a k-wave, the path of the solution in the domain Δ_{hu} will not be overtaken by a line of local maxima or minima of $\lambda^{(k)}$. A plot of $\partial_{\mathbf{V}}\left(\lambda^{(k)}\right)\cdot\mathbf{r}^{(k)}$ is shown in figure 5. It is observed that condition (35) holds for both the first and the second characteristic fields, as seen in figures 5(a) and (b).

In fact, except for a limited region near $h=0$ outside figure 5(a), $\partial_{\mathbf{V}}\left(\lambda^{(1)}\right)\cdot\mathbf{r}^{(1)}>0$ and $\partial_{\mathbf{V}}\left(\lambda^{(1)}\right)\cdot\mathbf{r}^{(1)}<0$. As for the third characteristic field, it can be shown that there is a line in Δ_{hu} where $\partial_{\mathbf{V}}\left(\lambda^{(3)}\right)\cdot\mathbf{r}^{(3)}=0$. However, Ferreira (2005), p. 357, demonstrated that when the Riemann problem takes the form of the dam-break problem, the third characteristic field does not develop contact discontinuities. All fields are, thus, genuinely non-linear.

If $\partial_{\mathbf{V}}\left(\lambda^{(k)}\right)\cdot\mathbf{r}^{(k)}$ is taken as a measure of the degree of the non-linearity of the field, it can be concluded from figure 5 that the strength of the 1-waves is expected to be larger than that of the 2-and 3- waves. For instance, if a shock develops as a solution of the $\lambda^{(2)}$-field, its strength is expected to be lower than the strength of a shock in the $\lambda^{(1)}$-field.

Ferreira (2005), pp. 363–366, showed that there are two wave structures depending on the values of the initial condition, namely the strength of the initial discontinuity in the bed and the difference in the upstream and downstream values of the water elevation. As seen in figure 6 (left), one, previously unnoticed, possible solution is composed of two shock waves and one rarefaction (or expansion) wave. This solution occurs when the initial downstream water depth is sufficiently large and when there is an initial discontinuity in the bed elevation. The second type of solution is that presented by Fraccarrollo & Capart (2002) and comprises two rarefaction waves and one shock (figure 6, right).

In order to determine the solution for a particular Riemann problem, it is sufficient to use the relations that involve the unknown quantities of the constant states upstream and downstream which bound a given wave, according to Lax's (1957) theorem. Across a rarefaction wave, the Riemann

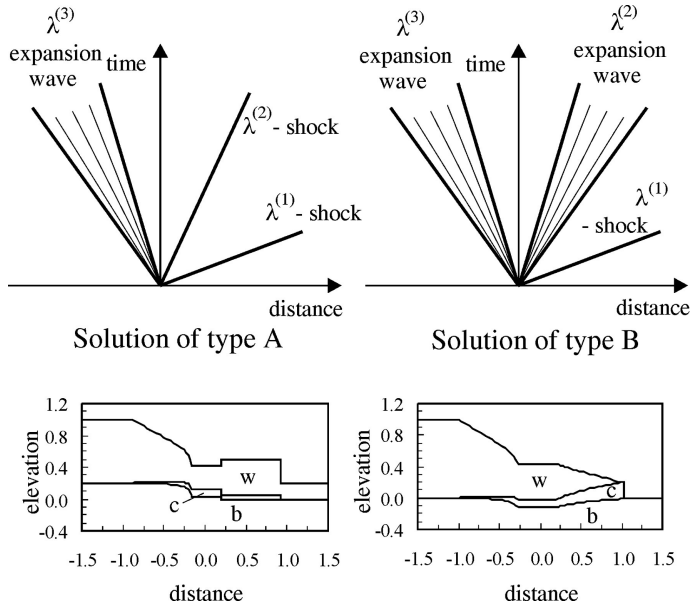


Figure 6. Possible wave structures of the Riemann solution for the geomorphic dam-break problem. Solution for $\lambda^{(2)}$-field: Left – shock wave, type A; right – rarefaction wave, type B.

invariants are computed from the differential equations

$$\frac{\mathrm{d}u}{\mathrm{d}h} = \frac{r_2^{(k)}}{r_1^{(k)}} = f_u^{(k)}(u, h; \cdot), \qquad \frac{\mathrm{d}Y_b}{\mathrm{d}h} = \frac{r_3^{(k)}}{r_1^{(k)}} = f_z^{(k)}(u, h; \cdot) \tag{36a,b}$$

Across the shocks the Rankine-Hugoniot conditions are applied

$$\left(\mathbf{U}_{\text{downstream}} - \mathbf{U}_{\text{upstream}}\right) S_k = \left(\mathbf{F}_{\text{downstream}} - \mathbf{F}_{\text{upstream}}\right) \tag{37}$$

where S_k is the velocity of shock k. Expressions for the eigenvectors and for the functions of equations (36) can be found in Ferreira (2005), pp. 367 and following.

Computed solutions and experimental results are shown in figure 7 (top: solution of type A; bottom: solution of type B) in terms of flow profiles. Both the experimental profiles and the theoretical solutions are shown in self-similar co-ordinates. It is observed that the thickness of the contact load layer and the minimum bed elevation are correctly estimated in almost all tests. Naturally, phenomena pertaining to soil mechanics, namely slope stability and en masse bed movement, could not have been reproduced by the theoretical solution since the governing equations do not incorporate formulations for those phenomena.

The effect of flow resistance is perceived in the observed profiles inasmuch they are not truly self-similar. This is particularly true for the flows depicted in figures 7(c) and 7(d). Furthermore, it is noticed that Froude similarity does not completely determine the flow profiles. This means that the solution is a function of the sediment density, diameter, and internal friction angle, other than initial conditions (Ferreira *et al*. 2006b).

While the theoretical solution is relevant to assess the scales of the phenomena and to provide a measure to the accuracy of numerical discretization schemes, engineering predicting tools for real geomorphic flows must be performed with numerical codes based on the full conservation equations. In the next section, numerical simulations of real geomorphic flows are presented.

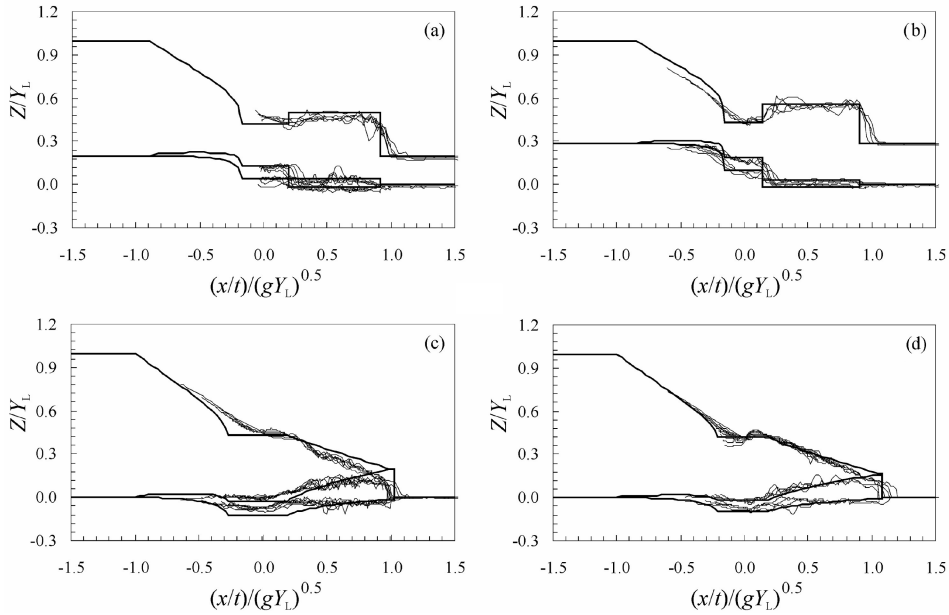


Figure 7. Flow profiles corresponding to dam-break laboratory tests performed in the Laboratory of Fluid Mechanics of the Université catholique de Louvain (details in Ferreira *et al.* 2006b, Spinewine 2005). Top: positive initial bed step, solution of type A; bottom: flat initial bed, solution of type B. Solid thick lines (——) stand for the theoretical solution; thin black lines (——) stand for experimental results (adapted from Ferreira *et al.* 2006b).

4 NUMERICAL SOLUTIONS

4.1 *Discretization techniques*

The solution procedure is presented in this text is based on a finite-difference explicit discretization closely following MacCormack's (1969) scheme. This widely used Lax-Wendroff type of scheme is second order in space and first order in time. Its properties are analysed in Hirsch 1990, pp. 239–245.

MacCormack's scheme is a fractional-step, second order, explicit, shock-capturing scheme widely used in hyperbolic problems, especially in the study of the dynamic of compressible fluids. The discretization of the geomorphic shallow-water equations (23)–(25) and (27)–(29) is straightforward – no adaptation is required since the system is fully hyperbolic such as the Euler/continuity set of equations. According to Garcia & Kahawita (1986), regressive and progressive differences should be alternated in the predictor and corrector steps, in order to better distribute the phase error and, thus, minimise the inevitable oscillations aroused by the second order precision in space. Hence, the discretization of the equations is performed in the following manner. In involves a predictor step,

$$\widehat{\mathbf{U}}_i = \mathbf{U}_i^n - \frac{\Delta t}{\Delta x}(\mathbf{F}_{i+1}^n - \mathbf{F}_i^n) + \Delta t \mathbf{G}_i^n \ \ (n \text{ odd}); \ \ \widehat{\mathbf{U}}_i = U_i^n - \frac{\Delta t}{\Delta x}(\mathbf{F}_i^n - \mathbf{F}_{i-1}^n) + \Delta t \mathbf{G}_i^n (n \text{ even}),$$

a corrector step,

$$\widehat{\widehat{\mathbf{U}}}_i = \mathbf{U}_i^n - \frac{\Delta t}{\Delta x}(\widehat{\mathbf{F}}_i - \widehat{\mathbf{F}}_{i-1}) + \Delta t \widehat{\mathbf{G}}_i \ (n \text{ odd}); \ \widehat{\widehat{\mathbf{U}}}_i = \mathbf{U}_i^n - \frac{\Delta t}{\Delta x}(\widehat{\mathbf{F}}_{i+1} - \widehat{\mathbf{F}}_i) + \Delta t \widehat{\mathbf{G}}_i \ (n \text{ even}),$$

where $\widehat{\mathbf{F}}_k = \mathbf{F}(\widehat{\mathbf{U}}_k)$ and $\widehat{\mathbf{G}}_k = \mathbf{G}(\widehat{\mathbf{U}}_k)$, and the update

$$\mathbf{U}_i^{n+1} = \tfrac{1}{2}(\widehat{\mathbf{U}}_i + \widehat{\widehat{\mathbf{U}}}_i) \tag{38}$$

In spite of the alternate regressive/progressive difference procedure, spurious oscillations near discontinuities still occur in second-order schemes. In order to achieve a monotone solution at discontinuities, TVD versions of MacCormack's scheme have been proposed, enhanced the scheme with a self-adaptive artificial viscosity that inhibits the formation of oscillations.

The TVD algorithm calls for building an approximate Jacobian matrix, $\tilde{\mathbf{J}}_{i+1/2}$, in order to obtain a linearised system at the boundary of two computational cells. This procedure was first developed by Roe (1981) and it is easy to derive for the clear-water shallow-water equations but is very difficult to obtain if the vectors $\mathbf{F}_{i+1/2}$ are those of the geomorphic shallow-water equations (Ferreira 2005, §5, Leal 2005). Indeed, the building of such an approximate matrix would require algebraic manipulations and solutions of highly non-linear equations, involving, in particular, the expression of the characteristics. Hence, in this work, the TVD correction is applied to the equations of conservation of total mass and momentum only. The formulæ used are those derived to the quasi-linear hyperbolic clear-water shallow-water equations (details in Ferreira 2005, §5, or Leal *et al.* 2003). The results of the TVD step are added to the classical MacCormack scheme, thus redefining the update step, which becomes

$$\mathbf{V}_i^{n+1} = \tfrac{1}{2}(\widehat{\mathbf{V}}_i + \widehat{\widehat{\mathbf{V}}}_i) + \frac{\Delta t}{\Delta x}\left(\mathbf{D}^n_{i+\frac{1}{2}} - \mathbf{D}^n_{i-\frac{1}{2}}\right) \tag{39}$$

where vector $\mathbf{D}$ is

$$\mathbf{D}^n_{i+\frac{1}{2}} = \tfrac{1}{2}\sum_{k=1,2}\left(1-\phi\left(a^{(k)}_{i+\frac{1}{2}}\right)\right)\Psi^{(k)}_{i+\frac{1}{2}-sgn}\left(1-\left|\tilde{C}r^{(k)}_{i+\frac{1}{2}}\right|\right)\alpha^{(k)}_{i+\frac{1}{2}}\tilde{\mathbf{r}}^{(k)}_{i+\frac{1}{2}} \tag{40}$$

the function $\phi(a^{(k)}_{i+1/2})$ stands for the flux limiter (see details *in* Ferreira 2005, p. 440). The argument of the flux limiter function (the inverse of ratio of consecutive flux differences) was defined as (Hirsch 1990, p. 551)

$$a^{(k)}_{i+\frac{1}{2}} = \left\{\tilde{\lambda}^{(k)}_{i+\frac{1}{2}-sgn}\left(1-\left|\tilde{C}r^{(k)}_{i+\frac{1}{2}-sgn}\right|\right)\alpha^{(k)}_{i+\frac{1}{2}-sgn}\right\} \Big/ \left\{\tilde{\lambda}^{(k)}_{i+\frac{1}{2}}\left(1-\left|\tilde{C}r^{(k)}_{i+\frac{1}{2}}\right|\right)\alpha^{(k)}_{i+\frac{1}{2}}\right\}$$

and incorporates Roe's (1981) approximations for depth-averaged flow velocity, u, and for the shallow-water propagation speed, c (*in* Ferreira 2005, p. 440). The signal sgn is positive/negative if the approximate characteristic is positive/negative. If $\tilde{\lambda}^{(2)}_{i+1/2} \approx 0$, *i.e.*, at critical flow points, the TVD correction is ineffective to hamper the non-monotone behaviour of the solution. At that point a discontinuity is expected to develop. Indeed, as pointed out by Roe (1981), the use of the approximate Jacobian can lead to unsound physical solutions such as stationary, entropy violating, shocks. In order to avoid these sonic shocks, notably those occurring in expansion waves, entropy corrections, namely those of Harten & Hyman, are prescribed (Hirsh 1990). Its application involves the substitution of $\left|\tilde{\lambda}^{(k)}_{i\pm1/2}\right|$ by a small number ε whenever its value is less than ε. Hence, the sonic point is displaced within a ball of radius ε. This promotes the transfer of information across the sonic point, *i.e.*, it diffuses the shock into small rarefaction (expansion) fans whose amplitude is ε, thus enforcing Oleinik's E condition (Ferreira 2005, p. 334). Such small expansion fan, represented by the function Ψ in equation (40), is of the order of magnitude of $\varepsilon_{i+\frac{1}{2}} = \max\{0, \tilde{\lambda}^{(k)}_{i+1/2} - \lambda(\mathbf{V}_i), \lambda(\mathbf{V}_{i+1}) - \tilde{\lambda}^{(k)}_{i+1/2}\}$.

The algorithm is complete with the inclusion of Jameson's Artificial viscosity (Hirsch 1990, p. 89) the equations of conservation of sediment mass (equations 25 and 29).

It is noted that the solution procedure is carried out for the conservative variables $Y = Y_b + h$, $R_m = \rho_m uh$ and $Z_b = (1-p)Y_b + C_c h_c$, in the case of the capacity model (equations 23–25), and $Y = Y_b + h$, $R_m = \rho_m uh$ and $H_c = C_c h_c$ in the case of the non-equilibrium model (equations 27–29). The primitive variables must be computed from these. In the case of the equilibrium model this implies solving the system of non-linear equations

$$R_m - \rho^{(w)}\left(1 + (s-1)\,C(h,u)\right)uh = 0 \tag{41}$$

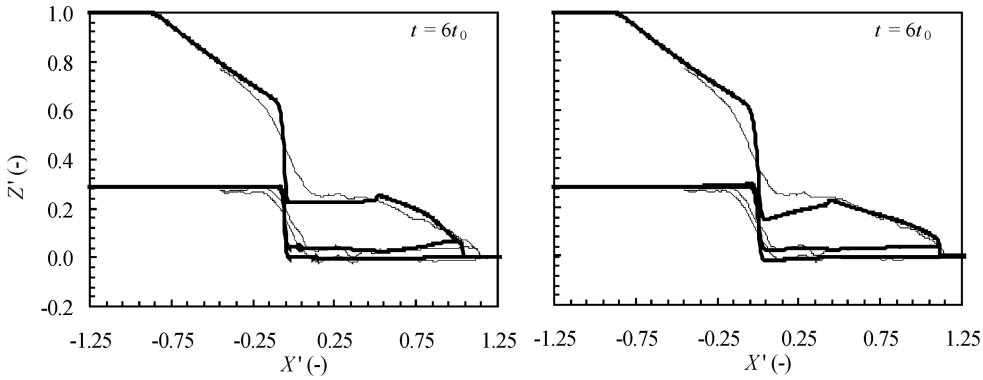


Figure 8. Comparison of experimental (___) and computed flow profiles (▬) for an initially discontinuous bed. Sand is the bed material. Left: results of the capacity model (equations 23–25); right: results of the non-equilibrium model (equations 27–29). Solutions plotted in a self-similar referential with $X' = (x/t)/\sqrt{gY_0}$ and $Y_0 = Y_L$.

$$Z_b - (1-p)(Y-h) + C_c(u)\,h_c(u) = 0 \tag{42}$$

with u and h as unknowns. The solution is obtained by a combination of Newton's method and fixed point iteration. In the case of the non-equilibrium model, the velocity is obtained from

$$R_m - \rho^{(w)}\,(uh + (s-1)\,C_c h_c(u) u_c(u)) = 0 \tag{43}$$

The sediment concentration is then retrieved from $C_c = H_c/h_c(u)$.

4.2 *Numerical simulations of geomorphic flows*

4.2.1 *Simulation of laboratory dam-break flood wave propagation over mobile beds (UCL)*

The results of the application of the above equations and algorithmic procedures are shown in figure 8 along with experimental results of Spinewine & Zech (2003), performed at the Laboratory of Fluid Mechanics of the Université Catholique de Louvain (UCL). Since this is a pure initial value problem, extrapolation boundary conditions (Hirsch 1990, p. 89) are used in both ends of the computational domain.

It is clear in the simulation results that the effect of the flow resistance is uniformly distributed over the time: the numerical waves of both solutions are slowed down at constant rate. The solution of the non-equilibrium model features a faster wave-front than the solution of the capacity model. Note that the friction coefficient was calibrated so that the solution of the non-equilibrium model would match the experimental result.

For this friction factor formulation, the numerical solution of the capacity model represents a good prediction of the observed water depths but not of the wave front velocity. It is also noted that the greater phenomenological complexity of the non-equilibrium model does not appear to be reflected in better prediction capabilities. This indicates that these flows may be able to promptly adapt to equilibrium conditions if the bed sediment are dense (*cf.* Ferreira 2005, §5).

4.2.2 *Numerical simulation of the 1996 flood in River Ha! Ha!, Quebéc, Canada*

Severe rainstorms scourged the Saguenay region, south Québec, Canada, between the 18th and the 21st of July 1996. Due to the overtopping and sequent failure of an earthfill dyke, River Ha! Ha! experienced a significative increase in the peak flood discharge. The dam-break wave, superimposed to the hydrologic flood, provoked massive geomorphic impacts in the downstream valley (Brooks & Lawrence 1999, Lapointe *et al*. 1998, Capart *et al.* 2003). The simulation of the full length of River Ha! Ha! is shown in Ferreira *et al.* (2005). A summary of the most relevant simulation

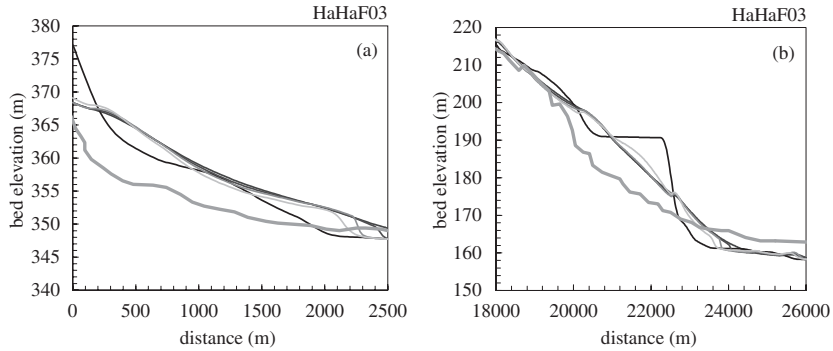


Figure 9. Morphologic evolution of the River Ha! Ha!. Black line (___) stands for the initial bed profile; grey thick line (___) stands for the field data expressing the final bed profile. Computed profiles at $t = 26$ h (___), $t = 32$ h (___), $t = 40\,h$ (___) and $t = 67.5$ h (___).

results is presented here. The conservation equations included in the model are a generalization of system (23–25) for non-prismatic channels. The TVD algorithm revealed inadequate to tackle oscillations born from strong momentum sources such as abrupt slope variations (details in Ferreira 2005, p. 000). Von Neumann's numerical viscosity (Von Neumann & Richtmyer 1950) was used in its place.

The pre- and post-flood measured bed profiles of River Ha! Ha! can be seen in figure 9, in the reaches were more important morphologic changes have occurred. The computed profiles of the bed elevation, corresponding to the rising limb of the discharge hydrograph, can be seen in the same figure. The magnitude of the morphologic impacts is impressive.

The dyke was completely eroded and so were the sediment deposits bellow it down to a depth of about 14 m bellow the initial bed elevation (figure 9a). Massive incision was registered in all convex profiles. In particular, in a chute at kilometre 22 the flood excavated a new channel around the rock outcrop (Lapointe *et al.* 1998) and severe upstream-progressing erosion took place (figure 9c). The new bed is, at some places, 20 m bellow the old river bed.

The model can deal with different flow regimes within the computational domain. In figure 10 – Froude number profiles –, up to eleven shocks (hydraulic jumps) and ten sonic points (or critical flow points) are observed at the end of the "warm up" and preliminary runs, *i.e.*, at $t = 0$ h (details in Ferreira 2005, p. 509–523).

Most of the hydraulic jumps and sonic points disappear during the rising limb of the hydrograph. In the same period, one new sonic point and one new hydraulic jump were created at the transition between fixed and mobile bed reaches (kilometre 15). It is noted that the upstream migration of the sonic points is due, as in the case of the knickpoint migration, to the upstream erosion of the point of minimum curvature.

The hydraulic jumps identified in figure 10, progress in the downstream direction. This is a consequence of the wave structure of system (23–25) shown in section 3.2: in mobile bed flows, discontinuities cannot be purely hydrodynamic because the bed elevation or the sediment concentration are dependent variables of the system of conservation laws. These hydraulic jumps occur in the characteristic wave associated to $\lambda^{(2)}$ and, hence, are associated to a discontinuity in the bed elevation that moves necessarily downstream at a slow velocity until eventually disappearing.

5 CONCLUSION

The fundamental premise of this text was that advances in computational fluvial hydraulics are closely linked to progresses in the characterization of fluvial phenomena. Bearing this principle in mind, a computational tool suitable to model highly unsteady geomorphic flows, *i.e.* flows

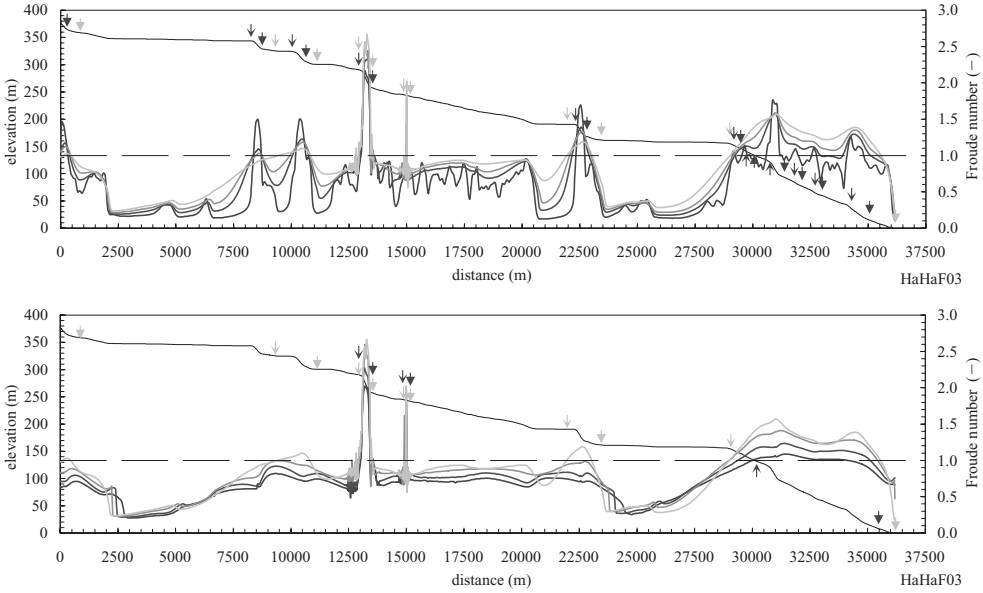


Figure 10. Froude number profiles. a) rising limb of the discharge hydrograph. Profiles at $t = 0$ h (___), $t = 14$ h (___), $t = 20$ h (___) and $t = 26$ h (___). b) receding limb; profiles at $t = 26$ h (___), $t = 32$ h (___), $t = 40$ h (___) and $t = 67.5$ h (___). Sonic points signalled by ↓; hydraulic jumps signalled by ↓. Black line (___) stands for the initial bed.

associated with severe morphological impacts, within an engineering framework was presented. It was shown that a proper articulation, at different stages of development, between empirical characterization, theoretical formulation and numerical simulation, is the key factor for the development of a sound conceptual model. Indeed, the theoretical characterization of the transport layer of geomorphic flows, based on extensive empirical and experimental work on granular and granular-fluid mixtures, performed in section 2, complemented with the numerical simulation of the two-dimensional vertical flow structure, allowed for the formulation of a successful one-dimensional conceptual model susceptible to become the core of an engineering predicting tool.

A proper mathematical analysis of the system of conservation equations allows for a better understanding of the simulation results, as exemplified with the characterization of the moving hydraulic jumps. Furthermore, the theoretical solutions drawn from this mathematical analysis, allow for the assessment of the quality of the numerical discretization schemes.

The proposed real-case modelling exercise was extremely demanding. The model was able to find non-oscillatory solutions even when subjected to important inputs of mass and momentum. One of the most accomplished characteristics of the model is its ability to deal with different flow regimes within the computational domain. In addition, the simulation reveals that the model can cope with the elimination of flow structures such as hydraulic jumps as a result of the profound morphologic changes. This shows that, despite being fast replaced by finite volume methods, for instance relying on HLLC solvers for Godunov fluxes, methods based on non-oscillatory second order finite difference schemes are still a reliable option for the simulation of essentially one-dimensional river phenomena.

REFERENCES

Alcrudo, F. & Garcia-Navarro, P., 1993. A High Resolution Godunov Type Scheme in Finite Volumes for the 2-D Shallow Water Equations. *J. Int. Methods in Fluids*. **16**: 489–505.

Armanini, A., Capart, H. Fraccarollo, L. & Larcher, M., 2004. Rheological stratification in experimental free-surface flows of granular–liquid mixtures. *J. Fluid. Mech*. **532**: 269–319.

Bermudez, A. & Vazquez, M. E. 1994. Upwind methods for hyperbolic conservation laws with source terms. *Computers & Fluids*, **23**: 1049–1071.

Brooks, G. R. & Lawrence, D. E., 1999. The drainage of the Lake Ha! Ha! reservoir and downstream geomorphic impacts along Ha! Ha! River, Saguenay area, Quebec, Canada. *Geomorphology*, **28**: 141–168.

Brufau, P., García-Navarro, P., Ghilardi, P., Natale, L. & Saviron, F., 2000. 1D Mathematical modelling of debris flow. *J. of Hydraul. Res.*, **38:** 435–446.

Caleffi, V. & Valiani, A. (2002) A mathematical model for dam-break over movable bed. *River Flow 2002*. Louvain-la-Neuve, Vol 2: pp. 991–1001.

Cao, Z., Day, R. & Egashira, S. 2002 Coupled and decoupled numerical modelling of flow and morphological evolution in alluvial rivers. *J. Hydraul. Eng.*, **128**(3): 306–321.

Cao, Z., Pender, G., Wallis, S. & Carling, P. 2004 Computational dam-break hydraulics over erodible sediment bed. *J. Hydraul. Eng.*, **130**(7): 689–703.

Capart, H., Spinewine, B., Young, D. L., Zech Y., Brooks, G. R., Leclerc, M. & Secretan, Y. 2003. The 1996 Lake Ha! Ha! breakout flood, Québec: Proposed test case for geomorphic flood models. *3rd IMPACT Workshop*, Louvain-la-Neuve, Belgium, 5–7 November (Cd-Rom).

Capart, H., Young, D. L., 1998 Formation of a jump by the dam-break wave over a granular bed. *J. of Fluid Mech.*, **372**: 165–187.

Chapman, S. & Cowling, T. G., 1970. *The Mathematical Theory of Non-Uniform Gases*, 3rd Edn. Cambridge University Press.

Correia, L. P., Krishappan, B. G. & Graf, W. H., 1992 Fully coupled unsteady mobile boundary flow model. *J. Hydraul. Eng.*, **118**(3): 476–494.

Cunge, J. A., Holly, Jr., F. M., Verwey, A., 1980 *Pratical aspects of computacional river hydraulics*. Pitman Advanced Publishing Program, Pitman Publishing Co., London, 420 pp.

de Vries, M., 1965. Considerations about non-steady bed-load transport in open-channels. *Proc. 11th Int. Congress Int. Assoc. for Hydraul. Res*. (IAHR), Leningrad, URSS, 3.8.1–3.8.11.

Ettema, R. & Mutel, C. F., 2004. Hans Albert Einstein: innovation and compromise in formulating sediment transport by rivers. *J. Hydr. Eng.*, **130**(6): 477–487.

Ferreira, R. M. F., 2005. River Morphodynamics and sediment transport. Conceptual model and solutions. PhD Thesis, Instituto Superior Técnico, Technical University of Lisbon.

Ferreira, R. M. L. & Cardoso, A. H., 1999. Numerical modelling of the upstream induced non-equilibrium sediment transport. Proc. of the 1st RCEM, Genova, Vol. I: 489–498.

Ferreira, R. M. L. & Leal, J. G. A. B., 1998. 1-D Mathematical Modelling of the Instantaneous-Dam-Break Flood Wave Over Mobile Bed: Application of TVD and Flux-Splitting Schemes. *CADAM Project Meeting*, Munich, Germany (edited in CD-rom).

Ferreira, R. M. L., Leal, J. G. A. B., Cardoso, A.H. & Almeida, A. B., 2003. Sediment transport by dam-break flows. A conceptual framework drawn from the theories for rapid granular flows. *3rd IMPACT Workshop*, Louvain-la-Neuve. Belgium, 5–7 November (Cd-Rom).

Ferreira, R. M. L., Amaral, S., Leal, J. G. A. B. & Spinewine, B., 2006. Discontinuities in geomorphic dam-break flows. *River Flow 2006*, Lisbon, Vol. 1: 1521–1530.

Ferreira, R.M.L., Leal, J. G. A. B., Cardoso, A. H., 2006. Conceptual model for the bedload layer of gravel bed streams based on laboratory observations. *River Flow 2006*, Lisbon, Vol. 1: 947–956.

Fraccarollo, L. & Capart, H. (2002) Riemann wave description of erosional dam-break flows. *J. Fluid Mech.* **461**: 183–228.

Fraccarollo, L. & Toro, E. F., 1995. Experimental and Numerical Assessment of the Shallow Water Model for Two-Dimensional Dam-Break Type Problems. J. of Hydraul Res., **33**(6): 843–864.

Fraccarollo, L., Capart, H. & Zech, Y., 2003. A Godunov method for the computation of erosional shallow-water transients. *Int. J. for Num. Methods in Fluids.*, **41**: 951–976.

Garcia, R. & Kahawita 1986 Numerical solution of the St. Venant equations with the MacCormack finite-difference scheme. *Int. J. for Numerical Meth. in Fluids.* **6**: 259–274.

García-Navarro, P. & Savirón, J., 1995. McCormack's method for the numerical simulation of one-dimensional discontinuous open channel flow. *J. of Hydraul. Res.*, **30**: 95–105.

Graf, W. H. 1998. *Fluvial Hydraulics: Flow and Transport Processes in Channels of Simple Geometry*. In collaboration with M.S. Altinakar, John Wiley and Sons, England, 681 pages. ISBN 0–471-97714–4.

Hirsch, C. 1990 *Numerical computation of internal and external flows. Vol. 2: Computational methods for inviscid and viscous flows*. 1994 Reprint, Willey-Interscience Publication.

Jenkins, J. T. & Askari, E. 1991 Boundary conditions for rapid flows: phase interfaces. *J. Fluid Mech.* **223**, 497–508.

Jenkins, J. T. & Richman, M. W. 1988 Plane simple shear of smooth inelastic circular disks: the anisotropy of the second moment in the dilute and dense limits. *J. Fluid Mech.* **192**, 313–328.

Jenkins, J. T. & Richman, M. W., 1985. Kinetic theory for plane flows of a dense gas of identical, rough, inelastic circular disks. *Phys. Fluids* **28**(12): 3485–3494.

Lapointe, M. F., Secretan, Y., Driscoll, S. N., Bergeron, N. & Leclerec, M., 1998. Response of the Ha! Ha! River to the flood of July 1996 in the Saguenay region of Quebec: large scale avulsion in a glaciated valley. *Water Resour. Res.* **34**(9): 2383–2392.

Lax, P. (1957) Hyperbolic systems of conservation laws II. Commun. Pure Appl. Maths **10**, pp. 537–566.

Leal, J. G. A. B., Ferreira, R. M. L., Cardoso, A. H., Almeida, A. B., 2003. Comparison between numerical and experimental results on dam-break waves over dry mobile beds. *Proc. 3rd IMPACT Workshop*, Louvain-la-Neuve, (edited in CD-Rom).

Leal, J. G. A. B., Ferreira, R. M. L., Franco, A.B., Cardoso, A.H., 2002. Experimental study on dam-break flood waves over movable bed channels. *Int. J. of Sediment Res.*, **17**(3): 186–196.

LeFloch, Ph. G., 1988. Entropy weak solutions to nonlinear hyperbolic systems in nonconservative form. *Comm. PDE*, **13**: 669–727.

LeVeque, R. J. 1990 *Numerical methods for conservation laws*. Lectures in Mathematics, ETH Zurich, Department of Mathematics.

Lun, K. K., Savage, S. B., Jeffrey, D. J. & Chepurny, N. 1984 Kinetic theories for granular flows: inelastic particles in Couoette flow and slightly inelastic particles in a general flow field. *J. Fluid Mech.* **140**: 223–256.

Lyn, D. A. (1987) Unsteady sediment transport modelling. *J. Hydraul. Eng.*, **113**(1): 1–15.

MacCormack, R. W., 1969 – *The effect of viscosity in hypervelocity impact cratering.* American Institute of Aeronautics and Astronautics, Paper 69–354.

Morris, P. H. & Williams, D. J. (1996) Relative celerities of mobile bed flows with finite solids concentrations. *J. Hydraul. Eng.*, **122**(6): pp. 311–315.

Nujic, M., 1995. Efficient implementation of non-oscilatory schemes for the computation of free-surface flows. *J. of Hydraul. Res.*, **33**: 101–101.

Preissmann, A. 1961 Propagation des intumescences dans les canaux et les rivières. *1er Congrès de l'Association française de calcul*, Grenoble, France, pp. 433–442.

Roe, P. L., 1981 – Approximate Riemann solvers, parameter vectors and difference schemes. *J. of Comp. Phys.*, **43**: 357–372.

Schlichting, H. & Gersten, K. 2000. *Boundary Layer Theory*. 8th Edition. Springer.

Sieben, J. 1999 A theoretical analysis of discontinuous flow with mobile bed. *J. Hydraul. Res.,* **37**(2): 199–212.

Sloff, C. J. 1993 Analysis of basic equations of sediment laden flows. *Communications on Hydraulic and Geotechnical Engineering, Rept. No 93–8*, Delft University of Technology, Delft, The Netherlands.

Spinewine, B. & Zech, Y., 2003. Experimental dam-break waves on movable bed: results and model comparisons from the IMPACT 'flat bed' benchmark. Proc. of 3rd IMPACT workshop, Louvain-la-Neuve (CD-ROM).

Spinewine, B., 2005. *Two-layer flow behaviour and the effects of granular dilatancy in dam-break induced sheet-flow*. Phd Thesis, Université catholique de Louvain, Belgium.

Sumer, B. M. & Deigaard, R. 1981 Particle motions near the bottom of in turbulent flow in an open channel initial. *J. Fluid Mech*., **109**: pp. 311–337.

Toro, E. F. 1997 *Riemann solvers and numerical methods for Fluid Dynamics. A practical introduction*. Springer-Verlag, Berlin.

Toro, E. F., Spruce, M. & Spears, W., 1994. Restoration of the contact wave in the HLL-Riemann solver. *Shock Waves*, **4**: 25–34.

Von Neumann, J. & Richtmyer, R. D. 1950 A method for the numerical calculations of hydrodynamical shocks. *J. Math. Phys.*, **21**(3): 232–237.

Yalin, M. S., 1977. *Mechanics of sediment transport*. 2nd Edition, Pergamon Press.

Zech, Y. Soares-Frazão, S., 2007. Dam-break flow experiments and real-case data. A database from the European IMPACT research, *J. of Hydraul Res.*, **45** Extra Issue, pp. 5–7.

Zech, Y., Soares-Frazão, S., Bousmar, D., Spinewine, B, Dal Cin, C. & Savary, C., 2006 Rivers and floods: a concern for people, a problem for engineers, a challenge for research. *River Flow 2006*, Lisbon, Vol 1: 13–27.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Numerical approach of sediment transport problems by high order finite volume methods

M.J. Castro
Dpto. Análisis Matemático, UMA, Málaga. Spain

E.D. Fernández-Nieto
Dpto. de Matemática Aplicada I, E.T.S. Arquitectura, Sevilla, Spain

A.M. Ferreiro
Universidad de A Coruña, Spain

C. Parés
Universidad de Málaga, Spain

ABSTRACT: This paper is concerned with the numerical approximation of bedload sediment transport due to water evolution and the tsumani generation by submarine avalanches. The bedload sediment transport models considered here have two components: The hydrodynamical component modeled by the shallow water equations and the morphodynamical component given by a continuity equation, which is defined in function of the solid transport discharge. We consider several deterministic models, such as Meyer-Peter & Müller, Van Rijn or Grass model, all of them formulated in an unified form. The model considered here for the tsunami generation by submarine avalances is a two-layer Savage-Huter type model (see (28)). Both kinds of models could be re-written as a non-conservative hyperbolic system. To discretize them, we consider finite volume methods and high order state reconstructions. Finally, several numerical tests are presented.

1 INTRODUCTION

This work is concerned with the numerical approximation of bedload sediment transport due to water evolution and the tsumani generation by submarine avalanches. The bedload sediment transport models considered here have two components: The hydrodynamical component modeled by the shallow water equations and the morphodynamical component given by a continuity equation, which is defined in function of the solid transport discharge. We consider several deterministic models, such as Meyer-Peter&Müller, Van Rijn or Grass model, all of them formulated in an unified form. The model considered here for the tsunami generation by submarine avalanches is a two-layer Savage-Huter type model (see (28)).

All models presented in this work are particular cases of a general class of hyperbolic systems of conservations laws with coupled and source terms that can be written in the form

$$\frac{\partial U}{\partial t} + \frac{\partial F_1}{\partial x_1}(U) + \frac{\partial F_2}{\partial x_2}(U) = B_1(U)\frac{\partial W}{\partial x_1} + B_2(U)\frac{\partial W}{\partial x_2} + S_1(U)\frac{\partial H}{\partial x_1} + S_2(U)\frac{\partial H}{\partial x_2}, \qquad (1)$$

where the unknown $U(\mathrm{x},t)$ is defined in $D \times (0,T)$, D being a domain of $\mathbb{R}^2$, and takes values on an open convex subset Ω of $\mathbb{R}^N$; F_i, $i = 1,2$ are two regular functions from Ω to $\mathbb{R}^N$; B_i, $i = 1,2$ are two regular matrix-valued function from Ω to $\mathcal{M}_N(\mathbb{R})$; S_i, $i = 1,2$ are two functions from Ω to $\mathbb{R}^N$; and finally $H(\mathrm{x})$ is a known function from $\mathbb{R}^2$ to $\mathbb{R}$.

Note that system (1) includes as particular cases: systems of conservation laws ($B_i = 0$, $S_i = 0$, $i = 1, 2$); systems of conservation laws with source term or balance laws ($B_i = 0$, $i = 1, 2$); and coupled systems of conservation laws.

The goal of this work is to present two-dimensional high-order finite volume numerical methods based on a generalized first order Roe scheme and on a high order reconstruction operator. To do this, first the system is written in the nonconservative form:

$$W_t + \mathcal{A}_1(W)W_{x_1} + \mathcal{A}_2(W)W_{x_2} = 0, \quad \mathbf{x} = (x_1, x_2) \in D, t \in (0, T), \tag{2}$$

being $\mathcal{A}_i$, $i = 1, 2$ two regular matrix-valued functions from $\Omega \times \mathbb{R}$ to $\mathcal{M}_{(N+1)\times(N+1)}(\mathbb{R})$. In effect, adding to (1) the equation

$$\frac{\partial H}{\partial t} = 0,$$

system (1) can be easily rewritten in this form (see (21), (13), (14), (16), (17)).

The nonconservative products involved in (2) do not make sense in general within the framework of distributions. Here, we follow the theory developed by Dal Maso, Le Floch and Murat in paper(10) to give a sense to these products as *Borel measures*. This theory is based in the choice of a family of paths.

First, we present a general writing of a Roe scheme for (2). To do this, the notion of Roe linearization introduced in (29) is used, which is also based on the choice of a family of paths. Once the Roe linearization chosen, a finite volume mesh of Ω is constructed and piecewise constant approximations of the solution are considered. These approximations are updated by considering, at any time level, a family of projected Riemann Problem in the normal direction to each edge of the mesh. These projected Riemann problems are associated to 1d hyperbolic nonconservative systems. These problems are then linearized by using the Roe linearization considered. The solutions of these 1d linear Riemann problems are then averaged in the cells to obtain the new piecewise constant approximation of the solution.

High order extensions are then obtained by extending to the 2d case the procedure developed in (4) for 1d problems: a reconstruction operator is considered, i.e. an operator that, given a family of constant values at the cells of the mesh, provides different values at the edges, in such a manner that if the values at the cell are the averages of a regular function, then the values at the edge are high order approximations of the value of the function.

This paper is organized as follows: in the second section we give the general expression of a Roe method for (2) and some general results concerning its well-balanced properties are presented.

In Section 3 we give the general form of a high order extension of Roe methods based on a reconstruction operator. The well-balance properties of such a scheme are also discussed.

The numerical schemes obtained in Sections 2 and 3 are particularized to systems (1). In the last section some numerical results are presented.

2 ROE METHODS FOR TWO-DIMENSIONAL NONCONSERVATIVE SYSTEMS

We consider the problem:

$$W_t + \mathcal{A}_1(W)W_{x_1} + \mathcal{A}_2(W)W_{x_2} = 0, \quad \mathbf{x} = (x_1, x_2) \in D \subset \mathbb{R}^2, t \in (0, T), \tag{3}$$

where $W(x, t)$ takes values on an open domain D of $\mathbb{R}^N$ and $\mathcal{A}_i$, $i = 1, 2$ are two smooth and locally bounded matrix-valued functions from Ω to $\mathcal{M}_{N\times N}(\mathbb{R})$.

Given an unitary vector $\eta = (\eta_1, \eta_2) \in \mathbb{R}^2$, we define the matrix

$$\mathcal{A}(W, \eta) = \mathcal{A}_1(W)\eta_1 + \mathcal{A}_2(W)\eta_2.$$

We assume that (3) is strictly hyperbolic, i.e. for all $W \in \Omega$ and $\forall \eta \in \mathbb{R}^2$, the matrix $\mathcal{A}(W, \eta)$ has N real and distinct eigenvalues

$$\lambda_1(W, \eta) < \cdots < \lambda_N(W, \eta).$$

As a result, $\mathcal{A}(W, \eta)$ is diagonalizable:

$$\mathcal{A}(W, \eta) = \mathcal{K}(W, \eta)\mathcal{D}(W, \eta)\mathcal{K}^{-1}(W, \eta),$$

being $\mathcal{D}(W, \eta)$ the diagonal matrix whose coefficients are the eigenvalues of $\mathcal{A}(W, \eta)$ and $\mathcal{K}(W, \eta)$ is a matrix whose columns are associated eigenvectors $R_j(W, \eta)$ to $\lambda_j(W, \eta), j = 1, \ldots, N$.

For discontinuous solutions W, the nonconservative products $\mathcal{A}_k(W)W_{x_k}$, $k = 1, 2$ do not make sense as a distribution. However, the theory developed by Dal Maso, LeFloch and Murat (see (10)) allows to give a rigorous definition of nonconservative products, associated to the choice of a family of paths Φ in Ω.

Suppose that a family of paths Φ in Ω has been chosen. Then, the nonconservative products in (3) can be interpreted as a Borel measure and then the notion of weak solution can be defined (see (10)). According to this definition, a piecewise regular function W is a weak solution if and only if the two following conditions are satisfied:

(i) W is a classical solution in the domains where it is $\mathcal{C}^1$.
(ii) At every point of a discontinuity W satisfies the jump condition

$$\int_0^1 (\sigma\mathcal{I} - \mathcal{A}(\Phi(s; W^-, W^+, \eta), \eta))\frac{\partial \Phi}{\partial s}(s; W^-, W^+, \eta)ds = 0, \tag{4}$$

where $\mathcal{I}$ is the identity matrix; σ, the speed of propagation of the discontinuity; η a unit vector normal to the discontinuity at the considered point; and W^-, W^+, the lateral limits of the solution at the discontinuity, and $\Phi(s; W^-, W^+, \eta)$ the path conecting W^- and W^+.

Together with the definition of weak solutions, a notion of entropy has to be chosen, as the usual Lax's concept or one related to an entropy pair.

The choice of the family of paths is important because it determines the speed of propagation of shocks. The simplest choice is given by the family of segments:

$$\Phi(s; W_L, W_R, \eta) = W_L + s(W_R - W_L), \tag{5}$$

that corresponds to the definition of nonconservative products proposed by Volpert (see (32)). In practical applications, it has to be based on the physical background of the problem. In paper (21) a clear motivation for the selection of the family of paths is provided when a physical regularization by diffusion, dispersion, etc. is available.

In order to construct a first order numerical scheme for (3), the extension of the the concept of Roe linearization introduced in (29) for 1d problems and which is also based on the use of a family of paths, is used (see (5)):

Definition 1 *Given a family of paths Ψ, a function $\mathcal{A}_\Psi : \Omega \times \Omega \times \mathcal{S}^1 \to \mathcal{M}_N(\mathbb{R})$ is called a* Roe linearization *of (3), if it verifies the following properties:*

1. *For each $W_L, W_R \in \Omega$ and $\eta \in \mathcal{S}^1$, $\mathcal{A}_\Psi(W_L, W_R, \eta)$ has N distinct real eigenvalues:*

$$\lambda_1(W_L W_R, \eta) < \lambda_2(W_L, W_R, \eta) < \cdots < \lambda_N(W_L, W_R, \eta).$$

2. $\mathcal{A}_\Psi(W, W, \eta) = \mathcal{A}(W, \eta)$, *for every $W \in \Omega$, $\eta \in \mathcal{S}^1$.*
3. *For any $W_L, W_R \in \Omega$, $\eta \in \mathcal{S}^1$:*

$$\mathcal{A}_\Psi(W_L, W_R, \eta)(W_R - W_L) = \int_0^1 \mathcal{A}(\Psi(s; W_L, W_R, \eta), \eta)\frac{\partial \Psi}{\partial s}(s; W_L, W_R, \eta)\, ds. \tag{6}$$

Notice that if $\mathcal{A}_k(W)$, $k=1,2$ are the Jacobian matrices of two smooth flux functions $F_k(W)$, $k=1,2$, (6) is independent of the family of paths and it reduces to the usual Roe property:

$$\mathcal{A}_\Psi(W_L, W_R, \eta)\cdot(W_R - W_L) = F_\eta(W_R) - F_\eta(W_L), \tag{7}$$

for any $\eta=(\eta_1,\eta_2)\in\mathcal{S}^1$, where

$$F_\eta(W) = \eta_1 F_1(W) + \eta_2 F_2(W), \tag{8}$$

represents the flux along the η direction.

Once a Roe linearization $\mathcal{A}_\Psi$ has been chosen, in order to discretize the system, first the domain D is decomposed into subsets with an easy geometry, called cells or finite volumes, $V_i\subset\mathbb{R}^2$. We assume here that the cells are closed polygons whose intersections are either empty or a complete edge. We will denote by $\mathcal{T}_h$ the mesh, i.e. the set of cells, and by NV the number of cells.

The following notation is considered: given a finite volume V_i, $N_i\in\mathbb{R}^2$ represents its center; $\mathcal{N}_i$ is the set of indexes j such that V_j is a neighbor of V_i; E_{ij} is the common edge to two neighbor cells V_i and V_j, and $|E_{ij}|$ represents its length; $\eta_{ij}=(\eta_{ij,1},\eta_{ij,2})$ is the normal unitary vector of edge E_{ij} pointing towards the cell V_j (see Figure 1). Each cell can be decomposed in triangular subcells $\{V_{ij}\}_{j\in\mathcal{N}_i}$: V_{ij} is the triangle defined by the center of the cell N_i and the edge E_{ij} (see Figure 1). $|V_i|$ and $|V_{ij}|$ represent, respectively, the areas of V_i and V_{ij}. Δx is the maximum of the diameters of the cells. Finally, W_i^n will represents the constant approximation of the averaged exact solution in cell V_i and time t^n provided by the numerical scheme:

$$W_i^n \cong \frac{1}{|V_i|}\int_{V_i} W_i(\mathbf{x},t^n)d\mathbf{x}.$$

Let us suppose that the approximations in the cells at time t^n, $\{W_i^n\}$, are already known. To progress in time, a projected Riemann Problem is considered at every edge E_{ij}: the projected Riemann Problem is a 1d non-conservative hyperbolic system as those studied in (25). Following this work, this one-dimensional problem is discretized by means of a generalized Roe scheme. W_i^{n+1} is then calculated by averaging the approximations obtained at every edge. The resulting scheme is as follows (see (6)):

$$W_i^{n+1} = W_i^n - \frac{1}{|V_i|}\sum_{j\in\mathcal{N}_i}|E_{ij}|\mathcal{A}_{ij}^-(W_j^n - W_i^n), \tag{9}$$

where $\mathcal{A}_{ij}$ is the Roe matrix associated to the states W_i^n, W_j^n, and the direction η_{ij},

$$\mathcal{A}_{ij} = \mathcal{A}_\Psi(W_i^n, W_j^n, \eta_{ij}).$$

and

$$\mathcal{A}_{ij}^- = \mathcal{K}_{ij}\mathcal{D}_{ij}^-\mathcal{K}_{ij}^{-1},$$

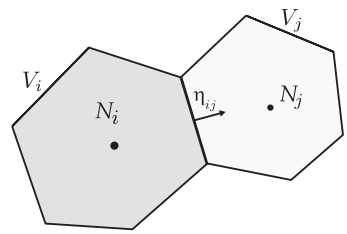


Figure 1. Finite volumes.

being $\mathcal{D}_{ij}^{-}$ the diagonal matrix whose coefficients are the negative parts of the eigenvalues of $\mathcal{A}_{ij}$ and $\mathcal{K}_{ij}$ is a $N \times N$ matrix whose columns are the associated eigenvectors. This is the general expression of a Roe scheme for problem (3).

Remark 1 *A C F L-like requirement must be imposed to guarantee the linear stability of the numerical scheme:*

$$\max\left\{\frac{|\lambda_{ij,k}|}{d_{ij}} : i = 1,\dots,NV, j \in \mathcal{N}_i, \quad k = 1,\dots,N\right\} \cdot \Delta t \leq \gamma, \tag{10}$$

with $0 < \gamma \leq 1$, being $\lambda_{ij,k}$, $k = 1,\dots,N$ the eigenvalues of the matrix $\mathcal{A}_{ij}$, and d_{ij} is the distance between the node N_i and the edge E_{ij}.

Remark 2 *As in the case of systems of conservation laws, when sonic rarefaction waves appear it is necessary to modify the approximate Riemann problem solver in order to obtain entropy-satisfying solutions. The Harten-Hyman Entropy Fix technique (see (18)), for instance, can be easily adapted to this problem.*

Remark 3 *The best choice of the family of paths Ψ appearing in the definition of Roe linearization is the family Φ selected for the definition of weak solutions. In effect, Roe methods based on the family of paths Φ can handle correctly with discontinuities in the following sense: let us suppose that the approximations at two neighbor cells, W_i^n and W_j^n, can be linked by an entropic discontinuity propagating at speed σ with normal direction η_{ij}; then, from (4) and (6) we deduce:*

$$\mathcal{A}_{ij}(W_j^n - W_i^n) = \sigma(W_j^n - W_i^n),$$

i.e., σ is an eigenvalue of the intermediate matrix and $W_j^n - W_i^n$ is an associated eigenvector. Nevertheless, the construction of these schemes with $\Psi = \Phi$ can be difficult or very costly in practice. In this case, a simpler family of paths Ψ has to be chosen, as the family of segments: In (25) it was remarked that, in this case, the convergence of the numerical scheme can fail when the weak solution to approach involves discontinuities connecting states W^- and W^+ when the paths of the families Φ and Ψ linking them are different.

2.1 *Well-balancing*

The well-balance properties of the one-dimensional Roe scheme here considered have been studied in (25). In general, these properties are not inherited by two dimensional schemes. Nevertheless some general results can be proved.

Definition 2 *Given an edge E_{ij}, we will denote by Γ_{ij} the set of integral curves of linearly degenerated fields of $\mathcal{A}(W, \eta_{ij})$ such that the corresponding eigenvalue vanishes along the curve.*

Theorem 1 *Let W be a regular stationary solution of (3) satisfying the following property: given two neighbor cells V_i, V_j the path*

$$s \in [0,1] \to \Psi(s, W(N_i), W(N_j), \eta_{ij})$$

is a parametrization of an arc γ of Γ_{ij}. Then the numerical scheme computes exactly the solution W (see (5) for the proof).

Corollary 1 *If Ψ is the family of segments (5), then the numerical scheme solves exactly any regular stationary solution W such that, for any pair of neighbor cells, V_i, V_j the segment linking $W(N_i)$ and $W(N_j)$ belongs to a curve γ of Γ_{ij}.*

Theorem 2 *Let us assume that $\mathcal{A}_1, \mathcal{A}_2$ are of class $\mathcal{C}^1$ with bounded derivatives. Let W be a regular stationary solution of the problem (3) with the following property: given two neighbor*

cells V_i, V_j the states $W(N_i)$, $W(N_j)$ belong to a same curve γ of Γ_{ij}, and it is possible to find a parametrization:

$$U : [0,1] \longrightarrow \Omega,$$

of class $\mathcal{C}^{k+1}$ of the arch of γ linking them, such that:

$$\int_0^1 \left| U'(s) - \frac{\partial \Psi}{\partial s}(s, W(N_i), W(N_j), \eta_{ij}) \right| ds = O(d(N_i, N_j)^{k+1}). \tag{11}$$

Then, the scheme approximates the solutions W with order k (see (5) for the proof).

Notice that the hypothesis concerning the stationary solutions in the previous theorems are much more restrictive than in the 1d case: while in 1d problems the values taken by a stationary solution at the centers of two neighbor cells always belong to the same integral curve of a linearly degenerate field whose corresponding eigenvalue vanishes along the curve (see (25)), this is not the case for general two dimensional stationary solutions. Nevertheless, this happens when the stationary solution is esentially 1d and the mesh is rectangular and properly oriented. For this kind of solutions and meshes the 2d schemes have the same well balanced properties that the one dimensional schemes from which they have been derived.

Nevertheless, we will see that Corollary 1 is enough to ensure the *C Property* of the schemes applied to shallow water systems of one or two layers, i.e. the stationary solutions corresponding to water at rest situations are preserved.

3 HIGH ORDER SCHEMES BASED ON RECONSTRUCTION OF STATES

In this section, we present a high order extension by state reconstructions of the previously proposed excheme.

Let us consider first the case of systems of conservation laws

$$\frac{\partial W}{\partial t} + \frac{\partial}{\partial x_1}(F_1(W)) + \frac{\partial}{\partial x_2}(F_2(W)) = 0. \tag{12}$$

These systems can be considered as a particular case of (3) in which the matrices $\mathcal{A}_i(W)$, $i = 1,2$ are the Jacobians:

$$\mathcal{A}_i(W) = \frac{\partial F_i}{\partial W}, \quad i = 1, 2.$$

High order methods based on the reconstruction of states can be built for (12) using the following procedure: given a first order conservative scheme with numerical flux function $G(U, V; \eta)$, a reconstruction operator of order p is considered, that is, an operator that associates to a given family $\{W_i\}_{i=1}^{NV}$ of values at the cells two families of functions defined at the edges:

$$\gamma \in E_{ij} \rightarrow W_{ij}^{\pm}(\gamma),$$

in such a way that, whenever

$$W_i = \frac{1}{|V_i|} \int_{V_i} W(\mathbf{x})\mathbf{dx} \tag{13}$$

for some smooth function W, we have that

$$W_{ij}^{\pm}(\gamma) = W(\gamma) + O(\Delta x^p).$$

Once the first order method and the reconstruction operator have been chosen, the method of lines can be used to develop high order methods for (12): the idea is to discretize only in space, leaving the problem continuous in time. This procedure leads to a system of ordinary differential equations which is solved using a standard numerical method.

Let $\overline{W}_i(t)$ denotes the cell average of a regular solution W of (12) over the cell V_i at time t:

$$\overline{W}_i(t) = \frac{1}{|V_i|}\int_{V_i} W(x,t)dx.$$

The following equation can be easily obtained for the cell averages:

$$\overline{W}_i^{'}(t) = -\frac{1}{|V_i|}\left(\sum_{j\in\mathcal{N}_i}\int_{E_{ij}} F(W(\gamma,t))\cdot\eta_{ij}d\gamma\right). \tag{14}$$

The first order method and the reconstructions are now used to approach the values of the fluxes at the edges:

$$W_i^{'}(t) = -\frac{1}{|V_i|}\left(\sum_{j\in\mathcal{N}_i}\int_{E_{ij}} G(W_{ij}^{-}(\gamma,t), W_{ij}^{+}(\gamma,t),\eta_{ij})d\gamma\right), \tag{15}$$

being $W_i(t)$ the approximation to $\overline{W}_i(t)$ provided by the scheme and $W_{ij}^{\pm}(\gamma,t)$ the reconstruction at $\gamma\in E_{ij}$ corresponding to the family $\{W_i(t)\}_{i=1}^{NV}$. It can be shown that (15) is an approximation of order p of (14).

In practice, the integral terms in (15) are approached by means of a numerical quadrature of order $\overline{r}\geq p$ at least:

$$\int_a^b f(s)ds = (b-a)\left(\sum_{l=1}^{n(\overline{r})}\omega_l f(x_l)\right) + O(\Delta^{\overline{r}}), \tag{16}$$

where $n(\overline{r})$ denotes the number of points, ω_l are the weights, and $x_l = a + s_l(b-a)$ with $s_l\in[0,1]$, represent the quadrature points. The expression of the corresponding semi-discrete numerical scheme is then as follows:

$$W_i'(t) = -\frac{1}{|V_i|}\sum_{j\in\mathcal{N}_i}|E_{ij}|\left(\sum_{l=1}^{n(\overline{r})}\omega_l G(W_{ij,l}^{-}(t), W_{ij,l}^{+}(t),\eta_{ij})\right). \tag{17}$$

where

$$W_{ij,l}^{\pm}(t) = W_{ij}^{\pm}(a_{ij} + s_l(b_{ij}-a_{ij}),t), \tag{18}$$

a_{ij} and b_{ij} being the extremes of the edge E_{ij}.

We assume here that the first order scheme is a Roe method, i.e.:

$$G(W_0,W_1,\eta) = \frac{F_\eta(W_1)+F_\eta(W_0)}{2} - \frac{1}{2}|\mathcal{A}(W_0,W_1,\eta)|(W_1-W_0), \tag{19}$$

where F_η is given by (8) and $\mathcal{A}(W_0,W_1,\eta)$ is the Roe matrix associated to the states W_i, W_j, and to the unit vector η, i.e. an intermediate matrix that verifies the Roe property:

$$F_\eta(W_1) - F_\eta(W_0) = \mathcal{A}(W_0,W_1,\eta)\cdot(W_1-W_0). \tag{20}$$

Let us now generalize the semi-discretes methods (15) or (17) for a nonconservative system (3). We will assume that the reconstructions are calculated as follows: given the family $\{W_i\}_{i=1}^{NV}$ of values at the cells, first an approximation function is constructed at every cell V_i, based on the values of W_i at some of the cells close to V_i (the *stencil*):

$$P_i(\mathbf{x}) = P_i(\mathbf{x};\{W_j\}_{j\in\mathcal{B}_i}),$$

for some set of indexes $\mathcal{B}_i$. If, for instance, the reconstruction only depends on the neighbor cells of V_i, then $\mathcal{B}_i = \mathcal{N}_i\cup\{i\}$. These approximations functions are calculated usually by means

of an interpolation or approximation procedure. Once these functions have been constructed, the reconstruction at $\gamma \in E_{ij}$ are defined as follows:

$$W_{ij}^{-}(\gamma) = \lim_{\mathbf{x}\to\gamma} P_i(\mathbf{x}); W_{ij}^{+}(\gamma) = \lim_{\mathbf{x}\to\gamma} P_j(\mathbf{x}); . \tag{21}$$

Clearly, for any $\gamma \in E_{ij}$ the following equalities are satisfied:

$$W_{ij}^{-}(\gamma) = W_{ji}^{+}(\gamma); W_{ij}^{+}(\gamma) = W_{ji}^{-}(\gamma).$$

We suppose that the state reconstruction operator satisfies the following properties:
H1$_{P_i}$: It is conservative, i.e. the following equality holds for any cell V_i:

$$W_i = \frac{1}{|V_i|}\int_{V_i} P_i(\mathbf{x})d\mathbf{x}. \tag{22}$$

H2$_{P_i}$: It is of order p.
H3$_{P_i}$: It is of order q in the interior of the cells, i.e. if the operator is applied to a sequence $\{W_i\}$ satisfying (13) for some smooth function $W(\mathbf{x})$, then:

$$P_i(\mathbf{x}) = W(\mathbf{x}) + O(\Delta x^q), \forall \mathbf{x} \in int(V_i). \tag{23}$$

H4$_{P_i}$: Under the assumption of the previous property, the gradient of P_i provides an approximation of order m of the gradient of W:

$$\nabla P_i(\mathbf{x}) = \nabla W(\mathbf{x}) + O(\Delta x^m), \forall \mathbf{x} \in int(V_i). \tag{24}$$

Let us denote by P_i^t the approximation functions corresponding to the approximations of the cell averages $W_i(t)$, i.e.

$$P_i^t(\mathbf{x}) = P_i(\mathbf{x}; \{W_j(t)\}_{j\in\mathcal{B}_i}).$$

$W_{ij}^{-}(\gamma,t)$ (resp. $W_{ij}^{+}(\gamma,t)$) is defined by

$$W_{ij}^{-}(\gamma,t) = \lim_{\mathbf{x}\to\gamma} P_i^t(\mathbf{x}); W_{ij}^{+}(\gamma,t) = \lim_{\mathbf{x}\to\gamma} P_j^t(\mathbf{x}); . \tag{25}$$

Using the Roe property (20), together with the definition of the reconstructions and the divergence theorem, we can rewrite (15) as follows:

$$\begin{aligned}
W_i'(t) =& -\frac{1}{|V_i|}\sum_{j\in\mathcal{N}_i}\int_{E_{ij}} (\mathcal{A}_{ij}^{-}(\gamma,t)(W_{ij}^{+}(\gamma,t) - W_{ij}^{-}(\gamma,t)))d\gamma \\
& -\frac{1}{|V_i|}\sum_{j\in\mathcal{N}_i}\int_{E_{ij}} (F_{\eta_{ij}}(W_{ij}^{-}(\gamma,t)))d\gamma \\
=& -\frac{1}{|V_i|}\sum_{j\in\mathcal{N}_i}\int_{E_{ij}} (\mathcal{A}_{ij}^{-}(\gamma,t)(W_{ij}^{+}(\gamma,t) - W_{ij}^{-}(\gamma,t)))d\gamma \\
& -\frac{1}{|V_i|}\sum_{j\in\mathcal{N}_i}\int_{V_i} \nabla\cdot(F\circ P_i^t)(\mathbf{x}))d\mathbf{x} \\
=& -\frac{1}{|V_i|}\sum_{j\in\mathcal{N}_i}\int_{E_{ij}} (\mathcal{A}_{ij}^{-}(\gamma,t)(W_{ij}^{+}(\gamma,t) - W_{ij}^{-}(\gamma,t)))d\gamma \\
& -\frac{1}{|V_i|}\sum_{j\in\mathcal{N}_i}\int_{V_i} \left(\mathcal{A}_1(P_i^t(\mathbf{x}))\frac{\partial P_i^t}{\partial x_1}(\mathbf{x}) + \mathcal{A}_2(P_i^t(\mathbf{x}))\frac{\partial P_i^t}{\partial x_2}(\mathbf{x})\right)dx;
\end{aligned} \tag{26}$$

where the following notation has been used

$$\mathcal{A}_{ij}(\gamma,t)=\mathcal{A}(W_{ij}^{-}(\gamma,t),W_{ij}^{-}(\gamma,t),\eta_{ij}).$$

Notice now that (15) can be easily generalized to obtain a numerical scheme for solving (3):

$$\begin{aligned}W_i'(t) = -\frac{1}{|V_i|}\Bigg[&\sum_{j\in\mathcal{N}_i}\int_{E_{ij}}(\mathcal{A}_{ij}^{-}(\gamma,t)\cdot(W_{ij}^{+}(\gamma,t)-W_{ij}^{-}(\gamma,t)))d\gamma\\&+\int_{V_i}\left(\mathcal{A}_1(P_i^t(x))\frac{\partial P_i^t}{\partial x_1}(x)+\mathcal{A}_2(P_i^t(x))\frac{\partial P_i^t}{\partial x_2}(x)\right)dx\Bigg];\end{aligned}\tag{27}$$

where now

$$\mathcal{A}_{ij}(\gamma,t)=\mathcal{A}_\Psi(W_{ij}^{-}(\gamma,t),W_{ij}^{+}(\gamma,t),\eta_{ij}),$$

being $\mathcal{A}_\Psi$ the chosen Roe linearization.

3.1 *Order of accuracy*

The cell averages of a smooth solution of (3), $\overline{W}_i(t)$, satisfy:

$$\overline{W}_i'(t)=-\frac{1}{|V_i|}\int_{V_i}(\mathcal{A}_1(W(\mathbf{x}))W_{x_1}(\mathbf{x})+\mathcal{A}_2(W(\mathbf{x}))W_{x_2}(\mathbf{x}))d\mathbf{x}.\tag{28}$$

Thus, (27) is expected to be a good approximation of (28). This fact is stated in the following result, whose proof is similar to the corresponding result for 1d problems (see (4)):

Theorem 3 *Let us assume that $\mathcal{A}_1$ and $\mathcal{A}_2$ are of class $\mathcal{C}^2$ with bounded derivatives and $\mathcal{A}_\Psi$ is bounded for all i, j. Let us also suppose that the reconstruction operator satisfies the hypothesis $\mathbf{H1}_{P_i}-\mathbf{H4}_{P_i}$. Then (27) is an approximation of order at least $\alpha=\min(p,q+1,r+1)$ to the system (28) in the following sense:*

$$\begin{aligned}&\sum_{j\in\mathcal{N}_i}\int_{E_{ij}}(\mathcal{A}_{ij}^{-}(\gamma,t)(W_{ij}^{+}(\gamma,t)-W_{ij}^{-}(\gamma,t)))d\gamma\\&+\int_{V_i}\left(\mathcal{A}_1(P_i^t(\mathbf{x}))\frac{\partial P_i^t}{\partial x_1}(\mathbf{x})+\mathcal{A}_2(P_i^t(\mathbf{x}))\frac{\partial P_i^t}{\partial x_2}(\mathbf{x})\right)d\mathbf{x}\\&=\int_{V_i}(\mathcal{A}_1(W(\mathbf{x},t))W_{x_1}(\mathbf{x},t)+\mathcal{A}_2(W(\mathbf{x},t))W_{x_2}(\mathbf{x},t))d\mathbf{x}+O(\Delta^{\alpha+1});\end{aligned}\tag{29}$$

for every solution W smooth enough, being $W_{ij}^{\pm}(\gamma,t)$ the associated reconstructions and P_i^t the approximation functions corresponding to the family

$$\overline{W}_i(t)=\frac{1}{|V_i|}\int_{V_i}W(\mathbf{x},t)d\mathbf{x}.$$

3.2 *Approximation of the integral terms*

In practice, the integral terms appearing in the expression of the numerical scheme are numerically approached. In this case, together with a 1d formula (16) for the integrals on the edges, it can be also necessary to choose a quadrature formula of order $\bar{s}$ for the integrals in the cells:

$$\int_{V_i}f(\mathbf{x})d\mathbf{x}=|V_i|\sum_{l=1}^{n(\bar{s})}\alpha_l^i f(\mathbf{x}_l^i)+O(|V_i|^{\bar{s}}).\tag{30}$$

In order to preserve the order of the numerical scheme, it is necessary to habe $\overline{r} \geq \alpha$ and $\overline{s} \geq \alpha$. The numerical scheme writes as follows:

$$\begin{aligned} W_i'(t) = -\frac{1}{|V_i|}\Bigg[& \sum_{j\in\mathcal{N}_i} |E_{ij}| \sum_{l=1}^{n(\overline{r})} w_l \mathcal{A}^-_{ij,l}(t)\cdot(W^+_{ij,l}(t) - W^-_{ij,l}(t)) \\ & + |V_i| \sum_{l=1}^{n(\overline{s})} \alpha_l^i \left(\mathcal{A}_1(P_i^t(\mathbf{x}_l^i)) \frac{\partial P_i^t}{\partial x_1}(\mathbf{x}_l^i) + \mathcal{A}_2(P_i^t(\mathbf{x}_l^i)) \frac{\partial P_i^t}{\partial x_2}(\mathbf{x}_l^i) \right) . \Bigg], \end{aligned} \tag{31}$$

where

$$\begin{aligned} W^{\pm}_{ij,l}(t) &= W^{\pm}_{ij}(a_{ij} + s_l(b_{ij} - a_{ij}), t), \\ \mathcal{A}_{ij,l}(t) &= \mathcal{A}_\Psi(W^-_{ij,l}(t), W^+_{ij,l}(t), \eta_{ij}). \end{aligned}$$

3.3 *Well-balance properties*

In this section we study the well-balanced properties of the schemes (27) or (31).

Definition 3 *We consider a semi-discrete method to approximate (3):*

$$\begin{cases} W_i'(t) = \dfrac{1}{|V_i|}\mathcal{H}(W_j(t), j \in \mathcal{B}_i), \\ \mathbf{W}(0) = \mathbf{W}_0, \end{cases} \tag{32}$$

where $\mathbf{W}(t) = \{W_i(t)\}_{i=1}^{NV}$ represents the vector of the approximations to the mean value of the exact solution, $\mathbf{W}_0 = \{W_i^0\}$ is the vector of the initial conditions, and $\mathcal{B}_i$ are the stencils. Given a smooth stationary solution W of the system, the numerical scheme is said to be exactly well-balanced for W if the vector or its cell averages is a critical point of (32), i.e.

$$\mathcal{H}(W_j, j \in \mathcal{B}_i) = 0; \tag{33}$$

ant it is said to be well-balanced with order p if,

$$\mathcal{H}(W_j, j \in \mathcal{B}_i) = \mathcal{O}(\Delta^{p+1}). \tag{34}$$

Let us introduce the concept of well-balance reconstruction operator:

Definition 4 *Given a smooth stationary solution of (3), a reconstruction operator is said to be well-balanced for $W(\mathbf{x})$ if the approximation functions P_i corresponding to the family of its cell averages coincide with W at every cell V_i, i.e.*

$$P_i(\mathbf{x}) = W(x), \quad \forall \mathbf{x} \in V_i, \quad i = 1, \ldots, NV.$$

The two following results can be easily proved:

Theorem 4 *Let W be a smooth stationary solution of (3). Let us suppose that both the first order Roe method and the reconstruction operator chosen are exactly well-balanced for W. Then the numerical schemes (27) and (31) are also exactly well-balanced for W.*

Theorem 5 *Under the hypotheses of Theorem 3, the schemes (27) and (31) are well-balanced with order at least $\alpha = \min(p, q+1, r+1)$.*

4 SYSTEMS OF CONSERVATION LAWS WITH NON CONSERVATIVE PRODUCTS AND SOURCE TERMS

In order to particularize the previous numerical schemes to problem (1), the following notation is used:

$$\mathcal{A}_k(W) = \mathcal{J}_k(W) - B_k(W), \quad k = 1, 2,$$

being $\mathcal{J}_k$, $k = 1, 2$ the Jacobian matrix of the flux functions F_k, $k = 1, 2$:

$$\mathcal{J}_k(W) = \frac{\partial F_k}{\partial W}, \quad k = 1, 2.$$

$$\begin{aligned}
\mathcal{A}(W, \eta) &= \eta_1 \mathcal{A}_1(W) + \eta_2 \mathcal{A}_2(W), \\
\mathcal{J}(W, \eta) &= \eta_1 \mathcal{J}_1(W) + \eta_2 \mathcal{J}_2(W), \\
B(W, \eta) &= \eta_1 B_1(W) + \eta_2 B_2(W), \\
S(W, \eta) &= \eta_1 S_1(W) + \eta_2 S_2(W).
\end{aligned}$$

In order to construct Roe matrices for the system (1), first of all a family of paths $\widetilde{\Psi}$ has to be chosen for an edge with unit normal vector equal to η. The following notation will be used to describe, given a vector η, the path linking two states $\widetilde{W}_0 = [W_0, H_0]^T$, $\widetilde{W}_1 = [W_1, H_1]^T$:

$$\widetilde{\Psi}(s; \widetilde{W}_0, \widetilde{W}_1, \eta) = \begin{bmatrix} \Psi(s; \widetilde{W}_0, \widetilde{W}_1, \eta) \\ \Psi_{N+1}(s; \widetilde{W}_0, \widetilde{W}_1, \eta) \end{bmatrix} = \begin{bmatrix} \Psi_1(s; \widetilde{W}_0, \widetilde{W}_1, \eta) \\ \Psi_2(s; \widetilde{W}_0, \widetilde{W}_1, \eta) \\ \vdots \\ \Psi_{N+1}(s; \widetilde{W}_0, \widetilde{W}_1, \eta) \end{bmatrix}.$$

Let us suppose that, given any unit vector η and two states $\widetilde{W}_j = [W_j, H_j]^T$, $j = 0, 1$ it is possible to calculate:

- A matrix $\mathcal{J}(W_0, W_1, \eta)$ such that:

$$\mathcal{J}(W_0, W_1, \eta) \cdot (W_1 - W_0) = F_\eta(W_1) - F_\eta(W_0),$$

  i.e. a Roe matrix for the flux function F_η.
- A matrix $B_{\widetilde{\Psi}}(W_0, W_1, \eta)$ satisfying:

$$B_{\widetilde{\Psi}}(W_0, W_1, \eta) \cdot (W_1 - W_0) = \int_0^1 B(\Psi(s; \widetilde{W}_0, \widetilde{W}_1, \eta), \eta) \frac{\partial \Psi}{\partial s}(s; \widetilde{W}_0, \widetilde{W}_1, \eta) ds; \tag{35}$$

- A vector $S_{\widetilde{\Psi}}(W_0, W_1, \eta)$ satisfying:

$$S_{\widetilde{\Psi}}(W_0, W_1, \eta) \cdot (H_1 - H_0) = \int_0^1 S(\Psi(s; \widetilde{W}_0, \widetilde{W}_1, \eta), \eta) \cdot \frac{\partial \Psi_{N+1}}{\partial s}(s; \widetilde{W}_0, \widetilde{W}_1, \eta) ds, \tag{36}$$

Then, if we define:

$$\begin{aligned}\mathcal{J}_{ij} &= \mathcal{J}(W_i^n, W_j^n, \eta_{ij}),\\ B_{ij} &= B_{\widetilde{\Psi}}(W_i^n, W_j^n, \eta_{ij}),\\ S_{ij} &= S_{\widetilde{\Psi}}(W_i^n, W_j^n, \eta_{ij}),\\ \mathcal{A}_{ij} &= \mathcal{J}_{ij} - B_{ij}.\end{aligned}$$

The corresponding generalized Roe scheme (9) could be rewritten as follows:

$$W_i^{n+1} = W_i^n - \frac{\Delta t}{|V_i|}\sum_{j\in\mathcal{N}_i}|E_{ij}|(G_{ij} - B_{ij}\cdot(W_j^n - W_i^n) - \mathcal{P}_{ij}^- S_{ij}(H_j - H_i)) \tag{37}$$

where

$$G_{ij} = \frac{1}{2}(F_{\eta_{ij}}(W_i^n) + F_{\eta_{ij}}(W_j^n)) - \frac{1}{2}|\mathcal{A}_{ij}|\cdot(W_j^n - W_i^n)$$

is the usual Roe flux and

$$\mathcal{P}_{ij}^- = \frac{1}{2}(\mathcal{I} - |\mathcal{A}_{ij}|\mathcal{A}_{ij}^{-1}).$$

This latter matrix can be also be written in the form

$$\mathcal{P}_{ij}^- = \frac{1}{2}\mathcal{K}_{ij}(\mathcal{I} - \mathrm{sgn}(\mathcal{D}_{ij}))\mathcal{K}_{ij}^{-1},$$

where $\mathrm{sgn}(\mathcal{D}_{ij})$ is the diagonal matrix whose coefficients are the signs of the eigenvalues $\lambda_{ij,1},\ldots,\lambda_{ij,N}$ of $\mathcal{A}_{ij}$ and $\mathcal{K}_{ij}$ is a $N\times N$ matrix whose columns are associated eigenvectors.

Using the equality $\mathcal{A}_{ij}^- = \mathcal{P}_{ij}^-\mathcal{A}_{ij}$ the numerical scheme can also be written in this way:

$$W_i^{n+1} = W_i^n - \frac{\Delta t}{|V_i|}\sum_{j\in\mathcal{N}_i}|E_{ij}|F_{ij}^-, \tag{38}$$

where

$$F_{ij}^- = \mathcal{P}_{ij}^-(\mathcal{A}_{ij}(W_j^n - W_i^n) - S_{ij}(H_j - H_i)). \tag{39}$$

In a similar way, the semi-discrete high order extension of the Roe scheme (37) based on a reconstruction operator (27) can be expressed as follows:

$$\begin{aligned}W_i' = &-\frac{1}{|V_i|}\sum_{j\in\mathcal{N}_i}\int_{E_{ij}}(G_{ij}(\gamma,t) - B_{ij}(\gamma,t)\cdot(W_{ij}^+(\gamma,t) - W_{ij}^-(\gamma,t))\\ &-\mathcal{P}_{ij}^-(\gamma,t)S_{ij}(\gamma,t)(H_{ij}^+(\gamma) - H_{ij}^-(\gamma)))d\gamma\\ &+\frac{1}{|V_i|}\int_{V_i}\left(\sum_{k=1}^{2}B_k(P_i^t(\mathbf{x}))\frac{\partial P_i^t}{\partial x_k}(\mathbf{x})\right.\\ &+\left.\sum_{k=1}^{2}S_k(P_i^t(\mathbf{x}))\frac{\partial P_{N+1,i}^t}{\partial x_k}(\mathbf{x})\right)d\mathbf{x},\end{aligned} \tag{40}$$

where the following notation has been used:

$$\mathbf{x} \in V_i \to \widetilde{P}_i^t(\mathbf{x}) = \begin{bmatrix} P_i^t(\mathbf{x}) \\ P_{i,N+1}^t(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} P_{i,1}^t(\mathbf{x}) \\ P_{i,2}^t(\mathbf{x}) \\ \vdots \\ P_{i,N}^t(\mathbf{x}) \\ P_{i,N+1}^t(\mathbf{x}) \end{bmatrix},$$

represents the approximation function at the cell V_i at time t;

$$\gamma \in E_{ij} \to \widetilde{W}_{ij}^{\pm}(\gamma, t) = \begin{bmatrix} W_{ij}^{\pm}(\gamma, t) \\ H_{ij}^{\pm}(\gamma) \end{bmatrix},$$

the reconstructions at the edge Γ_{ij} at time t;

$$\begin{aligned}
\mathcal{J}_{ij}(\gamma, t) &= \mathcal{J}(W_{ij}^-(\gamma, t), W_{ij}^+(\gamma, t), \eta_{ij}), \\
B_{ij}(\gamma, t) &= B_{\widetilde{\Psi}}(W_{ij}^-(\gamma, t), W_{ij}^+(\gamma, t), \eta_{ij}), \\
S_{ij}(\gamma, t) &= S_{\widetilde{\Psi}}(W_{ij}^-(\gamma, t), W_{ij}^+(\gamma, t), \eta_{ij}), \\
\mathcal{A}_{ij}(\gamma, t) &= \mathcal{J}_{ij}(\gamma, t) - \mathcal{B}_{ij}(\gamma, t), \\
G_{ij}(\gamma, t) &= \frac{1}{2}(F_{\eta_{ij}}(W_{ij}^-(\gamma, t) + F_{\eta_{ij}}(W_{ij}^+(\gamma, t))) - \frac{1}{2}|\mathcal{A}_{ij}(\gamma, t)| \cdot (W_{ij}^+(\gamma, t) - W_{ij}^-(\gamma, t)), \\
\mathcal{P}_{ij}^-(\gamma, t) &= \frac{1}{2}\mathcal{K}_{ij}(\gamma, t)(\mathcal{I} - \mathrm{sgn}(\mathcal{D}_{ij}(\gamma, t))\mathcal{K}_{ij}^{-1}(\gamma, t),
\end{aligned}$$

where $\mathrm{sgn}(\mathcal{D}_{ij}(\gamma, t))$ is the diagonal matrix whose coefficients are the signs of the eigenvalues $\lambda_{ij,1}(\gamma, t), \ldots, \lambda_{ij}(\gamma, t)$ of $\mathcal{A}_{ij}(\gamma, t)$ and $\mathcal{K}_{ij}(\gamma, t)$ is a $N \times N$ matrix whose columns are associated eigenvectors.

Finally, the numerical scheme (31) can be rewritten in this particular case as follows:

$$\begin{aligned}
W_i' = &- \frac{1}{|V_i|} \sum_{j \in \mathcal{N}_i} |E_{ij}| \sum_{l=1}^{n(\overline{r})} w_l (G_{ij,l}(t) - B_{ij,l}(t) \cdot (W_{ij,l}^+(t) - W_{ij,l}^-(t)) \\
&- \mathcal{P}_{ij,l}(t)^- S_{ij,l}(t)(H_{ij,l}^+(t) - H_{ij,l}^-(t))) \\
&+ \frac{1}{|V_i|} \sum_{l=1}^{n(\overline{s})} \alpha_l^i \left(\sum_{k=1}^{2} B_k(P_i^t(\mathbf{x}_l)) \frac{\partial P_i^t}{\partial x_k}(\mathbf{x}_l) \right. \\
&+ \left. \sum_{k=1}^{2} S_k(P_i^t(\mathbf{x}_l)) \frac{\partial P_{N+1,i}^t}{\partial x_k}(\mathbf{x}_l) \right),
\end{aligned} \tag{41}$$

using the obvious notation.

5 APPLICATIONS

5.1 *Bed-load sediment transport in shallow water systems*

In this section we apply the proposed scheme to the case of shallow water equations with sediment transport, which can be written as an hyperbolic system with non-conservative term, under the structure of system (1).

We consider the system,

$$\begin{cases} \dfrac{\partial h}{\partial t} + \dfrac{\partial q_1}{\partial x_1} + \dfrac{\partial q_2}{\partial x_2} = 0, \\ \dfrac{\partial q_1}{\partial t} + \dfrac{\partial}{\partial x_1}\left(\dfrac{q_1^2}{h} + \dfrac{1}{2}gh^2\right) + \dfrac{\partial}{\partial x_2}\left(\dfrac{q_1 q_2}{h}\right) = gh\dfrac{\partial Z}{\partial x_1} - S_{F,x_1}, \\ \dfrac{\partial q_2}{\partial t} + \dfrac{\partial}{\partial x_1}\left(\dfrac{q_1 q_2}{h}\right) + \dfrac{\partial}{\partial x_2}\left(\dfrac{q_2^2}{h} + \dfrac{1}{2}gh^2\right) = gh\dfrac{\partial Z}{\partial x_2} - S_{F,x_2}, \\ \dfrac{\partial Z}{\partial t} - \xi\dfrac{\partial q_{b,x_1}}{\partial x_1} - \xi\dfrac{\partial q_{b,x_2}}{\partial x_2} = 0. \end{cases} \tag{42}$$

where the unknowns are the height of the water column $h(\mathbf{x},t)$, the discharge $q(\mathbf{x},t) = (q_1(\mathbf{x},t), q_2(\mathbf{x},t))$ and $Z(\mathbf{x},t) = H(\mathbf{x}) - z_b(\mathbf{x},t)$, where H is the bathimetry of the fixed bottom and z_b is the height of the sediment layer column. $\xi = \frac{1}{1-\rho_0}$, where ρ_0 denotes the porosity of the sediment layer and g the constant acceleration gravity. By q_{b,x_1} and q_{b,x_2} we denote the solid transport discharge for the two components x_1 and x_2 respectively. The definition of q_b depends on the considered model and usually depends on the velocity field of the fluid layer, the sediment grain size, the Manning coefficient, etc. In this paper we have considered Grass model and Meyer-Peter & Müller model (see (24), (30) and (26) for more details). S_{F,x_1} and S_{F,x_2} denote the sediment-water friction terms. Those terms are discretized in a semi-implicit way (see (3)), so they are not present in the finite volume scheme and in this section they are supposed to be zero. Therefore, the system of equations (42) can be rewritten under the structure of system (1) where,

$$W = \begin{bmatrix} h \\ q_1 \\ q_2 \\ S \end{bmatrix} = \begin{bmatrix} h \\ q_1 \\ q_2 \\ H - z_b \end{bmatrix}, F_1(W) = \begin{bmatrix} q_1 \\ \dfrac{q_1^2}{h} + \dfrac{1}{2}gh^2 \\ \dfrac{q_1 q_2}{h} \\ -\xi q_{b,x_1} \end{bmatrix}, F_2(W) = \begin{bmatrix} q_2 \\ \dfrac{q_1 q_2}{h} \\ \dfrac{q_2^2}{h} + \dfrac{1}{2}g\,h^2 \\ -\xi q_{b,x_2} \end{bmatrix},$$

$$B_1(W) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g\,h \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, B_2(W) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g\,h \\ 0 & 0 & 0 & 0 \end{bmatrix}, S(W) = 0.$$

In order to apply the numerical schemes (37) or (40) described in the previous section, the following matrices are used:

$$\mathcal{A}_{ij} = \mathcal{J}_{ij} - B_{ij}, \quad S_{ij} = 0,$$

where

$$\mathcal{J}_{ij} = \begin{bmatrix} 0 & \eta_{ij,1} & \eta_{ij,2} & 0 \\ (-u_{ij,1}^2 + c_{ij}^2)\eta_{ij,1} - u_{ij,1}u_{ij,2}\eta_{ij,2} & 2u_{ij,1}\eta_{ij,1} + u_{ij,2}\eta_{ij,2} & u_{ij,1}\eta_{ij,2} & 0 \\ -u_{ij,1}u_{ij,2}\eta_{ij,1} + (-u_{ij,2}^2 + c_{ij}^2)\eta_{ij,2} & u_{ij,2}\eta_{ij,1} & u_{ij,1}\eta_{ij,1} + 2u_{ij,2}\eta_{ij,2} & 0 \\ -\xi\dfrac{\partial q_b}{\partial h}(h_{ij}, u_{ij}) & -\xi\frac{\partial q_b}{\partial q_1}(h_{ij}, u_{ij}) & -\xi\frac{\partial q_b}{\partial q_2}(h_{ij}, u_{ij}) & 0 \end{bmatrix}, \tag{43}$$

and

$$B_{ij} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{ij}^2\eta_{ij,1} \\ 0 & 0 & 0 & c_{ij}^2\eta_{ij,2} \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{44}$$

being

$$\begin{aligned} \frac{\partial q_b}{\partial h}(h_{ij}, u_{ij}) &= \left(\frac{\partial q_{b,x_1}}{\partial h}\eta_{ij,1} + \frac{\partial q_{b,x_2}}{\partial h}\eta_{ij,2}\right)(h_{ij}, u_{ij}), \\ \frac{\partial q_b}{\partial q_1}(h_{ij}, u_{ij}) &= \left(\frac{\partial q_{b,x_1}}{\partial q_1}\eta_{ij,1} + \frac{\partial q_{b,x_2}}{\partial q_1}\eta_{ij,2}\right)(h_{ij}, u_{ij}), \\ \frac{\partial q_b}{\partial q_2}(h_{ij}, u_{ij}) &= \left(\frac{\partial q_{b,x_1}}{\partial q_2}\eta_{ij,1} + \frac{\partial q_{b,x_2}}{\partial q_2}\eta_{ij,2}\right)(h_{ij}, u_{ij}); \end{aligned} \tag{45}$$

and,

$$c_{ij} = \sqrt{g\mathrm{h}_{ij}}, \;\; u_{ij,\alpha} = \frac{\sqrt{h_i}u_{i,\alpha} + \sqrt{h_j}u_{j,\alpha}}{\sqrt{h_i} + \sqrt{h_j}}, \;\; \alpha = 1, 2 \;\; \mathrm{y}h_{ij} = \frac{h_i + h_j}{2}. \tag{46}$$

Proposition 1 *The schemes (37) or (40) with matrices (43) and (44) solves exactly the solutions corresponding to water at rest.*

5.2 *Tsunami generation model*

In this section we present a model to study the tsunami generation by submarine avalanches, given by a two-layer Savage-Huter type model (see (28)). For the sake of shortness, we present here a simplified model formulation. A more complex model is discussed in (11):

$$\begin{cases} \dfrac{\partial h_1}{\partial t} + \dfrac{\partial q_{1,1}}{\partial x_1} + \dfrac{\partial q_{1,2}}{\partial x_2} = 0, \\ \dfrac{\partial q_{1,1}}{\partial t} + \dfrac{\partial}{\partial x_1}\left(\dfrac{q_{1,1}^2}{h_1} + \dfrac{g}{2}h_1^2\right) + \dfrac{\partial}{\partial x_2}\left(\dfrac{q_{1,1}q_{1,2}}{h_1}\right) = -gh_1\dfrac{\partial h_2}{\partial x_1} + gh_1\dfrac{dH}{dx_1}, \\ \dfrac{\partial q_{1,2}}{\partial t} + \dfrac{\partial}{\partial x_1}\left(\dfrac{q_{1,1}q_{1,2}}{h_1}\right) + \dfrac{\partial}{\partial x_2}\left(\dfrac{q_{1,2}^2}{h_1} + \dfrac{g}{2}h_1^2\right) = -gh_1\dfrac{\partial h_2}{\partial x_2} + gh_1\dfrac{\partial H}{\partial x_2}, \\ \dfrac{\partial h_2}{\partial t} + \dfrac{\partial q_{2,1}}{\partial x_1} + \dfrac{\partial q_{2,2}}{\partial x_2} = 0, \\ \dfrac{\partial q_{2,1}}{\partial t} + \dfrac{\partial}{\partial x_1}\left(\dfrac{q_{2,1}^2}{h_2} + \dfrac{g}{2}h_2^2\right) + \dfrac{\partial}{\partial x_2}\left(\dfrac{q_{2,1}q_{2,2}}{h_2}\right) = -grh_2\dfrac{\partial h_1}{\partial x_1} + gh_2\dfrac{dH}{dx_1} + \mathcal{T}_{x_1}, \\ \dfrac{\partial q_{2,2}}{\partial t} + \dfrac{\partial}{\partial x_1}\left(\dfrac{q_{2,1}q_{2,2}}{h_2}\right) + \dfrac{\partial}{\partial x_2}\left(\dfrac{q_{2,2}^2}{h_2} + \dfrac{g}{2}h_2^2\right) = -grh_2\dfrac{\partial h_1}{\partial x_2} + gh_2\dfrac{\partial H}{\partial x_2} + \mathcal{T}_{x_2}. \end{cases} \tag{47}$$

In these equations, index 1 makes reference to the upper layer (fluid) and index 2 to the lower one (granular sediment). The fluid-sediment system is assumed to occupy a two dimensional domain. t is the time, and g is the gravity. $H(\mathbf{x})$ represents the depth function measured from a fixed level of reference. Each layer is assumed to have a constant density, ρ_i, $i = 1, 2$ ($\rho_1 < \rho_2$), $r = \frac{\rho_1}{\rho_2}$. The unknowns $q_{i,j}(\mathbf{x},t)$ and $h_i(\mathbf{x}, t)$ represent respectively the mass-flow in the direction j and the thickness of the i-th layer at point $\mathbf{x}$ at time t.

$\mathcal{T}_{x_1}$ and $\mathcal{T}_{x_2}$ denote the Coulomb friction terms. We observe that those terms must be understood as:

$$\text{If } |\mathcal{T}_\alpha| \geq \sigma_c \quad \Rightarrow \quad \mathcal{T}_{x_\alpha} = -g(1-r)h_2 \frac{u_{2,\alpha}}{|u_{2,\alpha}|} \tan(\delta_0), \quad \alpha = 1, 2. \tag{48}$$

$$\text{If } |\mathcal{T}_{x_\alpha}| < \sigma_c \quad \Rightarrow \quad u_{2,\alpha} = 0, \quad \alpha = 1, 2. \tag{49}$$

where $\sigma_c = g(1-r)h_2 \tan(\delta_0)$, being δ_0 the Coulomb friction angle.

Again problem (47) can be written under the form (1) taking:

$$W = \begin{bmatrix} h_1 \\ q_{1,1} \\ q_{1,2} \\ h_2 \\ q_{2,1} \\ q_{2,2} \end{bmatrix}, \tag{50}$$

$$F_1(W) = \begin{bmatrix} q_{1,1} \\ \frac{q_{1,1}^2}{h_1} + \frac{1}{2}gh_1^2 \\ \frac{q_{1,1}q_{1,2}}{h_1} \\ q_{2,1} \\ \frac{q_{2,1}^2}{h_2} + \frac{1}{2}gh_2^2 \\ \frac{q_{2,1}q_{2,2}}{h_2} \end{bmatrix}, \quad F_2(W) = \begin{bmatrix} q_{1,2} \\ \frac{q_{1,1}q_{1,2}}{h_1} \\ \frac{q_{1,2}^2}{h_1} + \frac{1}{2}gh_1^2 \\ q_{2,2} \\ \frac{q_{2,1}q_{2,2}}{h_2} \\ \frac{q_{2,2}^2}{h_2} + \frac{1}{2}gh_2^2 \end{bmatrix}, \tag{51}$$

$$B_1(W) = \begin{bmatrix} 0 & 0\,0 & 0 & 0\,0 \\ 0 & 0\,0 & -gh_1 & 0\,0 \\ 0 & 0\,0 & 0 & 0\,0 \\ 0 & 0\,0 & 0 & 0\,0 \\ -rgh_2 & 0\,0 & 0 & 0\,0 \\ 0 & 0\,0 & 0 & 0\,0 \end{bmatrix}; B_2(W) = \begin{bmatrix} 0 & 0\,0 & 0 & 0\,0 \\ 0 & 0\,0 & 0 & 0\,0 \\ 0 & 0\,0 & -gh_1 & 0\,0 \\ 0 & 0\,0 & 0 & 0\,0 \\ 0 & 0\,0 & 0 & 0\,0 \\ -rgh_2 & 0\,0 & 0 & 0\,0 \end{bmatrix}; \tag{52}$$

$$S_1(\mathbf{x}, W) = \left[\, 0;\ gh_1;\ 0;\ 0;\ gh_2;\ 0 \,\right]^T; \tag{53}$$

$$S_2(\mathbf{x}, W) = \left[\, 0;\ 0;\ gh_1;\ 0;\ 0;\ gh_2 \,\right]^T. \tag{54}$$

We apply again the construction of the numerical scheme developed in Section 4 taking as Roe linearizations at the edges for the one-dimensional problems the ones presented in (25). In this case

the matrices appearing in the definition of the numerical scheme can be written as follows:

$$\mathcal{A}_{ij} = \left[\begin{array}{c|c} J_{ij}^1 & -B_{ij}^{1,2} \\ \hline -B_{ij}^{2,1} & J_{ij}^2 \end{array}\right], \tag{55}$$

where

$$J_{ij}^k = \begin{bmatrix} 0 & \eta_{ij,1} & \eta_{ij,2} \\ (-u_{k,ij,1}^2 + c_{k,ij}^2)\eta_{ij,1} - u_{k,ij,1}u_{k,ij,2}\eta_{ij,2} & 2u_{k,ij,1}\eta_{ij,1} + u_{k,ij,2}\eta_{ij,2} & u_{k,ij,1}\eta_{ij,1} \\ -u_{k,ij,1}u_{k,ij,2}\eta_{ij,1} + (-u_{k,ij,2}^2 + c_{k,ij}^2)\eta_{ij,2} & u_{k,ij,2}\eta_{ij,2} & u_{k,ij,1}\eta_{ij,1} + 2u_{k,ij,2}\eta_{ij,2} \end{bmatrix}, \tag{56}$$

$$B_{ij}^{1,2} = \begin{bmatrix} 0 & 0 & 0 \\ -c_{1,ij}^2\eta_{ij,1} & 0 & 0 \\ -c_{1,ij}^2\eta_{ij,2} & 0 & 0 \end{bmatrix}, B_{ij}^{2,1} = \begin{bmatrix} 0 & 0 & 0 \\ -rc_{2,ij}^2\eta_{ij,1} & 0 & 0 \\ -rc_{2,ij}^2\eta_{ij,2} & 0 & 0 \end{bmatrix}, \tag{57}$$

and

$$S_{ij} = \begin{bmatrix} 0 \\ gh_{1,ij}\eta_{ij,1} \\ gh_{1,ij}\eta_{ij,2} \\ 0 \\ gh_{2,ij}\eta_{ij,1} \\ gh_{2,ij}\eta_{ij,2} \end{bmatrix}, \tag{58}$$

being

$$c_{k,ij} = \sqrt{gh_{k,ij}}, u_{k,ij,\alpha} = \frac{\sqrt{h_{k,i}}u_{k,i,\alpha} + \sqrt{h_{k,j}}u_{k,j,\alpha}}{\sqrt{h_{k,i}} + \sqrt{h_{k,j}}}, \alpha = 1, 2 \text{ y } h_{k,ij} = \frac{h_{k,i} + h_{k,j}}{2},\ k = 1, 2. \tag{59}$$

The numerical treatment of the Coulomb friction terms is critical in order to simulate property a tsunami generation and to preserve the well-balanced properties. The main idea is to treat explicitly the upwinding terms corresponding to Coulomb friction law by rewriting them as

$$\mathcal{T}_{x_\alpha} = \frac{1}{2}\mathcal{T}_{x_\alpha}\left(\frac{\partial x_1}{\partial x_1} + \frac{\partial x_2}{\partial x_2}\right), \quad \alpha = 1, 2,$$

while the centered terms are treated semi-implicitly.

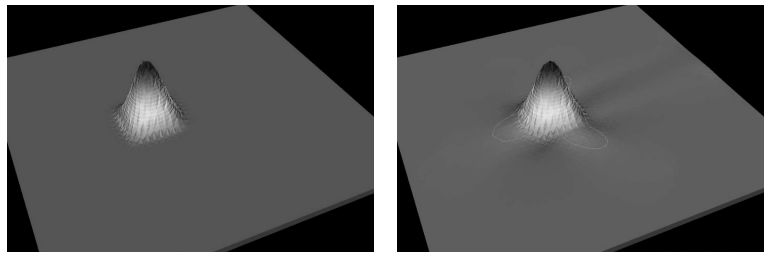
In this case, it is also easy to probe the corresponding *C Property*.

6 NUMERICAL EXPERIMENTS

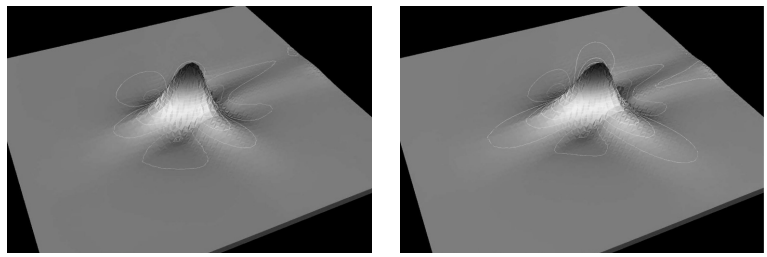
6.1 *Two-dimensional simulation of a conical dune of sand evolution*

This test purely two-dimensional (see (20)) studies the evolution behavior of a sand conical dune in a channel of dimensions $1000\,m \times 1000\,m$. Initial conditions are:

$$h(x, y, 0) = 10.1 - z_b(x, y, 0), \quad q_x(x, y, 0) = 10, \quad q_y(x, y, 0) = 0;$$



(a) 0 s. (b) 75000 s.



(c) 2600000 s (d) 3600000 s

Figure 2. Sediment layer evolution. Lateral view and color scale.

and the sediment layer is a sand dune with a conical form,

$$z_b(x_1,x_2,0)=\begin{cases}0.1+\sin^2\left(\dfrac{\pi(x_1-300)}{200}\right)\sin^2\left(\dfrac{\pi(x_2-400)}{200}\right) & \text{if } \begin{matrix}300\le x_1\le 500,\\ 400\le x_2\le 600,\end{matrix}\\ 0.1 & \text{other case.}\end{cases}$$

The layer evolution is modeled by Grass formulae (see (15))). We use the high order numerical scheme (40) with a non-regular MUSCL state reconstructions (see (12)). We consider a weak interaction between fluid and sediment layer, imposing $A_g=0.001$. Sediment porosity considered is 0.4. CFL condition is equal to 0.8. We use a mesh of 7600 edge finite volumes. As boundary conditions, discharge $q(0,x_2)=(q_1(0,x_2),q_2(0,x_2))=(10,0)$ and sediment layer thickness $z_b(0,x_2)=0.1$ is imposed at $x_1=0$. Free boundary condition is imposed at $x_1=1000$. At lateral walls we impose sliding condition $\mathbf{q}\cdot\eta=0$. With this data the experiment evolutes during 100 hours.

The sediment layer evolution is shown in Figure 2. Observe that the bed-load sediment transport produces a deformation of the original dune in a star form (see Figure 2).

Sediment layer is deformed gradually towards a star form, expanding along time with a certain angle (see (20)). It is possible to get an analytical approximation angle, supposing a weak interaction between sediment layer and fluid, that is, $A_g<0.01$ (see (20)). Analytical approximation of this expansion angle is,

$$\alpha=\tan^{-1}\left(\frac{3\sqrt{3}}{13}\right)=21.786789^\circ.$$

To measure the expansion angle, we have superposed level curves obtained in different time instants ($0\,s$, $120000\,s$ y $295000\,s$), and then expansion angle obtained is 23° (Figure 3). Note that we do not get a perfect symmetric dune figure in this simulation, as a non-symmetric unstructured mesh is used.

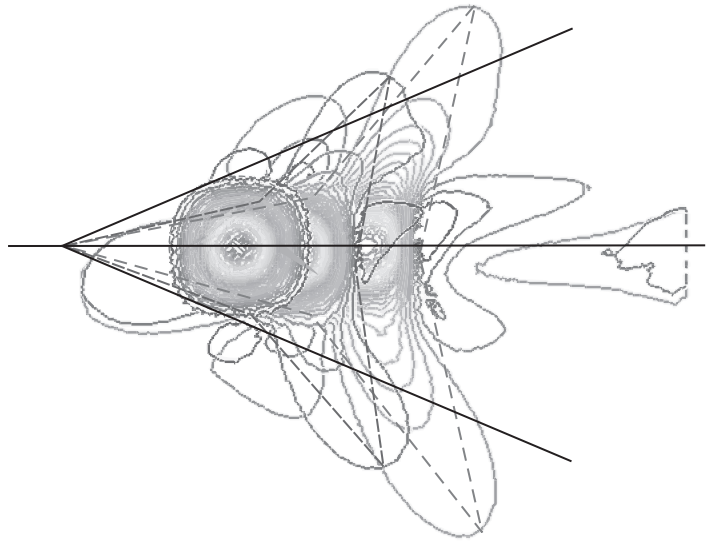


Figure 3. Spread angle estimation (using 20 level curves).

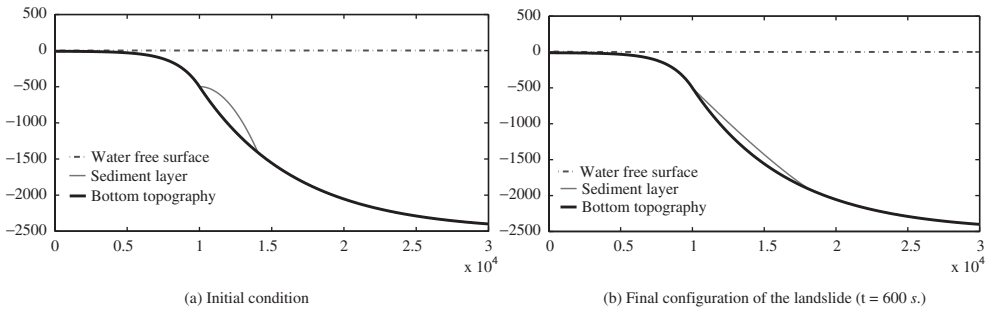


Figure 4. Tsunami experient: Longitudinal section at $x_2 = 0$.

6.2 *Tsunami generation experiment*

This numerical example is inspired in the one presented in the paper of Heinrich et al. (see (19)) and the objective it is to understand the main mechanisms of water wave generation and propagation produced by submarine landslides. To do that, a rectangular domain of 30 km long and 1 km width is considered, where a simplified topography is considered:

$$H(x_1,x_2) = \begin{cases} 10 + 490\exp^{-6.1429\cdot 10^{-4}(10000-x_1)} & \text{if } x_1 \leq 10000 \\ 2500 - 2000\exp^{-1.5050\cdot 10^{-4}(x_1-1000)} & \text{otherwise} \end{cases}$$

As initial condition we imposed $q_{i,j}(x_1,x_2) = 0$, $i,j = 1,2$,

$$h_2(x_1,x_2) = \max\left(H(x_1,x_2) - (500 + \frac{(x_1-10000)^2}{1.8\cdot 10^4}), 0\right), \text{ and } h_1(x_1,x_2) = H(x_1,x_2) - h_2(x_1,x_2).$$

Figure 4(a) shows a longitudinal cut of the initial condition at $x_2 = 0$. The Coulomb friction angle is set to $\delta_0 = 12^\circ$ (as in (19)). The density ratio is set to $r = 0.2$, CFL condition is equal to 0.8. We

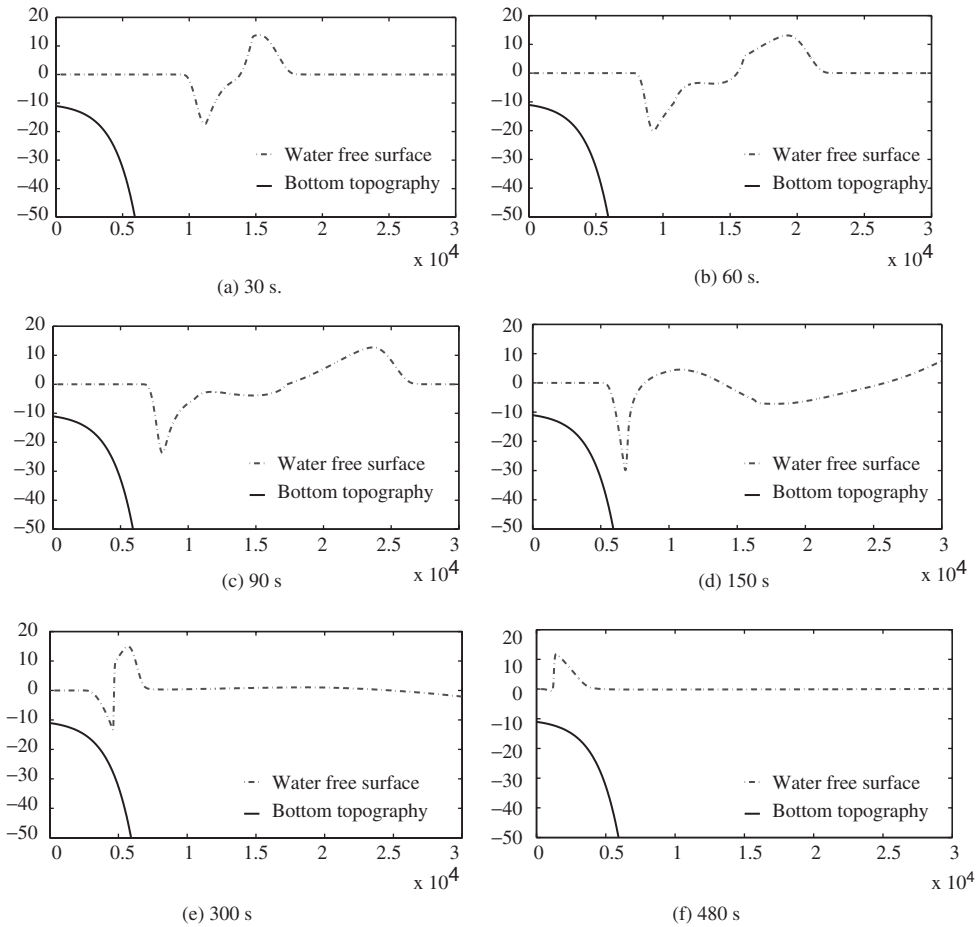


Figure 5. Tsunami evolution (zoom): Longitudinal section at $x_2 = 0$.

use a mesh of 10000 edge finite volumes. As boundary conditions, free boundary condition are imposed at $x_1 = 0$ and $x_2 = 30000$. At lateral walls sliding condition $\mathbf{q}_\alpha \cdot \eta = 0, \alpha = 1, 2$ is imposed. With this data the experiment evolves during 600 s.

Figure 5 shows a longitudinal section at $x_2 = 0$ of the generation and propagation of the tsunami wave. Observe that, water ahead of the front face of the slide is pushed away, creating a positive wave in the slide direction that represents the most of the tsunami energy. Above the landslide, water is sucked, which creates a large trough splitting into two waves, one wave propagating shorewards and the other offshore. Later, the shoreward-propagating wave is followed by a positive wave, responsible for coastal inundation and on which attention will be focused. Note that the second positive wave is of about 10m height when approaching to coastal line (see Figure 5(f)).

Finally, Figure 4(b) shows the final position of the landslide.

ACKNOWLEDGEMENTS

This research has been partially supported by the Spanish Government Research projects MTM2006-08075 and MTM2006-01275.

REFERENCES

Bermúdez, A., Vázquez-Cendón, M.E. 1994. *Upwind methods for hyperbolic conservation laws with source terms.* Computers and Fluids, 23(8):1049–1071.

Bouchut, F. 2004. *Nonlinear stability of finite volume methods for hyperbolic conservation laws and well-balanced schemes for sources.* Birkhäuser.

Brufau, P. 2000. *Simulación bidimensional de flujos hidrodinámicos transitorios en geometrías irregulares*. Phd Thesis, Universidad de Zaragoza.

Castro, M.J., Gallardo, J.M, Parés, C. 2006. *High order finite volume schemes based on reconstruction of states for solving hyperbolic systems with nonconservative products. Applications to shallow water systems.* Math. Comp., 75:1103–1134.

Castro, M.J., Fernández-Nieto, E.D., Ferreiro, A.M., García-Rodríguez, J.A., Parés, C. 2007. *Roe schemes for two dimensional nonconservative hyperbolic systems*. Preprint.

Castro, M.J., García, J.A., González, J.M., Parés, C. 2006. *A parallel 2D finite volume scheme for solving systems of balance laws with nonconservative products: application to shallow flows.* Comp. Meth. Appl. Mech. Eng., 196:2788–2815.

Castro, M.J., Macías, J., Parés, C. 2001. *A Q-Scheme for a class of systems of coupled conservation laws with source term. Application to a two-layer 1-D Shallow Water system.* ESAIM: M2AN, 35(1):107–127.

Chacón, T., Domínguez, A., Fernández, E.D. 2003. *An entropy correction free solver for non-homogeneous Shallow Water equations.* ESAIM: M2AN, 37(3):755–772.

Chacón, T., Domínguez, A., Fernández, E.D. 2004. *Asymptotically balanced schemes for non-homogeneous hyperbolic systems – application to the Shallow Water equations.* C. R. Acad. Sci. Paris, Ser. I, 338:85–90.

Dal Maso, G., LeFloch, P.G., and Murat, F. 1995. *Definition and weak stability of nonconservative products*. J. Math. Pures Appl. 74:483–548.

Fernández-Nieto, E.D., Bouchut, F., Bresch, D., Castro, M.J., Mangeney, A. 2007. *A new Savage-Hutter type model for submarine avalanches*. Preprint.

Ferreiro Ferreiro, A. 2006. *Desarrollo de técnicas de post-proceso de flujos hidrodinámicos, modelización de problemas de transporte de sedimentos y simulación numérica mediante técnicas de volúmenes finitos*. Phd Thesis, Universiad de Sevilla.

Gosse, L. 2000. *A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms.* Comp. Math. with Applic. 39:135–159.

Gosse, L. 2001. *A well-balanced scheme using non-conservative products designed for hyperbolic system of conservation laws with source terms.* Mat. Mod. Meth. Appl. Sc. 11:339–365.

Grass, A.J. 1981. *Sediments transport by waves and currents*. SERC London Cent. Mar. Technol., Report No. FL29.

Greenberg, J.M., and LeRoux, A.Y. 1996. *A well balanced scheme for the numerical processing of source terms in hyperbolic equations.* SIAM J. Numer. Anal. 33:1–16.

Greenberg, J.M., LeRoux, A.Y., Baraille, R., Noussair, A. 1997. *Analysis and approximation of conservation laws with source terms.* SIAM J. Numer. Anal. 34:1980–2007.

Harten, A., Hyman, J.M. 1983. *Self-adjusting grid methods for one-dimensional hyperbolic conservation laws.* J. Comp. Phys. 50:235–269.

Heinrich, Ph., Piatanesi, A., and Hébert, H. 2001. *Numerical modelling of tsunami generation and propagation submarine slumps: the 1998 Papua New Guinea event.* Geophys. J. Int, 145:97–111.

Hudson, J. 2001. *Numerical technics for morphodynamic modelling*. Phd Thesis, University of Whiteknights.

LeFloch, P.G. 1989. *Shock waves for nonlinear hyperbolic systems in nonconservative form*, Institute for Math. and its Appl., Minneapolis, Preprint 593.

LeFloch, P.G. 2002. *Hyperbolic systems of conservation laws, the theory of classical and nonclassical shock waves*. ETH Lecture Note Series, Birkauser.

Mangeney-Castelnau, A., Bouchut, B., Vilotte, J.M., Lajeunesse, E., Aubertin, A.,and Pirulli, M. 2005. *On the use of Saint-Venant equations for simulating the spreading of a granular mass*. J. Geophys. Res., 110:B09103.

Meyer-Peter, E., Müller, R. 1948. *Formulas for bed-load transport*. Rep. 2nd Meet. Int. Assoc. Hydraul. Struct. Res., Stockholm: 39–64.

Parés,C., Castro, M.J. 2004. *On the well-balance property of Roe's method for nonconservative hyperbolic systems. Applications to shallow-water systems*. ESAIM: M2AN, 38(5):821–852.

Peña González, E. 2002. *Estudio numérico y experimental del transporte de sedimentos en cauces aluviales.* Tesis Doctoral. Universidade da Coruña. Grupo de Ingeniería del agua y del medio ambiente.

Roe, P.L. 1981. *Approximate Riemann solvers, parameter vectors and difference schemes.* J. Comp. Phys., 43:357–371.
Savage, S.B., Hutter, K. 1991. *The dynamics of avalanches of granular materials frominitiation to run-out.* Acta Mech. 86:201–223.
Toumi, I. 1992. *A weak formulation of Roe's approximate Riemann Solver*. J. Comp. Phys. 102(2): 360–373.
Van Rijn, L.C. 1984. *Sediment transport (I): bed load transport.* J. Hydraul. Div., Proc. ASCE, 110: 1431–56.
Vázquez-Cendón, M.E. 1999. *Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry.* J. Comp. Phys. 148:497–526.
Volpert, A.I. 1967. *Spaces BV and quasilinear equations*, Math. USSR Sbornik, 73:255–302.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Modeling and real time flash flood forecasting in a Mediterranean basin

I. Villanueva
Ofiteco consultancy for the Mediterranean Andalusian real time flood forecasting, Málaga, Spain

ABSTRACT: During the last years there has been a digital data information revolution in hydrologic-hydraulic modeling through the fast processing of remote sensing techniques and their results distribution, particularly the airborne laser altimetry or LIDAR and its associated surface models. While Geographical Information Systems services and computers are ready to store and deal with those amounts of data, there are several remaining questions about how to use or upgrade mathematical predictive models for flood prevention. The author experience for joining and exploiting those results is explained in the context of a real time flood forecasting system focused on flash flood events located in Southern Spain. Other operational issues related to the introduction of uncertainty in models, global warming threats and future strategies or policies are also commented.

1 INTRODUCTION

Numerical modeling in hydrology and hydraulics has been feasible and reliable for calibration, modeling, capacity design of structures, etc., during the last decades, or from another point of view it has usually managed to reduce and simplify the problem definition from its formulation and offer valuable answers to engineers according to society needs with the increasing support of computational power growth, Abbott (1998). Nowadays the current amount of topographic and real time data or basically the digital input data for modeling reality is overwhelming the processing capacity and scientists and engineers are assessing the ways to cope with it, so from a past identification of essential data with estimated resolution to be compiled or surveyed we have reached the saturation point where selection or reduction between all the available and increasing digital data is necessary. On the contrary, other essential and required data are still poorly provided due to maintenance cost or other operational issues. Weak points or limitations are described within the selected models and solutions, particularly within the flood modeling using the two-dimensional shallow water equations in a mountainous region characterized by strong flash flood events.

The nature of the basin, discussed in Romero-Cordón et al. (2002) and Liquete et al. (2005), with presence of heavy rains and torrential flood events during autumn or spring seasons, makes convenient to select numerical models specialized in highly transient states, in other topographic conditions hydrodynamic simplified models have been extensively used, Estrela and Quintas (1994), Bates and De Roo (2000), but in this working basin complete shock-capturing schemes are needed to catch on the physics related to steep beds and propagating shores. For hydrological calculations, traditional methods like the US Soil Conservation System (SCS) have been substituted by schemes which include the use of digital terrain models like topmodel or topographical index based ones, Beven (2001).

After all of these considerations, according to the professional and research experience of the author, the present contribution is intended to be more focused in practical knowledge or applications rather than the description of new academic or theoretical models, and it is distributed as

follows: a second and third section about applied numerical models (hydraulic, hydrologic), a fourth one about realistic or operational considerations and future and upgrading tools for hydrological monitoring, warning and response systems and their potential threats, ending with a conclusions section.

2 HYDRAULIC MODELLING

2.1 *1D versus 2D hydrodynamic*

The author's own hydrodynamic modeling presented in this contribution is always based on the shallow water equations, well described with their assumptions and simplifications in García-Navarro and Jubillar (1999), or assumed to be concisely detailed in other chapters of this book.

In particular the selected techniques of explicit shock-capturing schemes based on the approximate Riemann solver of Roe have been developed, improved, and extensively tested during the last fifteen years by many authors: Alcrudo and García-Navarro (1993), Zhao et al. (1996), Anastasiou and Chan (1997), Burguete and García-Navarro (2001), Toro (2001), Connell et al. (2001), Beffa and Connell (2001), Bradford and Sanders (2002), Erduran et al. (2002), Guinot (2003), Brufau et al. (2004), and Murillo et al. (2006). The author's model is a simple explicit first order in space and time for irregular one-dimensional (1D) rivers García-Navarro et al. (1999) and a two-dimensional (2D) for rectangular grids Villanueva and Wright (2006a) which can operate separately or in a hybrid approach Villanueva and Wright (2006b, 2006c), and allows performance within multi-processor machines. The basic mass and momentum conservations in conservative formulation update the flow variables (U) in two dimensions according to:

$$U_{ij}^{n+1} = U_{ji}^{n} - \frac{\Delta t}{\Delta x}\left[F_{i+1/2j}^{*} - F_{i-1/2j}^{*}\right] - \frac{\Delta t}{\Delta y}\left[G_{ij+1/2}^{*} - G_{ij-1/2}^{*}\right] + \Delta t\left[H_{ij}^{*}\right] \tag{1}$$

with complete definitions and specifications in the cited references, and calibrations done with real flood inundations captured with radar airborne imagery Villanueva and Wright (2006a, 2006b).

Being this contribution more focused on operational issues rather than model definitions, in the following subsections results from 1D models against 2D models will be compared, as normally flood modelers in consultancies or agencies are still 'addicted' to 1D models; both steady or unsteady ones. Three simple cases are selected to highlight the advantages of the 2D modeling in particular geometries located in practical scenarios, and section 2.2 proposes a well known cluster computing methodology for reducing the 2D computational cost with another three testing cases. All is described with the purpose of showing that quality 2D computations with complete models is normally made in less than 12 hours, or equivalently during a night and the flood process is better covered and understood.

In Southern Spain most of the dam safety studies or urban building schemes done in the last 10 years concerning to the risk of flooding and delimitation of hazard areas were done mainly with one-dimensional commercial schemes as they were available at relatively cheap prices and with low computational demands; Intecsa-Inarsa (2003), Junta Andalucía (2007). Alternatively many of them are now being reviewed with 2D models. Most of the reservoirs are small ones (less than 100 Hm3 capacity) or are located so close to the seaside that the differences between 2D and 1D are expected to be not very important, as well as many rivers or streams are short and steep enough to present mainly 1D properties, but among dozens of studies we selected three representative cases where the differences are relevant and represent effects of common morphologies or structures that could appear or guide other studies or authors in the future.

2.1.1 *La Caleta and Velez study cases*

A couple of cases where the comparison is between a steady 1D back water profile (obtained using the popular HEC-RAS software) and a steady or converging solution from the proposed 2D hydrodynamic model are explained below.

The interest of La Caleta case is to compare the 'Y' type junction of two natural bending streams converging to a third straight channel in a high slope terrain (0.023), the junction is also a constriction as the final channel is 15 meters width and the upstream ones are over 20 meters both, figure 1a, indeed an under pressure design in a populated residential area.

The 1D cross sections were selected every 25–29 meters along two Kilometers over a 1×1 meters terrain grid, which is considered an accurate selection for HEC-RAS which computes results immediately, and the 2D study was done over that original 1 meter resolution grid with a total of 3.4 million cells for a few hours, with a 3.2 GHz speed processor.

Figure 1b, displays in yellow the 1D extrapolated extent and in blue the 2D one, for both upstream branch discharges of 150 m^3/s, (estimated originally with 1000 years return period, but



Figure 1. 1a, 1b: La Caleta case. Steep junction of two channels, steady water extent with 150 m^3/s each channel (1000 years return period, under review), in yellow with HEC-RAS, in blue with the proposed 2D hydrodynamic model.

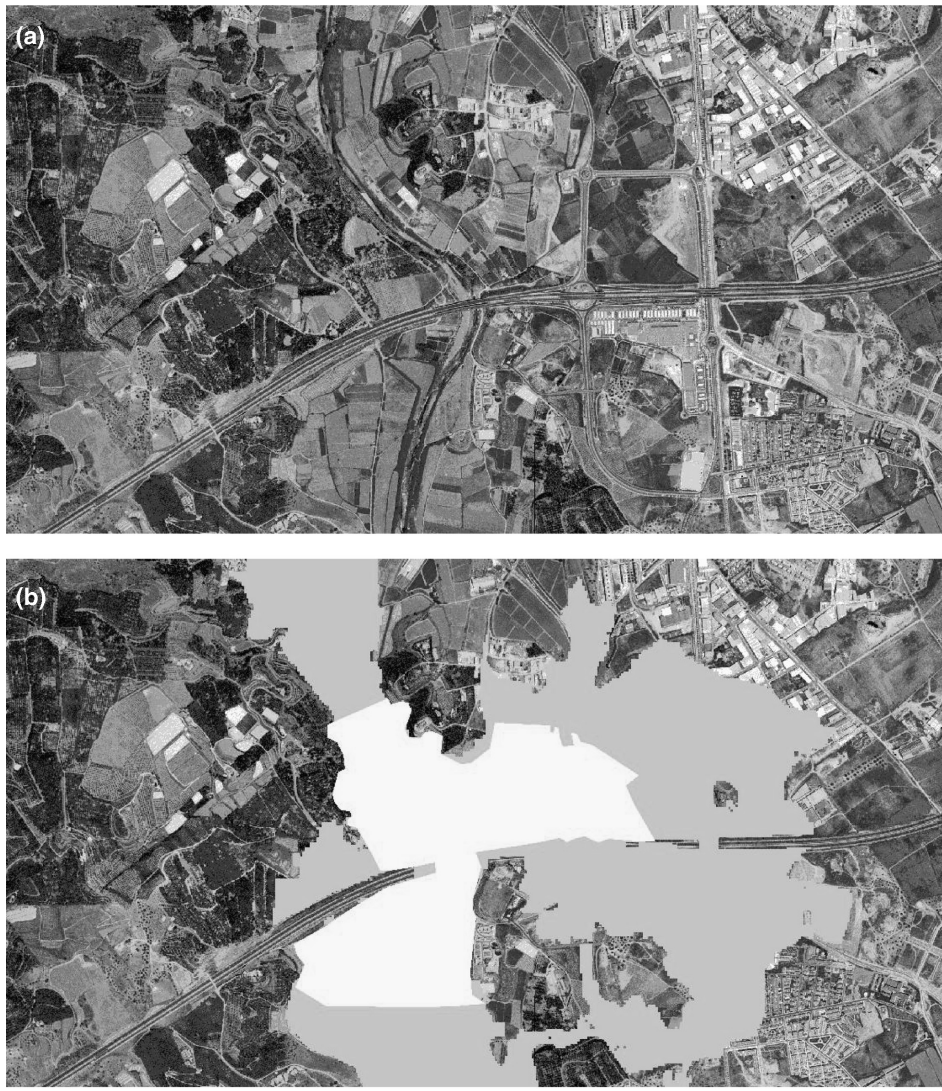


Figure 2. Velez river case. A river flow of 3700 m^3/s (500 years return period) crossing a highway embankment designed with an opening breach of 200 meters, flooding extent in yellow obtained with HEC-RAS, and in blue with the proposed 2D hydrodynamic model.

under review). It is remarkable the difference between the two models and their flood consequences over the area, due to the captured 2D back water effects over the upstream bends before the junction, resulting in a bigger and riskier flood extent area while using the 2D model which seems reasonable.

The Velez river study is focused on the effect of a highway embankment structure crossing over a dry river bed. Being the river bed with no meanders it can be considered itself as a one-dimensional structure but its interaction with the embankment is really two-dimensional, figure 2a.

In this case the 1D cross sections were taken every 75 meters, over a 13 Kilometers river length, within a 10×10 meters resolution grid, of totally 1.7 M cells.

There was no embankment or bridge overtopping in the study for an upstream 3700 m^3/s discharge (associated to a 500 years period return), but figure 2b displays the great differences on the flooding

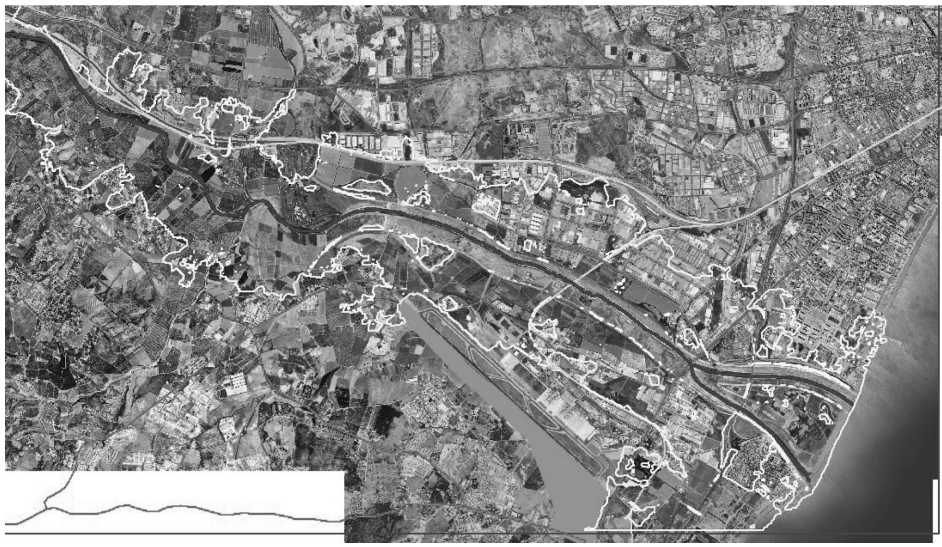


Figure 3. Flooding extent for a period return of 500 years at Guadalhorce downstream (yellow line), crossing a network of roads (red lines), bridges, airport and railways (light blue line).

extent, mainly due to the wall or reflection effect of the embankment captured with the 2D model in blue, which allows the water to spread and trespass downstream areas through the right hand side roundabout and road connections. It is remarkable the interest of this example due to the vast presence of these structures across highly developed areas with many road networks, see also as an example of complexity the downstream Guadalhorce river figure 3.

2.1.2 *Rules dam break*

A dam break study where the 1D hydrodynamic or transient model Mike11 from DHI software, Intecsa-Inarsa (2003), is compared with the proposed 2D hydrodynamic, as it follows.

A 90 meters height dam collapse with a stored volume of 130 Hm^3 is proposed as an extreme breaking hypothesis. Whereas the 2D well known model shows a slower arrival time of around 10–20 minutes over, figures 4a and 4b, this quantity has no important implications for the emergency strategy in a stream length of 19 Kilometers. It is assumed that 2D models are in general slower in peak propagation than 1D ones as reported in CADAM (2000), but it is difficult to generalize the reasons as 2D models are strongly dependent on the available topography.

Concerning the flooding maps, the 2D models will show in general broader flooding extents as they do not have the one-dimensional boundaries at each cross section, although in the delta area the 1D calculation simulates a wider area due to the intersection of cross-section areas and its interpolation in a GIS software, as it is shown in figure 4a, apart from that, the maximum extent is very similar and the differences do not affect urban areas in this case.

The computational two-dimensional cost is around 11 hours for a grid size of 10 meters and 6 million cells (at 3.2 GHz), but the short time taken in the pre-processing of data is really an advantage compared with the one-dimensional modeling which requires the splitting or branching of the downstream area and careful inspection in doing so and an interpolation process for flood curves within the GIS software.

It is not surprising that 2D transient models present in general lower peak results and they are softer or slower in propagation times than the 1D (or they need to increase the boundary values to reproduce strictly 1D results), due to the multiple reflections and expansions in the 2D domain and filling-emptying of flood plains, unless the 1D model tries to reproduce all those 2D topographic

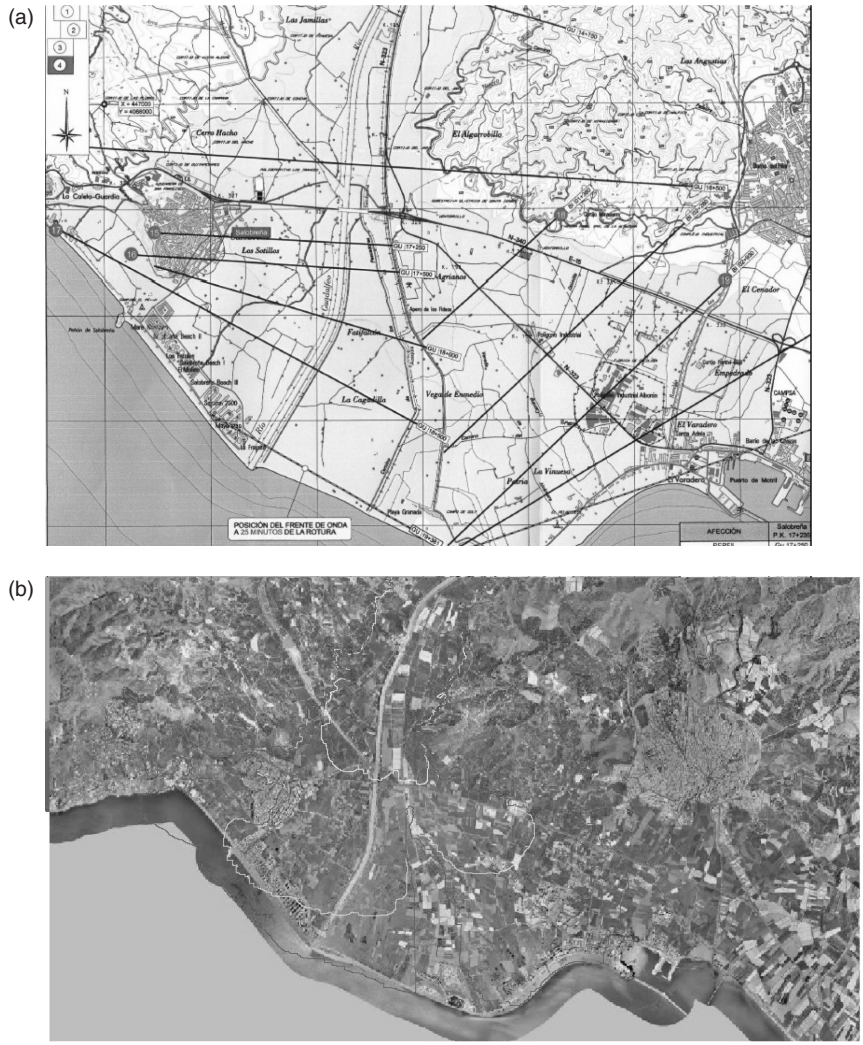


Figure 4. Dam break results with 1D Mike11(above, with pictured Xs) and 2D (below) showing the 29 (yellow), 36 (blue) and 49 (red) shore lines in minutes after the dam collapse, which are delayed around 10–20 minutes compared with the 1D calculations.

features (with concepts as flood storage cells) which is practically unrealistic and time consuming for the input data processing, see all references in CADAM (2000). That does not invalidate past 1D dam-break studies or make them wrong as their results are generally stronger and of riskier consequences for human interests.

Errors for mass conservation in 2D modelling were always less than 1%, this fact should be always checked within 2D modelling due to the presence of irregular topographies, Brufau et al. (2004).

The reported cases try to show three examples where the differences in the modeling approach are significant for the final results. They clarify that the consuming input data processing time in one-dimensional modeling does not result in improved quality, and the use of standard personal

computers with 2D models for small areas is not so expensive (1 night or less than 12 hours commonly).
Summarizing some clear advantages of 2D modeling are:

- Full employment of commonly adopted laser mapping or altimetry, although introduces another uncertainties (see sections below).
- Realistic consideration of volumes over the flood plains.
- Saving of hydraulic input data processing time once the provided digital model is accurate enough for the scale of the study, although to obtain that detailed digital surface model with a removal of vegetation and noise filtering process from remote sensing techniques is time consuming and introduces uncertainties, Neelz et al. (2006).
- Immediate integration of streets, junctions, deltas or river-coastal transitions or bays areas, with no interpolation process for output results.

While other aspects make one-dimensional models to survive like:

- Free internet distribution (as HEC family) and fastness delivery of results, easiness in defining concepts and boundaries, and extensively tested modules like the bridge calculation module that covers different regimes and threshold values for overtopping.
- More appropriated, simplified or well defined friction law according to Manning's law, which is usually extrapolated into 2D models.
- Operational issues like that one-dimensional modeling has to be kept within agencies or common practitioners for fast response models, conservation of many older calibrated events or because of the presence of clearly 1D hydraulic structures, or mainly due to the absence of bathymetry, Whitlow (2006). For hybrid modeling see also Villanueva and Wright (2006a, 2006b).
- Probably stronger hazard delimitation effects in terms of faster propagation time and higher peaks, at the expense of resulting in smaller flood extent areas, which is not always identified or expected before running the models.

Note that in this particular basin, the absence of bathymetry is no longer a problem as the majority of the river beds are dry, but it remains as a problem for reservoir capacity calculations.

Against the 2D modeling for big areas it is argued that fine grid resolution and its associated computational cost are very time consuming.

The right resolution issue is also linked directly with the presence of embankments, bridges, roads, railways, weirs, or culverts, as large grid cell resolutions erase, or do not capture the real height of those elements and it is needed to include them, being the results clearly dependent on those findings.

Moreover, urban or nearby urban flood modeling appears as a complex web of hydraulic structures, very difficult to classify or give priority only with remote sensing techniques or flight based techniques, see for instance downstream Guadalhorce in figure 3, and a field survey with GPS assistance is required with its associated further post-processing time and overlapping problems, which can delay considerably technical report deadlines. As a practical guide it is first recommended to the modeler to visually identify and keep those flow barriers in every resolution, modifying the terrain models with GIS programs (also known as burning of digital terrain models).

Resolution grid effects in urban modeling including buildings with small grids of less than 5 meters, are key subjects within the FRMRC activities for years to come, with some preliminary considerations described in Neelz et al. (2006) and Neelz and Pender (2006). In addition, the importance of the interaction with sewage or underground systems should be considered.

At this point the strong dependence on fine grids for quality modeling or capturing details is expected to be solved or rely on the increasing hardware or computational resources through the use of cluster facilities and associated tools in universities or research centers or alternatively cheaper multi-processor boxes which are becoming popular at consultancy offices, as it is described in the next section.

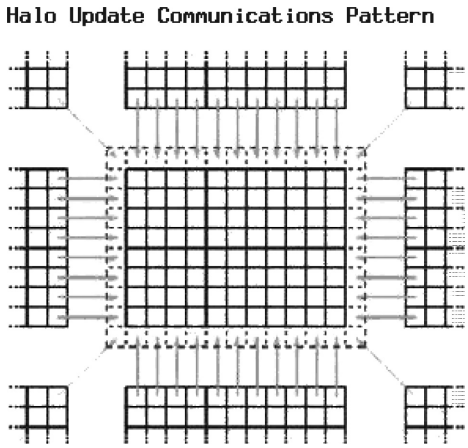


Figure 5. Domain division technique, with the assigned domain for one processor and the neighbouring communications (classic Halo division).

2.2 *2D hydrodynamic with clusters*

In order to decrease the computational cost of 2D Riemann solver based calculations, the basic domain division concept was adopted, it is well known on atmospheric or oceanographic flow calculations and it is also known as halo decomposition, MacDonald et al. (2005): the model domain is split up into 'chunks' and each processor works on a chunk, communicating with surrounding processors for necessary data exchange as water depth and velocities from neighbour cells and selection of common parameters like the minimum time step. The simplest and selected case for these studies is the rectangular grid division, being easily scalable depending on the number of available processors, figure 5, and reference Villanueva and Wright (2006c).

The transfer of water stage cell values is very simple for a first order scheme in space: it just involves the first neighbour column or row between adjacent processors, and the MPI (Message Passage Interface) libraries include many standard subroutines for common problems and synchronization processes making the implementation an easy task.

The presence of dry-wet interfaces or even complete dry domains make the computational cost or load no regularly obvious or difficult to define with a rule, at this moment the simple one equation 2 was detected for the cases run at Nottingham University high performance computing grid: the urban Glasgow case and the Siberian ice dam collapse of Altai.

It relates the ratio between two simulations (named 1 and 2) with different grid resolution or number of processors (or divided domains) as:

$$\frac{T_2}{T_1} = \left[\frac{\Delta x_1}{\Delta x_2}\right]^3 \frac{N_1}{N_2} C_f \tag{2}$$

where, T_1 is the computational cost with Δx_1 resolution grid defined in simulation 1, T_2 is the computational cost with Δx_2 resolution grid defined in simulation 2, N_1 is the number of processors for simulation 1 and N_2 is the number of processors for simulation 2. It is due to the squared ratio in number of cells and the explicit CFL criteria of linear time step dependence on grid resolution and includes the effect of a C_f parameter which accounts on the communication or synchronization delay factor that it is expected to grow if N_2 increases respect to N_1 or the simulation exceeds many days, and for the finished studies it was between 1.1 and 2.45 excluding three simulations from that mean expected range, at Nottingham cluster.

Table 1. Grid computing facility cluster, massive flood Siberian study numbers.

Dam height (m)	400	500	600	700
Reservoir volume ($M m^3$)	10 353	19 181	41 200	77 918
Run time (36 procs. at 2.2 GHz)	25 h : 12 m	31 h : 30 m	44 h : 5 m	73 h : 32 m

Unfortunately at the moment, much of this research is to be published elsewhere within the context of the FRMRC whole system modeling group or additional cooperations, Carling et al. (2007), and only the main magnitudes and tables will be reported in the following subsections as a demonstration of the robustness of the proposed techniques.

2.2.1 *The Altai Siberian flood: a palaeoflood calibration*

This study is thought to be the biggest ever freshwater flood discharge in planet earth history, it happened during the last deglacial period, Pleistocene and early Holocene, 40–13 thousand years ago in the mountains located in the common borders of Russia, Kazakhstan, China, and Mongolia. Nowadays through the conservation of several sediment bars and run-up deposits is possible to calibrate and evaluate through hydrodynamic simulations the maximum peak heights and reservoir volumes with their associated dam heights like in a 'forensic' methodology, Carling et al. (2007).

The scenario covers a main lake, its ice dam, a main channel downstream of 120 Kilometres with a downstream bifurcation of two rivers of 53 and 62 Kilometres respectively, with typical widths of the main valley of around a few Kilometres. The event produced a wave reaching heights of 250 meters, maximum speeds of 45 m/s and maximum discharges of 9–18 Million m^3/s. All the study area was covered with a terrain model of 50 meters resolution grid and 4.2 Million cells.

All these magnitudes were dealt with Riemann solvers without special modifications, allowing CFLs numbers of at least 0.5. While 1D hydrodynamic calculations suggest partial breaches finishing in a complete dam collapse between 350–400 meters height, the 2D calculations estimate a complete and higher dam body collapse of a maximum 700 meters height, which is more realistic according to the geometry and ice dam structure evolution, showing again that 2D models require stronger boundaries for reproducing 1D peak results.

The 1D model takes a few minutes to run along the main valley study, and the complete 2D model including the bifurcation with both downstream rivers takes for an instantaneous collapse and one day simulation time the following costs for 36 processors, see Table 1.

As commented before the main conclusions will be reported elsewhere and here it is just briefly described as a top magnitude case.

2.2.2 *A small urban test case*

In this case, a much more fitted to real life event than the previous one is studied; resolving surface water movement through urban areas requires resolution of complex flow paths around buildings, representation of micro-scale topographic and blockage effects (e.g. kerbs and walls, etc.) and numerical schemes capable of dealing with high velocity flows of often very small depth. A small urban catchment in Glasgow, Scotland, UK, was proposed as a benchmark study for several 2D models within the FRMRC whole system modeling group.

The selected as feasible resolution for a single desktop calculation is 2 m and contains 98 Kilo cells, whereas the selected testing multi-processor scale is 0.25 m which multiplies by 64 the total number of cells. This was a used as a validation check for the author's domain-division technique and confirmation that propagation times and peak heights are invariant through the use of appropriated different scales within the finite volume technique adopted or 2D model.

Table 2 displays the computational costs in hours and minutes using 2.2 GHz processor nodes within the grid facility described at the bottom.

Table 2. Urban flood computational costs for 0.25 m, and 2 m, with 1, 6, 12, 18 processors.

2 m (98 K cells)	1 proc.	1 h 37 m
0.25 m (6.3 M cells)	6 procs.	198 h 51 m
	12 procs.	141 h 9 m
	18 procs.	112 h 42 m

Table 3. Simple cluster computational costs in seconds for uniform hydrographs during 12 hours at 10–100 meters resolution with 1–6 processors at 3.2 GHz, for an approximate 50 kms river length, and 723 km^2 extent, (73 K cells at 100 m resolution, 7.3 M cells at 10 m resolution).

Q in (m^3/s)	T hydrograph (hours)	Dx = 100 m, P = 1	Dx = 10 m, P = 1	Dx = 10 m, P = 6
1000	12	5 m	168 h	64 h
5000	12	10 m	296 h 23 m	96 h
10 000	12	13 m	322 h 30 m	151 h 39 m
15 000	12	18 m	Est. 528 h	192 h 25 m

The hydrograph is 2 hours long with a peak of 10 m^3/s, and different approaches with different numerical schemes and results will be carefully described and published elsewhere.

As a technical note these techniques were started at the computer cluster available at Nottingham University consisted of a 256-processor tower, with 128 units with dual 2.2 GHz and local memory of 2Gb. Communication was through a dual port gigabit network. This resource is part of the Nottingham High-performance Computing Grid which in summer 2005 was ranked in position 109th of the top500 supercomputer list.

2.2.3 *Downstream Guadalhorce river*

A real case related to the main reservoir system (Conde de Guadalhorce, Guadalteba and Guadalhorce joint reservoir system) of the Mediterranean Andalusian river basin is now described. Instead of a University cluster facility an affordable Pc-box with two double core 64 bits processor was employed for the calculations.

Several dewatering or even dam break discharges were released (simulated) in order to compute warning response times and rural flooding extents.

A table for comparison of resolutions, number of processors and computational times is shown above, Table 3, for a desktop at 3.2 GHz with a maximum of 6 processors employed. Within that architecture the coefficient C_f referred in equation 2 reaches a top value of 2.8, a bit higher than the optimized cluster at Nottingham.

The convenience and improvement of the 10 meters grid against the 100 meters resolution can be appreciated in figure 6, although differences in propagation times at the control points were not found.

At this stage the author does not know about commercial packages that deal with cluster facilities within flood prevention schemes, and consultancies and agencies are still reluctant to evaluate or cooperate about these bespoke and university research techniques. Its introduction will require specialized tutoring and technical assistance offered by university groups, with the support and popularity of current available ensembles of processors at 64 bits, free GNU software as Linux operative systems and portable multi-parallel libraries as MPI, which allows to welcome with relative hope the uncertainty 'flood' associated with boundary conditions definition, lack of structure definitions in terrain models and other issues commented below in the uncertainty section.

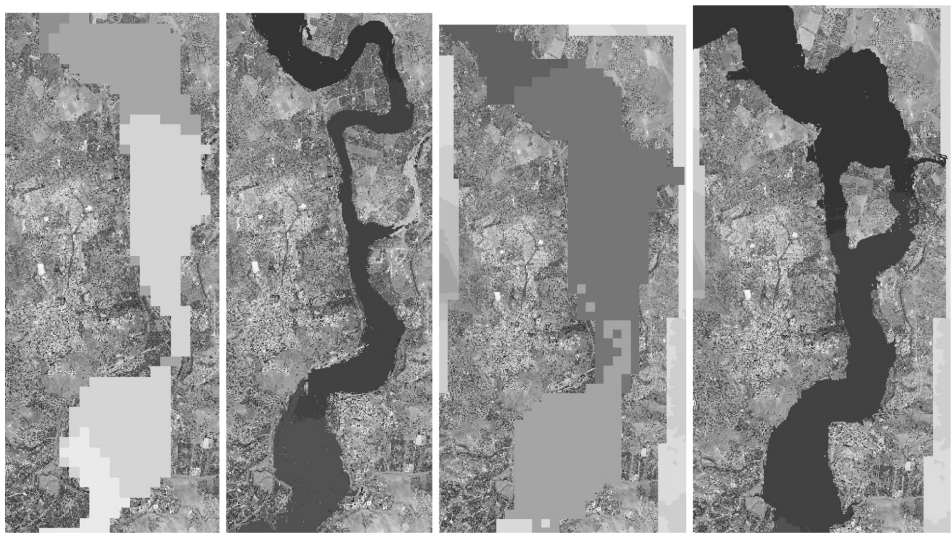


Figure 6. Flooding extents with different resolutions, (1st and 3rd blocks in green colour are with 100 meters, 2nd and 4th blocks in blue are 10 meters), for 1000 m^3/s (1st and 2nd), and 10 000 m^3/s, (3rd and 4th).

2.3 *Scales-resolution, and uncertainty*

No appropriate scales or grid resolutions gold rules can be addresses at this point, although a few advices and methodologies are described and could guide as help or reference, with a more industrial or end user based than a theoretical approach.

2.3.1 *The right resolution answer*

With all the described examples, bearing in mind that this research was done with raster or rectangular grids, avoiding the extra-computational efforts of unstructured meshes, a gold rule is not provided but a few conclusions are outlined.

Warning propagation times for reservoirs dewatering or dam break waves in the 50 Kms length downstream Guadalhorce river have been used with resolutions of 10 and 100 meters; the fastest approach of 100 meters is good enough and the 10 meters resolution does not provide additional information in terms of peak propagation for a warning response, in the Siberian case a grid of 50 meters was used for peak calibration. On the contrary, during flood extent calculations interaction with structures like embankments and roads is relevant, see figure 3, and the 10 meters resolution is at least needed for rural scenarios and a 5 meters one for urban; 2 and 0.25 meters were used in Glasgow case with no critical differences.

Classical uncertainties in river modeling like the Manning's end or vegetation estimation are reduced with remote sensing photographs, but others related to overtoppings are still present: embankments, connection structures below embankments like culverts, and bathymetries which are usually interpolated with old cross sections.

Uncertainties can appear as well in the boundary hydrographs, mainly if they are obtained using curve-rates from a water stage gauge, due to the wrong knowledge of the vegetation or friction slope at the moment of the event, pier effects if it is located over a bridge deck, or simply because when there is overtopping of the main channel cross section it is very difficult to determine the velocity across the complete section. Sometimes even there are inconsistent calibrations for the discharge between two gauges along the same stream.

All the commented issues require a large number of simulations to calibrate their influence on results (sensibility analysis), and make the process of flooding calculation a continuous or iterative process where flexible, no limited number of nodes and bespoke codes with open access

to the programming lines are essential for medium-long term projects, in addition, areas where fast building of infrastructures and development pressure are present (Mediterranean coastal ones) are continuously modifying their communications hub or 'terrain models' making the flood predictions obsolete in a few years.

A few tables and resolutions with commented examples were at least described as guidance or preparation for intensive computing cycles.

3 HYDROLOGICAL MODELLING

From the hydrological point of view, the nature of the involved Mediterranean basin where most of the river beds are dry and correspond to arid or semi-arid areas, with the presence of big aquifers and maritime intrusion problems present the total hydrological problem as a very complex one. Covering a total surface of 18 000 squared Kilometers, from west to east the annual mean rainfall varied across the registered rainfall gauges from 1608 mm (mountains of Gibraltar strait) to 156 mm (deserts of Almería), during the last 12 years, covering many micro-climates. Due to the presence of reservoirs and a registered desertification process spreading from eastern areas, the majority of the water courses are dried-out most of the seasons, with occasional flash floods in autumn, over low vegetation areas.

Being focused in the surface runoff problem, in order to make the best of the basic digital model at 10 meters resolution, a simplified distributed formulation based in topmodel concepts and assumptions; Beven (2001) and Piñol et al. (1997) for Mediterranean cases, has been selected, as a more advanced conceptual model than the SCS traditionally applied, and keeping a low computational cost.

It defines basically a mean catchment storage deficit D which evolves explicitly in time as:

$$\overline{D}_t = \overline{D}_{t-1} + [Q_b - Q_v]\frac{\Delta t}{A} \tag{3}$$

Defining the base flow Q_b as:

$$Q_b = Ae^{-\gamma}e^{\frac{\overline{D}}{m}} \tag{4}$$

with Q_v associated to the rainfall recharge, A the catchment surface, m and adjustable parameter related to the rate of transmissivity in the soil profile, and *gamma* the mean topographic index or geometric coefficient obtained as result of a sweep across the catchment raster model with the definition:

$$\gamma = \frac{1}{A}\sum_i a_i \ln\left(\frac{A_i}{T_o \tan\beta_i}\right) \tag{5}$$

with another parameter to adjust T_o related to horizontal fluxes, and A_i the upstream drainage contributing area for each cell of surface a_i and $\tan(\beta_i)$ as its medium slope.

The topographical index plot for every cell is also a good estimator of flood risk prone areas as symbolizes areas where the run-off flow should accumulate, even it can be used for searching for preferential areas where locate reservoirs using its numerical criteria. It has been shown that during the last years the reservoirs which highest percentage of filled capacity are the ones located in catchments where that index is higher in the Mediterranean Andalusian basin.

A yearly calibration for one of the few wet streams (a western one), river Guadiaro is selected as a successful calibration with a few parameters required, and with a good potential to be further improved on particular events. With a mean annual rainfall of 838 mm during the last 12 years and a surface of 623 km^2, a water stage gauge point is selected to be compared against the result of the described hydrological model. Although the conceptual modeling is a distributed one, it must be noticed that due to maintenance costs is very difficult to keep operational a dense network of rainfall gauges, moreover in mountainous regions where the access is very difficult, meaning that for every catchment calibration there are usually no more than two rainfall gauges as input data.

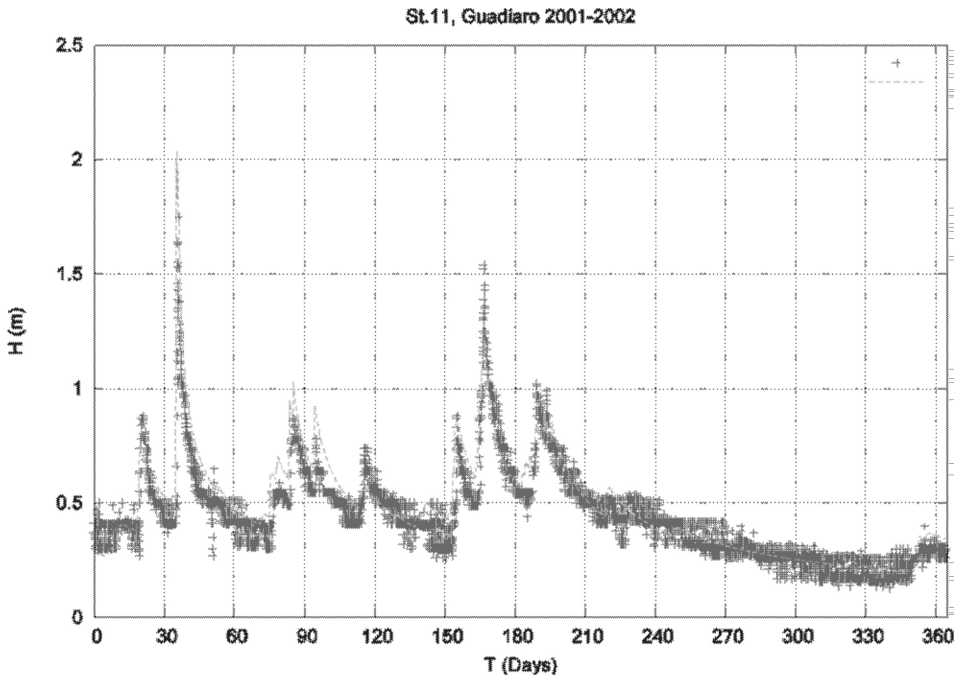


Figure 7. Hydrologic, yearly calibration for Guadiaro river, measured water stage (red points), topmodel estimation (green), hourly rainfall and timestep, with 10 meters resolution grid and gama = 11.27, t = 0.1, m = 0.02, surface = 623 Km^2.

Figure 7 shows the best results for the year between October 2001 and September 2002. The model run for 27 Million cells at 10 meters grid resolution with an hourly time step and rainfall input, in a few minutes with a 3.2 GHz processor.

After a calibration process for every year (starting from October) since 1994, the water depth itself and the calibration coefficients clearly show a decrease or slow drying out process: *m* coefficient varies gradually from 0.05 to 0.02, which means a fastening in the response time of the catchment or a loss of vegetation cover.

As summary on the search for a few parameters models the selected topmodel one has shown good agreements with real data, although a deeper analysis will be done in the coming years, larger catchments will be studied in order to couple run-off with transport models, and detailed flash or torrential scenarios will be selected. Mainly, this modeling technique reflected that once the catchment calibration for a non-dry river is done, the hydrographs are quickly generated provided an input of rainfall, and they agree with the observations at a gauge point.

Additionally many uncertainty factors could be enumerated, even more than in hydraulic modeling, due to the nature of hydrologic parameters related to infiltration and ground processes. Many authors as Beven (2001) explains in detail, are approaching them in terms of ensemble modeling and uncertainty work frames, including Monte-Carlo simulations and data-based models relying on artificial intelligence.

4 REAL-TIME SYSTEMS AND FUTURE TRENDS

In order to understand the integration of the previously described predictive models for a right exploitation within a real time flood forecasting system framework, many practical issues should

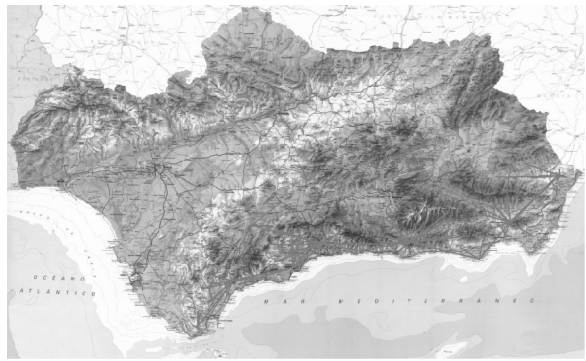


Figure 8. Mediterranean Andalusian Basin radio links between hydrologic stations in red colour. It is remarkable the irregular topography with the presence of mountains close to the sea with several 'white' peaks over 3000 meters height.

be understood or highlighted within that practical context, these are mainly related to the limited number of control or calibration points and their expensive cost of maintenance, their distribution in space according to non always the best 'hydrological location', the communication of uncertainty in the input and in the output of modeling, interaction with society or other information services like weather forecasting or emergency services, upgrading with recent technologies, etc. This section tries to enumerate many of them as a starting point for years to come, trying to identify weakest points rather than provide solutions or answers.

Firstly a particular basic description of the basin is needed, in the case of the Mediterranean Andalusian basin, available maps with gauge locations, photographs and descriptions of the architecture system and data exploitation are on line at hidrosur network website, for instance figure 8 displays the radio links between hubs and stations for exchange of data. That figure confirms that the distribution of stations was done according to direct eye-contact between points, locating the main hubs over the top of several mountains and preferably close to power supply networks (although they are assisted by batteries and a few are solar powered), that policy neglected many interesting points from an hydrological point of view. On the other hand an independent radio network service with own supply through batteries is highly appreciated in case of a total collapse of traditional communications and it is more stable during intense and heavy rainfalls than satellite based ones.

It is well accepted, and we cite references from Hadley (2005) and Tyndall (2006) research centers, that global warming will generate a decrease in rivers discharge following global climate models that predict a decrease in rainfall. All these effects are already confirmed in most of the whole Mediterranean Andalusian basin as many rivers run dry most of the seasons. From the modeler point of view these are also very bad news as less data will be available for future calibration of surface water models, with no possibility of feedback and with the prospect of very intense rains or flash flood events to happen.

As commented in the previous sections, detailed and accurate flood modeling is time-consuming, and can not be delivered on real time during an emergency, the actions would rely on fast and simple conversions of real time gauge water levels to discharges and associated flooding maps calculated previously for the main streams, for the minor ones without gauge control it is expected a warning message depending of nearby rainfall measurements and mainly a regular policy of prevention by cleaning the river beds.

Regarding to flooding maps, the best dissemination tool nowadays is done through web servers and client browsers or navigators. Particularly a free software tool like Mapserver has been tested successfully, providing the essential tools for layer and background management within a GIS, and including the portability and easiness of a common web browser. A clear policy of displaying hazard

areas and other flood related information on internet is desirable, Koch et al. (2006), although due to uncertainty and other social issues like investment in depressed areas or deals with insurance companies, sometimes it is difficult to publish all the results, but in any case, they should be at least easily accessed for emergency services.

Other threat is related to the self-maintenance of these systems, as they are becoming expensive for large areas, require a 24 hours service, the technology and consultancy markets evolve towards a fragmentation of the tasks, sharing satellite bands, land lines, gauges, at the expense of loosing unity and coordination every day and even worse during the presence of a natural disaster, as commented in NEDIES project (2002) and USGS web pages. A strong and coordinated steering committee is also needed for their effective continuity and fund raising.

Other unknowns related to recent technologies are for instance the contribution from mobile phone emergency calls, this information appropriately managed by a communication center could be very useful for location of flooded areas or propagation times, although as other satellite networks could collapse during an extreme event. The internet use and reliability under this pressure scenarios remains also as a doubt while in normal operation is the essential way of distributing information.

The way that the uncertainty chain propagates downstream: starting from the weather forecast to the emergency warning and plans concerning the population, it depends on so many factors, that is still very difficult to quantify or evaluate without real or experienced cases or lessons learnt from another similar services. Prevention, dissemination of learning to live with risk advices and attitudes should be available for the entire population.

Finally, as a more technical summary of technical points that are particularly needed to be strengthened in the future of Málaga's real time flood forecasting service, the following ones are cited:

- Rainfall gauges density: with a current one of 0.55 every 100 square Kilometers, is one of the highest in the Mediterranean arch, but there are many small-medium catchments with only one rainfall gauge, this could be improved for making better use of hydrological distributed models. The maintenance cost is the main problem for widening the network, although recent low consume electronic devices within wireless networks are expected to make cheaper networks, moving the main concern to the practical one of physical and legal installation in safe places against robbery, vandalism or animal damage. The difficult topography would imply the distribution of many hubs or concentration points for covering the whole area. It is expected not to depend on satellites as they could fail during very intense rains and because that increases the external dependency on telecommunication providers and decentralizes the organization of the warning system in case of emergency.
- Discharge gauge stations: the absence of flow meters in the main rivers makes the discharge measurement to rely only on water stage signals, and forces to estimate the discharge through rate curves assuming the vegetation distribution is known, this is reliable if periodical cleanings or debris removal over the dry beds are done. In reality most of the water stage gauges are placed over bridge decks and the modeler has to understand the bridge pier effects or cross section contraction.
- Rainfall radar and Meso-scale modeling: being the Mediterranean Andalusian basin a highly complex climatologic system with many sub-systems (ranging from 1500 milimeters rainfall per year in the Gibraltar West point and less than 300 at the eastern Cape of Gata), due to the number or mountains, strong winds and sea currents including the Gibraltar strait one, a dense rainfall radar network for avoiding wall echoes and highly detailed computational models would be required. Both techniques should be implemented for a long term basis and within research programs joining universities and/or research institutes. The current radar rainfall reflectivity images, provided by the National Institute of Metheorology network radars located at Málaga, Sevilla and Almería are insufficient. The presence of traditional rainfall gauges will be still required for calibration purposes.
- The commented basin is also becoming more dependent in aquifers for water supply and their consume and modeling needs to be controlled for long term periods.

– Internet communication service: it is the present façade of the service and an every day tool for reservoir monitoring and control of water resources, which are public demanded, it must provide GIS data layers dynamically, restricted real time data and deliver efficiently hydrological data from past events like rain distribution in tables and maps.

5 CONCLUSIONS

Some issues concerning flood prediction tools used in a monitoring and warning response system have been described. Due to the nature of the congress, the contribution has focused mainly on applications of well known numerical models adapted to accurate digital terrain models. It is considered that the provided solutions are neither ultimate nor cutting edge research, but they are practically interesting and cheap to implement considering the nature of the involved basin (flash floods in Mediterranean hilly areas) and the available limited resources. This could help as guidance or basis for similar installations which are expected to grow in developing countries due to global warning consequences and needs to control water resources.

In particular, it can be concluded that 2D complete hydrodynamic flood modeling techniques are mature enough to be applied extensively in water monitoring agencies in terms of robustness, but they require all the assumptions and the explained associated uncertainties in the input data to be carefully considered. New techniques like digital terrain data manipulation, grid computing or domain division will be extended requiring detailed learning periods for common practitioners.

ACKNOWLEDGEMENTS

As a former PhD student who attended the first edition of these congress proceedings in 1999, the author acknowledges Prof. García-Navarro's organization and research work, and hopes this present second edition is as useful as the first one and guides and encourages young researches as much as it did with the present author.

Several talks and meetings with members of FRMRC during a 3 years postdoctoral stay, particularly with Prof. N.G. Wright, currently at IHE-Delft, are appreciated. Hidrosur network data basis operational since 1991, field experience and practical advices are also gratefully acknowledged.

REFERENCES

Abbott, M.B., 1998. *Computational Hydraulics*, 2nd Edition, Aldershot : Ashgate.

Alcrudo, F. and Garcia-Navarro, P., 1993. A high-resolution Godunov-type scheme in finite volumes for the 2D shallow-water equations. *International Journal for Numerical Methods in Fluids*, Vol. 16, pp 489–505.

Anastasiou, K. and Chan, C.T., 1997. Solution of the 2D shallow water equations using the finite volume Method on Unstructured Triangular Meshes. *International Journal for Numerical Methods in Fluids*, Vol. 24, pp 1225–1245.

Bates, P.D. and De Roo, A.P.J, 2000. A simple raster-based model for floodplain inundation. *Journal of Hydrology*, 236, pp 54–77.

Beffa, C. and Connell, R.J, 2001. Two dimensional flood plain flow. I: Model Description. *Journal of Hydrologic Engineering,* Vol. 6, No. 5, pp 406–415.

Beven, K., 2001. *Rainfall-Runoff Modeling: The Primer*. Chichester: John Wiley.

Bradford, S.F. and Sanders, B.F., 2002. Finite-Volume model for Shallow-Water Flooding of Arbitrary Topography. *Journal of Hydraulic Engineering*, Vol. 128, No. 3.

Brufau, P., Garcia-Navarro, P. and Vazquez-Cendon, M.E., 2004. Zero mass error using unsteady wetting-drying conditions in shallow water flows over dry irregular topography. *International Journal for Numerical Methods in Fluids*, Vol. 45, pp 1047–1082.

Burguete, J. and Garcia-Navarro, P., 2001. Efficient construction of high-resolution TVD conservative schemes for equations with source terms: application to shallow water flows. *International Journal for Numerical Methods in Fluids*, Vol. 37, pp 209–248.

Carling, P., Villanueva, I., Herget, J. and Morvan, H., 2007. Unsteady 1-D and 2-D hydraulic models with ice-dam break for Quaternary megaflood, Altai Mountains, southern Siberia. Accepted In *4th International Palaeoflood workshop: Proc. intern. symp. Crete, 24–30 June, 2007*.

Connell, R.J, Painter, D.J. and Beffa, C., 2001. "Two dimensional flood plain flow. II: Model Validation", *Journal of Hydrologic Engineering*, Vol. 6, No. 5, pp 397–405.

Erduran, K.S., Kutija, V. and Hewett, C.J.M., 2002. "Performance of finite volume solutions to the shallow water equations with shock-capturing schemes", *International Journal for Numerical Methods in Fluids*, Vol. 40, pp 1237–1273.

Estrela, T. and Quintas, L. , 1994. Use of a GIS in the modeling of flows on floodplains. In *Proc. 2nd International Conference on River Flood Hydraulics*, York.

European Commission, 2002. *Guidelines on Flash Flood Prevention and Mitigation*. Colombo, A.G. and others. NEDIES project, Available at: http://nedies.jrc.it/doc/FlashFloods_Final.pdf

European Union Concerted Action on Dam Break Modelling: *CADAM*, 1998–2000. http://www.hrwallingford.co.uk/projects/CADAM/CADAM.

Garcia-Navarro, P., Fras, A. and Villanueva, I., 1999. Dam-break flow simulation: some results for one-dimensional models of real cases. *Journal of Hydrology*. 216(3–4): pp 227–247.

Garcia-Navarro, P. and Jubillar, E.P., (ed.), 1999. In *1st International Workshop on Numerical Modeling of Hydrodynamic Systems: Zaragoza, 21–24 June, 1999*.

Guinot, V., 2003. *Godunov-type schemes: an introduction for engineers*. Elsevier.

Hadley Centre, 2005, *Climate change, rivers and rainfall. Recent research on climate change*. Available at: http://www.metoffice.com/research/hadleycentre/pubs/brochures/B2005/COP11.pdf

Intecsa-Inarsa, 2003. *Emergency plan for Rules Dam*. Consultancy study for the South Hydrographic Confederation, Málaga, Spain. (In Spanish)

Junta de Andalucía, 2007. Estudio Hidráulico para la ordenación de las cuencas de la Costa del Sol Oriental (Málaga), Exp. A6.803.676/0411. Mediterranean Andalusian Basin, Málaga, Spain. (In Spanish)

Koch, A., Tzschirner, M. and Jupner, R., 2006. Flood management system- A newly developed instrument based on WebGIS and CMS and the improvement of basis data. In *Riverflow 2006. Proc. intern. symp: Lisbon, September 6–8, 2006.*

Liquete, C., Arnau, P., Canals, M. and Colas, S., 2005. Mediterranean river systems of Andalusia, southern Spain, and associated delta: A source to sink approach. *Marine Geology*, 222-223, 471–495.

MacDonald, N. and others, 2005. *Writing Message Passing Parallel Programs with MPI*, Edinburgh Parallel Computing Center Course Notes, Edinburgh.

Murillo, J., Garcia-Navarro, P., Brufau, P. and Burguete, J., 2006. Extension of an explicit finite volume method to large time steps (CFL > 1): application to shallow water flows, *International Journal for Numerical Methods in Fluids*, Vol. 50, pp 63–102.

Neelz, S. and others, 2006. Using remotely sensed data to support flood modelling, *ICE Journal of Water Management*, Vol. 159, March.

Neelz, S. and Pender, G., 2006. *The influence of errors in Digital Terrain Models on flood flow routes*. In *Riverflow 2006. Proc. intern. symp: Lisbon, September 6–8, 2006.*

Piñol, J., Beven, K. and Freer, J., 1997. Modelling the hydrological response of Mediterranean catchments, Prades, Catalonia. The use of distributed models as aids to hypothesis formulation. *Journal of Hydrological Processes*, Vol. 11, 1287–1306.

Romero-Cordon, E. and others, 2002. The catastrophic flood event of the 18th of October, 1973, in the rambla of Albuñol (La Rabita, Granada, Spain). In *Proceedings of the PHEFRA Workshop on Palaeofloods, Historical Floods and Climatic Variability: Applications in Flood Risk Assessment. Barcelona, 16–19 October, 2002.*

Toro, E.F., 2001. *Shock-Capturing Methods for Free-Surface Shallow Flows*. Chichester: Wiley.

Tyndall centre for Climate Change Research, 2006. *Understanding the regional impacts of climate change*. Available at: http://www.tyndall.ac.uk

Villanueva, I. and Wright, N.G., 2006a. Linking Riemann and storage cell models for flood prediction, *ICE Journal of Water Management*, Vol. 159, March.

Villanueva, I. and Wright, N.G., 2006b. Performance of several hybrid numerical schemes to determine flooding extent. In *Riverflow 2006. Proc. intern. symp: Lisbon, September 6–8, 2006.*

Villanueva, I. and Wright, N.G., 2006c. Efficient multiprocessor solver for the 2D shallow water equations. In *Hydroinformatics 2006, Proc. intern. symp: Nice, 4–8 September, 2006*.

Whitlow, C., 2006, personal communication within FRMRC meetings.

Zhao, D.H., Shen, H.W., Lai, J.S. and Tabios III, G.Q., 1996. Approximate Riemann Solvers in FVM for 2D Hydraulic Shock Wave Modeling, *Journal of Hydraulic Engineering.* Vol. 122, No. 12.

FURTHER READING, TOOLS OR IMAGES ON INTERNET:

British Flood Risk Management and Research Consortium (FRMRC), online information at http://www.floodrisk.org.uk
Hidrosur network website, http://hidrosur.agenciaandaluzadelagua.es
Mapserver web gis server, online information at http://mapserver.gis.umn.edu
MPI libraries, on line information at http://www-unix.mcs.anl.gov/mpi
Scientific Linux operative system, http://www.scientificlinux.org
United States Geological Survey, online information at http://www.usgs.gov
United States Hydrologic Engineering Centers River Analysis System modeling software (HEC-RAS), http://www.hec.usace.army.mil/software/hec-ras
United States MODIS rapid response system, images of Southern Spain, on line at http://rapidfire.sci.gsfc.nasa.gov/subsets/?AERONET_El_Arenosillo
United States NOAA air resources laboratory. http://www.arl.noaa.gov

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Water quality in urban hydrology

J. Puertas & J. Suárez
ETS Ingenieros de Caminos, Canales y Puertos, Universidade Da Coruña, Spain

ABSTRACT: This paper includes a fast overview of the problems included into "urban drainage" and "urban pollution control" topics. The governing equations are presented, and some real data are commented. The poor agreement between numerical and real data is pointed out, in order to encourage researchers to improve the equations and methods used, and in order to show some phenomena that are far from being properly modelled.

1 INTRODUCTION

1.1 *Urban water pollution*

Urban drainage projects are mainly concerned with the protection of the quality habitat of receiving waters. A major reason for the long term persistence of poor quality waters is the continued existence of uncontrolled or poorly controlled discharges from combined sewer overflows (CSOs) and/or surface water runoff. The impact on rivers and lakes caused by CSO events is accepted as an important source of pollution. This problem was first pointed out in the 1970s and has been studied over the course of the last few decades (Butler and Davies, 2000).

European regulations concerning water quality standards in rivers, which classifies the rivers according with the standards that different kind of fish need to live in, are mainly broken because of CSO discharges. Some countries have made an effort to determine pollution levels in the sewers as well as receiving waters, and structures and best management practices (BMPs) have been designed to control this pollution source (i.e. WRC, 1987).

Generally two types of the wet weather pollutant sources exist:

- Discharges form combined sewer systems (CSOs), by which a mixture of industrial wastewater, urban surface runoff, domestic wastewater and sewer deposits are discharged into receiving waters;
- Discharges form separate systems, by which mainly the runoff from urban surfaces is discharges.

These types of discharges are different with respect to the volume discharged, the concentrations of pollutants in the water phase and the periodicity of discharge. It is very important to know the maximum concentration of every pollutant, which permits to evaluate the acute impact (immediate effects) over the receiving waters and the habitats involved, and the event mean concentration, which permits to know the total mass discharged (immediate effects and long-term effects). These contaminants will affect stream organisms and alter the character of the ecosystem.

Pollution from rainwater discharges to receiving waters was first identified in early 1970s and this subject has been studied in several research projects, such as the Nationwide Urban Runoff Program, US-EPA (1983). From these projects, an order of magnitude was obtained to evaluate pollution from rainwater discharges. However the main result is the evidence of the complexity of the problem, with many degrees of freedom, and pollution values that show enormous variations

from one catchment to another. This is also true if we compare CSO pollution data. Soil use, population density, traffic intensity, etc., are parameters whose variations produce huge differences in pollution amounts (Butler and Davies, 2000).

As geographic and demographic aspects appear to be important, there may be noticeable differences among countries which have different city planning and sanitary policies. In Europe, some field work has been done. The projects by Ellis (1991) and others presented in Valiron & Tabuchi (1992) are worthy of mention. Both separate and combined sewer networks have been studied. The results differ substantially from one country to another, and even from one catchment to another within one country. No general trends have been defined. Some of the results on mean values are summarized below:

- Ellis et al. (1989): SS: 425 mg/l; BOD: 90 mg/l; COD: 380 mg/l
- Lager et al. (1977): SS: 370 mg/l; BOD: 115 mg/l; COD: 367 mg/l
- Metcalf & Eddy (1991): SS: 270–550 mg/l; BOD: 60–220 mg/l; COD: 260–480 mg/l
- NWRW (1991): SS: 105–320 mg/l; BOD: 40–124 mg/l; COD:148–389 mg/l

Even in Spain, some field work has been done to measure pollution loads in both combined and separate systems. The pollution values above have been considered too low for some real catchments in Spain, whose measured pollution values seem to be quite a bit higher. As these measurements are isolated tests done on a small number of CSO events, more accurate information about these phenomena would be necessary in order to ascertain the real COD concentrations of CSO events in Spanish catchments. The possible reasons for the increase in COD values in CSO events may be attributed to the removal of sediments from the sewer network. These sediments cannot be removed by wastewater, whose hydraulic energy is moderate, but can be removed by combined waters, which have more discharge, velocity, and shear stress.

1.2 *First flush*

If the removal of pollution from pavements is produced during the first stages of the storm, first flush phenomena have to be considered not only as a wash off of the catchment but also as a result of the cleaning of the sewer network. If sediment levels are substantial and if any colloidal or cohesive structures have been formed, the removal processes could be very slow and difficult to analyze, possibly generating rather large amounts of pollution during the full process without any relevant first-flush effects.

The first flush of pollutants has been identified in a relatively high proportion of the total storm pollution load that occurs in the initial part of the combined sewer runoff. Thornton and Saul (1986) and Pearson et al. (1986), mentioned by Gupta and Saul (1996), defined the first flush as the initial period of storm flow during which the concentration of pollutants was significantly higher than those observed during the latter stages of the storm event. Previous studies have highlighted the fact that the previous dry weather conditions, the length of the previous dry weather period, the magnitude and pollutant characteristics of the dry weather and storm flows, together with the characteristics of the sewer system, including the area of the catchment, soil use, cleaning policy, slopes of the catchment and the sewer system, all influence the temporal variability in the concentration and the load of the pollutants (Saget et al.,1996; Bertrand-Krajewski et al., 1998). These authors defined the first flush as being when at least 80% of the pollution load is transferred in the first 30% of the volume. From an analysis of 197 rainfall events in 12 separate and combined sewer systems, the authors showed that in separate sewer systems, 80% of the total pollutant mass is transported in the first 74% of the total volume for 50% of the rainfall events. In combined sewer systems, 80% of the total pollutant mass is transported in the first 79% of the total volume for 50% of the rainfall events. Whether or not a first flush does exist is very important when attempting to design CSO control structures.

The variation of the pollutant concentration during an event is described by the pollutograph. For a comparison of different events, a non-dimensional curve is more appropriate (Saget et al., 1996).

For this reason it is necessary to study the pollution load distribution. For each event and each pollutant, we used the cumulative load divided by the total pollution load versus the cumulative volume divided by the total volume of the event curves (curves L(V)). Each curve can be fitted by a function of the $Y = X^a$ type, where "Y" is the fraction of the discharged pollution load and "X" is the corresponding fraction of volume. The value of parameter "a" characterizes the deviation of the curve from the diagonal. These curves permit an analysis of the existence of first flush. Figure 7 includes two of these curves.

1.3 *Indicators of the pollution. Graphical analysis. Parameters*

Water quality parameters, which are indicators of pollution, can be grouped into families:

- COD, BOD5, TOC are considered indicators of oxygen consumption, as global organic and/or oxidizable matter.
- Turbidity, Total solids (ST) Suspended solids (SS), Dissolved solids (SD) and volatile fractions (SSV, SDV). Volatile solids are considered to have an organic origin. Suspended solids have adsorption qualities, which attract most of COD and metal loads. Their nature, however, is essentially inorganic.
- Kjeldahl Nitrogen (NTK), Ammonia (NH_4+), Total Phosphorus (P). These indicators, usually known as nutrients, have a domestic (bathroom, …) or similar origin.
- Metals (Cu, Pb, Zn, Cd). Their origins are the streets with heavy traffic loads. They are associated with suspended solids.
- Total Hydrocarbons (HC). Their origins are also the traffic effects.

These parameters have been measured (Chapter 3) both in dry weather conditions and during rainfall events. The graphic information consists of a series of hydrographs and associated pollutographs (i.e. Figure 10), where we can see if there is any decay in the pollutant concentration.

Apart from the graphic results, which give us a notion of the evolution of pollution during the event, general quantitative parameters have been extracted for each event, including:

- Maximum concentration (C.max) (mg/l) of each pollutant
- Event mean concentration (EMC) (mg/l) of each pollutant during the full event
- Total mass per surface unit (g/m^2) of each pollutant during the full event
- Mass maximum flux (kg/s) of each pollutant

The kind of information that can be extracted from these parameters is considerably different. So, maximum concentration and maximum flux tell us about pollution peaks. They may show a broad variation from one catchment to another, and they have some dependence on hydraulic values such as the slope of catchments and sewers. So, it is not possible to compare data if the catchment characteristics are not similar.

Also, the maximum flux depends on the magnitude of the catchment (area), since the full mass per second is measured in the control section. So, this value is very difficult to compare or extrapolate even in similar places if catchment sizes are not the same.

The total mass per surface unit provides a global value for the whole event, which is not conditioned by singular values, as is the case of maximum values. Given that this is not a concentration value, it is possible to have an idea of the pollutant mass which has been removed. If there is some source of clean water in the catchment, such as infiltration water through phreatic levels, concentration values decrease, but total mass per surface unit values do not change.

A reference value in the literature is usually the event mean concentration (EMC) which gives a mean weighted value for the concentration during the event. It is a value that can be extrapolated regardless of the size of the catchments, as it is a concentration value and it does not depend on the magnitude of the catchment.

2 EQUATIONS WORLD

2.1 *The phenomena and their equations*

Urban drainage is a very complicated matter, as the problems concerned are difficult and the degrees of freedom involved are, in most of cases, very numerous.

The equations involved, however, are well known, and they are not very difficult to solve. The phenomena are fully unsteady, but this is not a major problem. Some problems appear when trying to simplify reality into a model, which will be explained in chapter two, but once a model is generated, there are no numerical problems.

In fact, there are some commercial codes, as SWMM, MOUSE or INFOWORKS, which integrate the full problem, they are user-friendly and no problems can be reported when using them.

The equations that are usually used can be divided into three categories: hydraulic equations, dispersion equations, and solid transport equations. In the next paragraphs these equations are presented.

2.2 *Hydraulic equations*

Sewer networks can be considered as 1-D channel networks, whose hydraulic behaviour are modeled by the well-known Saint-Venant equations:

$$\frac{\partial y}{\partial t} + v\frac{\partial y}{\partial x} + \frac{A}{B}\frac{\partial v}{\partial x} = 0 \tag{1}$$

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} + g\frac{\partial y}{\partial x} + g\,(I - i) = 0 \tag{2}$$

where y: water depth; v: water velocity; A: surface area; B: width; I: energy slope; i: bed slope.

The water enters the network from the surface of the catchments. The flow over this surface can be modeled by using hydrologic methods (i.e. unit hydrographs) or, most commonly, by using hydraulic approaches, as kinematic wave.

Under these approaches, streets, roofs and parks are a sort of channel network, whose equations can also be (1) and (2), some simplified form of them, or even a 2D shallow water system, which is a more accurate way of modelling, but is not so common.

Most of commercial models, as SWMM, MOUSE, Hydroworks,. . . , apply Saint-Venant equations to sewer networks, and a simplified pack of equation to overland flow. Some calibration results will be presented in chapter 3, in order to evaluate the accuracy of the results.

2.3 *Dissolved pollutants*

Advection-diffusion equation is usually used to consider the behaviour of dissolved pollutants, including the dissolved solids, which are an important amount of the total solid load.

The equation must consider turbulent diffusion terms, as molecular diffusion is negligible. A vectorial form of the equation is presented (Equation 3), but it is often enough to consider a 1D equation, as mixing can be consider as complete in a few meters from the source point in most of cases.

$$\frac{\partial\,\overline{c}}{\partial\,t} + \nabla\cdot(\overline{c}\,\overline{\mathbf{u}}) = \nabla\cdot((\mathbf{D}_m + \mathbf{D}_t)\cdot\nabla\overline{c}) + s \tag{3}$$

where c: pollutant concentration; s: source; $\mathbf{D}_m$: molecular diffusion; $\mathbf{D}_t$: turbulent diffusion.

Some reactions can be expected, even in the sewer system; it is not common, however, to consider reactive models in such a turbulent media. There is not information enough in the literature.

As sewer systems admit a huge amount of pollutant sources, whose pollution degree is not easy to determinate, the practical application of this equation is not possible in most of cases.

2.4 *Solids removal. Bed load*

Suspended solids from the catchment are usually deposited in the sewer systems, so sediment transport equations can be used to analyze the removal and transport of these loads, which have been accumulated during dry periods, and are mobilized during short, wet periods (rainfall events). Equations of removal and transport consist of a set of two equations: one explaining the continuity of the sediment and the order explaining the relationship between shear stress and solid (bed load) discharge.

Continuity equation has been developed by Exner (1925). The one presented below have been modified by García (2000):

$$(1-p_0)\frac{\partial Z}{\partial t}+\frac{\partial q_b}{\partial x}+w_s(E_s-c_b)=0 \tag{4}$$

where p_0: porosity; Z: bed elevation; q_b: solid discharge; w_s: solid fall velocity; E_s: turbulent resuspension factor; c_b: equilibrium bed load concentration.

Shear stress equations are not analytical equations, but experimental or dimensional analysis-based algebraic equations (formulae could be a better name). Some dozens of such formulae can be found in the literature. Some of them are more famous, but there is not a full agreement about which is the best. In fact, there is some agreement about some of them are better than others under certain conditions, and worse under some other ones. In order to visualize the typical compositions of these formulae, two of them have been provided: a classical one, by Meyer, Peter & Muller (1948), which is implemented in many commercial codes:

$$\frac{q_b}{\sqrt{(G-1)gd_i^3}}=8(\tau_b^*-\tau_c^*)^{1.5} \tag{5}$$

where $G=$ relative density of the sediment, d_i: diameter of the sediment; τ_b: bed shear stress

$$\tau=\gamma R_h I \tag{6}$$

where: R_h: hydraulic radius, and the one by Nielsen (1992):

$$q_b=\left(\frac{12\tau_b}{\rho(G-1))gd_i}-0.05\right)\sqrt{\frac{\tau_b}{\rho(G-1))gd_i}}\sqrt{(G-1)gd_i^3} \tag{7}$$

As most of pollution is added to suspended solids, the proper evaluation of this flow is very important to model the full phenomenon. Unfortunately, the accuracy of these equations is very poor.

Besides, an important factor that takes place because of the organic nature of some solids is a certain degree of cohesion. Cohesion depends strongly on the chemical characteristics of the materials involved, and on the time from sedimentation. It is very difficult (almost impossible) to obtain parameters to model the cohesion equations. A minimum shear stress can be fixed as a threshold, but it is very difficult to find its value.

2.5 *Wash load*

Solids are generated in the catchment. They can be seen as "dirty and dust", which cumulate in the streets and the roofs and they are removed when it rains. The cumulation equations are exponential algebraic equation as:

$$M(t)=M_0(1-e^{-\alpha t}) \tag{8}$$

where $M(t)$: mass cumulated per surface unit; M_0: initial mass per surface unit; t: time; α: exponent.

The washing equations are also algebraic equations, as:

$$W=rI^{\beta}M \tag{9}$$

Where W: mass washed; M: mass available; I: rainfall intensity; r, β: fitting coefficients. The values of these coefficients depend strongly on factors as wind velocity, slope, and so on.

3 REAL WORLD

In order to analyze the goodness of the equations presented, some real data are to be presented, which will cover most of the phenomena involved. The data have been obtained in different cities, and in different experimental campaigns.

3.1 *Field work*

This paper includes a synthesis of information related to flow and contamination on the basis of three different sampling and analysis campaigns covering a total of six pilot catchments. Given that the objectives of each campaign varied and that they were conducted at different times, the data from the different campaigns are somewhat heterogeneous. The measurements had been taken in the lowest points of the different catchments, in the sewer systems in all cases. Measurements were not taken of runoff water or roof water.

The first data series is from the CSO Monitoring Program (PROMEDSU), funded by the Spanish Ministry of Environmental Protection. The results of this program, as related to COD, BOD and SS, already have been published (Suarez and Puertas, 2005). The above-mentioned publication describes the general methodology used in the sample taking. Since the other catchments were analyzed by the same research team, the methodology used was the same in all cases.

The main focus of the PROMEDSU program was to determine pollutant charges during CSO events in several experimental catchments in Spain, in order to obtain a pollution pattern which could be applied to any other catchment in Spain or, at the very least, to get an idea of the order of magnitude of pollution loads in Spanish catchments. Another topic of interest was the existence or non-existence of first-flush.

The catchments were chosen to cover different climatic regions and areas all over Spain. They all have combined sewer systems and the main use of the soil is residential, with some differences in population density.

One of the topics to be taken into account is the control of the inputs and outputs into the catchments. Preliminary work was done to ensure that the catchments observed had a single low point in their sewer systems, and that no parasite waters (inflows) were entering the system. In many cases, these preliminary studies were based on sparse or obsolete data about the sewer system, which have since been revised and completed during this project. The cities and catchments studied are presented in table 1.

Barcelona and Valencia are two of the largest cities on the Mediterranean coast, and their climates may be considered totally Mediterranean, with moderate annual rainfall heights, but a very great risk of high intensity storm events, with normal values of 100-200 mm/h per hour, for return periods of no more than 5 years.

In addition, both cities are quite flat, with sewer slopes of about 0.1–0.2%, so their sewer systems consist of enormous sections designed as flood evacuation structures, with CSO pollution being

Table 1. Cities and catchments analyzed.

City	Madrid	Sevilla	Vitoria	Valencia	Santiago	Santiago
Catchment	Arroyo del Fresno	Los Remedios	Almendra	Malvarrosa	Santiago-Este	Cancelón
Surface area (ha)	3800	135	132	89	265	80
Use	Urban – yard	urban	urban	urban	Urban and yard	
Climate	Continental	Semi-arid	Cont-Oceanic	Mediterranean	Atlantic	
Rainy days per year	63	52	103	44	141	
Annual rainfall (mm)	436	534	779	454	1886	

a secondary problem in most cases. During the last 10 years some effort has been made in both cities, and particularly in Barcelona, to limit CSO events by constructing huge reservoirs with a double aim: to limit the flood peaks and drive most of water to the WWTP.

Sevilla, located in the south of Spain, has a climate that can be considered semi-arid, with high temperatures and a moderate rainfall regime, and no pronounced episodes of heavy rainfall. As floods are not a major problem in the city, future investments are planned to reduce CSO pollution.

Madrid lies in the very center of Spain. Its climate is continental, with a low annual rainfall height, hot summer temperatures and cold winter temperatures. The receiving river is the Manzanares, whose source can be located in one of the WWTP of the city. Quality is the main issue in the municipality of Madrid, since the quality of the Manzanares river is extremely poor during CSO events.

Vitoria, located in the north central part of the country, has an oceanic climate, with rainy winters and moderate intensities. Vitoria is considered to have one of the highest standards of living of all the cities in Spain, and this municipality is highly concerned with environmental conservation. The river Zadorra, a tributary of the Ebro river, is the receiving river, and the improvement of the quality of this river is of top priority for the municipality of Vitoria.

The Atlantic area was not covered in this study, since its climate is markedly different from that of the rest of Spain. The other catchments included in this paper correspond to this area.

The aim of this program was not to determine the concentrations of metals, therefore the analysis of these types of pollutants was not carried out in depth. Daily mean samples were taken to measure the contamination of wastewater, and mean event-composed bottles were used for rainfall events.

In addition to the catchments included in PROMEDSU, two other catchments were studied – both located in the city of Santiago de Compostela, in northwest Spain. The first, "Santiago-Este", started operating between 1998 and 2000, while the Cancelón catchment began between 2001 and 2003. The characteristics of these catchments as well as their geographic location are also found in table 1. The city of Santiago is one of the wettest in all of Spain, with an average annual rainfall of over 1800 mm.

3.2 *Flow analysis*

This paper is mainly concerned with water quality, but a proper quantitative control is a previous condition that should be accomplished. Some calibration campaigns have been done, with models covering an important degree of complexity (kinematic wave-based hydrological models and Saint-Venant equations) and others much simpler.

The transformation rain water-runoff involves a process by which the precipitation that has fallen in an area runs down the streets and off the roofs until it reaches the inlets of the drainage network. After this process, the water circulates through the network of conduits divided into branches that meet and join together until it reaches an element downstream, either a wastewater treatment plant or a receiving system such as a river, a lake, the ocean, or a structure for the control of rain water.

3.2.1 *Rainfall-runoff transformation models*

Focusing on the process of rain water-runoff transformation, there are different calculation philosophies, two of the most commonly applied of which will be discussed here.

The method based on unit hydrographs has been more commonly used in rural catchments than in urban ones. However, references to this method are commonly found in the bibliography, and the commercial calculation models generally include a calculation package based on this philosophy. The unit hydrograph has a limited number of parameters (between 1 and not more than 3 or 4) which encompasses the hydraulic and hydrological information of the catchment. The purely geometric variables, such as the surface, are generally accepted as unquestionable, whereas the remaining parameters must be calculated and estimated. Anyone who has worked with these parameters is aware of the degree of uncertainty that arises once they are set.

The very nature of synthetic unit hydrographs (it is unusual to have unit hydrographs available based on real data) leads to the question of separating what is conceptual from what is real. The unit hydrograph proposed by Nash, based on waterfall tanks, accepts a non-whole number of tanks as a parameter, which means that his conceptual deduction results in a mathematical

formulation, which is more oriented towards flexibility than to the reproduction of physical phenomena. Many unit hydrographs practically coincide with probability distribution functions, i.e., functions whose form fluctuates widely with the variation in the value of their parameters. And this is where their ability to reproduce physical phenomena lies. What is debatable, however, is the physical sense of these parameters, their relationship to the natural phenomenon they seek to reproduce.

Perhaps the intention of giving a physical sense to the parameters is not in the nature of unit hydrographs, although, in order to evaluate them, it is indeed, necessary to appeal to their relationship to geometric and hydraulic variables. It is not uncommon for a parameter of a unit hydrograph to be associated with a complicated list of physical variables, obtained through a series of experiments, which, are in reality, merely an adjustment or a resolution of what is known as "back analysis" in optimization (refer to the calculations models implemented in Infoworks for the time constant K of the unit hydrograph of waterfall tanks -Nash-).The synthesized parameter does not have a concrete physical sense, and the relationship obtained, which does not come from a dimensional analysis, is nothing more than the best fit to a problem that does not appear to have a sole solution applicable to any catchment, due to the number of variables at stake.

Unlike the models based on a unit hydrograph, models based on the more or less reliable reproduction of the hydraulic processes that really occur in the catchment are presented. These are the most commonly used models and they are based on the Saint-Venant equations or on some of the simplified versions thereof, notably the application of kinematic wave equations or similar formulations (such as those used in SWMM model, which is, without a doubt, the most widely used on an international level.

In these types of models, as previously commented in chapter 2, the catchment is idealized in a series of maps that represent the roofs and street sections, connected to each other, each one having its own hydraulic characteristics (initial depression storage, friction coefficient, etc.), with a known physical sense. These hydraulic parameters, along with the idealized geometry make up a network of channels that are analyzed using the equations discussed earlier, accepting that rainfall is a runoff flow distributed over the subcatchment surfaces.

A model of rainfall-runoff transformation like the one presented here, naturally connects to the transport module of the runoff collected by the drainage network, using Saint-Venant's equations (or one of their simplifications, although in this case the full equations are generally used). In fact, from this standpoint, there is no conceptual difference between one phenomenon and the other. The transit calculation of the hydrographs collected at the inlets along the drainage network require a knowledge of their geometry, hydraulic characteristics (friction coefficients) and the existence of specific points (manholes, weirs, . . .).

3.2.2 *Information and simplification*

Modelling with the equations of Saint Venant for the calculation of flows generated by a rainfall event requires a good knowledge of the drainage network, as well as the characteristics of the flow inlet structures and urban topography, etc. If we attempt to structure a catchment in elements having homogeneous hydraulic characteristics, the discretization process leads strictly to the definition of an infinite number of small elements whose characteristics must be defined individually.

Since, for reasons of economy, it is not possible to undertake this task, a reasonable number of subcatchments are usually defined, each of which should include information of several processes. For example, even if we divide a block of houses into four subcatchments corresponding to their four slopes, if applicable, it is not viable to model each drainpipe and gutter and their connection to the drainage system. Even if we want to work with the utmost accuracy, this system is generally replaced by another with an hydraulic detention time corresponding to the entire transit for example. The same would be true of the water detained in these elements, etc.

The modelling process requires discretization and also simplification, and the parameters that have, strictly speaking, a clear physical sense (i.e., conduit's roughness coefficient) end up encompassing others whose effects are similar (head loss at manholes, retention time at inlet structures. . .).

If we accept this simplification, we also accept the fact that the parameters that we are working with, which are a synthesis of the infinite number actually involved, are not representative of specific phenomena, but rather additions. Therefore, their values cannot, or at least, should not be, what would be expected if we could really separate their effect.

Modelling a drainage network, including surface runoff and transport throughout the network, requires values for each and every parameter selected. From the point of view presented here, selecting "reasonable" parameters for the Manning coefficient of a conduit, for instance, is not a trivial issue, since a parameter may or may not be reasonable depending on the conceptual charge it has been assigned. This, in turn, depends on the discretization carried out, on the number of parameters considered to be representative, etc. The true problem is that this is not always taken into account when modelling and it is possible to make the mistake of assigning values to parameters with the help of a handbook, according to their designation. So, if we are evaluating Manning's coefficient for a concrete conduit, we would assign it an interval value of 0.012–0.016. It would be too gross an error for an engineer to forget to include the effect of the manholes; however, there are other implicit factors than may be (and often are) overlooked.

A method which has been accepted as reliable and which is used to the degree in which the means and time limits permit, is the comparison with real data. This is obviously not a valid tool for design development, since the comparison is carried out in the operational stage, but it is useful in the analysis of how a network functions, or to study its possible connection to other structures (detention tanks, flood tanks, first flush tanks) downstream.

The comparison is based on equipping the catchment with one or more rain meters and one or several flow meters. If only one is available, it should be placed at the lowest point of the network, where all of the waters intersect. The hydrograph recorded by the flow meter may override the one presented by the numerical model, after the rainfall record has been provided. By comparing the two, we will be able to determine whether or not the model reproduces the real situation in a sufficiently reliable manner, and the degree to which some parameters have been over or underestimated. Thus, if the recorded peaks and measurements do not coincide in time, we would have to assume that those parameters than entail delay (friction) have not been accurately estimated. If the final and initial volumes do not coincide, the problem would lie in the parameters that deal with infiltration and detention as well, and so forth.

With the data of a rainfall event (precipitation and flow), it is quite simple to adjust the parameters so that the measured and calculated hydrographs have a similar appearance. However, this does not guaranty that these corrected parameters will offer an accurate result in another event. If the same exercise is carried out on a succession of rainfalls, it would be logical to assume that the result would be a final set of averaged parameters with forecasting ability, even though the fit for each specific rainfall would not be as good as the fit that would be obtained if it were only applied to this event.

At this point, it would be helpful to know how many parameters we are talking about. In a numerical model, it is not very difficult to divide a catchment into fifty or a hundred units, each one having two or three significant parameters (apart from the geometrical ones, such as surface, form and mean slope, which would leave room for much discussion. In this paper, however, they are accepted as accurately measured). It is perfectly feasible to assign "reasonable" values to each one of these parameters with the help of a handbook, which results in a hydrograph. When we compare this hydrograph with the one that has been measured and accepted as real, and decide that the parameters are going to be modified, are we really going to change two or three hundred parameters? Even if it were humanly possible to do so accurately, it would be highly unlikely that a new set of parameters would have a good response to the next event.

The truth is that the engineer creating the model does not really choose this large number of parameters; the characteristics of the subcatchments are generally grouped together automatically depending on their use: gardens, streets, roofs of one type or another, etc. Each group is assigned a set of parameters with a maximum of twenty or thirty, which have, indeed, been chosen consciously and are different from each other.

When it comes time to fit these parameters, the number of recordings available (a hydrograph at the depression point) is obviously not enough for the number of parameters to be fit, in other

words, there is a large number of combinations of values of twenty parameters that will allow a hydrograph to be adjusted. It is no longer necessary to debate the "reasonable" character of these parameters, since, in a limited number of data, a phenomenon is being condensed with a greater degree of freedom, which is much more complex. The parameters now represent more than what their name indicates.

If, for a moment, we compare the two methods than have been presented for the estimation of hydrographs in a rain water runoff transformation, the unit hydrograph and the use of hydraulic equations, we can see that the number of parameters involved in the two methods is clearly off balance: there are no more than four parameters in a unit hydrograph and lots of them when the hydraulic equations are used.

Added to this is the fact that the conceptual structure of a unit hydrograph is very fragile, and its form is strongly dependant on its parameters, while the use of hydraulic equations and the application of the correct geometry, with fine discretization, is able to generate solid results, somewhat independent of the parameters. Unless they adopt unusual values, they serve to add precision to the value of the peaks and delays marked by the equations themselves. Thus, the three or four parameters of a unit hydrograph are in charge of defining the structure and the values of a real hydrograph, competing with a method whose structure is much clearer and which has a higher number of parameters. However the joint experience of many researchers shows that it is possible to obtain results having the same level of reliability with either method, if there is a comparison that can adjust them.

Is it really necessary to work with enormous quantities of parameters whose physical sense has actually been lost? Are we not allowing ourselves to be sidetracked by aesthetic aspects, when the basic question involves being able to predict flows associated with precipitation, regardless of the parameters of the catchment?

It is important to determine whether we are applying the parsimony principle philosophy to our studies of drainage systems, or if our intention is to use more parameters than those that are really necessary to explain the problem. We are faced with a problem of non-linear optimization, with an objective function defined in a twenty dimensional space, or the specific number of parameters we are working with. Both mathematicians and engineers who are specialized in optimization have addressed this problem with the utmost respect, whereas many urban drainage engineers accept the first combination of twenty parameters chosen for estimation, whose graph resembles the real result, and they consider each one to be correct. They even raise the question of whether or not their value is "reasonable".

3.2.3 *Black box models. AI based models*

In both the models based on the unit hydrograph and the models that reproduce hydraulic equations, there is a conceptual basis, or at least an a priori structure, so that the method of computation is predetermined to give a specific answer (flow) in the presence of a specific impulse (rainfall). The very essence of the unit hydrograph method is based on assigning a form (speaking in terms of graphs) to an exact excitation. This basic form is the a priori structure, and implies an attenuation, a transit, a delayed response to a specific action. Saint-Venant's equations also obviously require a structure – much stricter than the one mentioned above – since the equations have a greater conceptual basis.

These models compete with a family of forecasting methods that have no conceptual basis – they are the methods based on artificial intelligence, as neural networks and those known as genetic algorithms, which broaden the focus of the former.

The artificial neural networks are used almost routinely in some fields of engineering, and are starting to become a common tool in the field of hydraulics. The articles published on these models use their own terminology, which is closely linked to the computer field, distancing itself from civil engineering terminology, making the essential calculation method hard to understand.

Without going into the mechanical details, we will state that artificial neural networks are not based on any physical principle whatsoever and have no a priori structure. There are no component equations, the result from an input signal is not predetermined, either in form or magnitude, and

it depends entirely on the learning process of the system, containing mute functions whose values depend only on their parameters, and to which data is supplied on rainfall and runoff pairs which must come from real measurements. In this method, calibration precedes prediction, contrary to what occurs in the other methods described.

The use of these methods imposes the predominance of real data over the a priori hypotheses: Now, not only is the physical sense of the parameters questioned, but also the validity of the model itself, of its structure. The only active element of this method, which cannot be analyzed (since it is nothing more than a set of numbers, that make no sense if they are considered separately) or extrapolated (since the parameters are only valid for one catchment alone-remember that even the catchment area is imprinted on the connections) is its forecasting ability. The plain fact is that once trained, the system can accurately reproduce – as accurately as the other methods – the specific hydrograph of a rainfall event at the control section, and it does this economically in the sense that it saves a lot of human effort.

The number of parameters involved (there may be fifteen, twenty or even more, depending on the problem) is subject to the laws of the parsimony principle: However, we should not scrimp, since this set of numbers must be able to reproduce a structure – remember that there are no significant equations – in addition to establishing the specific values of the flow. So, the connections are more than just parameters. All the information, including the laws of physics, is imprinted on them.

3.2.4 *Calibration*

In the context of a more general study in terms of their objectives, the authors of this paper have monitored some of the urban catchments in the city of Santiago de Compostela. This paper presents two: the first has a combined sewer system ("Santiago Este") and the second has a separative type ("Fontiñas"). The study of the two catchments included the use of rain meters and flow records at downstream points. The catchments were discretized using a highly detailed geometric definition as a criterion, according to which the buildings were divided by slopes and each street segment was established as a subcatchment. This gave rise to divisions of around two hundred subcatchments for each case. This division is shown in Figure 1 for Fontiñas catchment. The discretization of the drainage network is similar.

The first calibration model used was SWMM. This model includes a hydrological type model with a philosophy that resembles the kinematics wave approach of Saint-Venant's equations to determine the surface runoff and a module which is based on Saint Venant's equations for the transport throughout the drainage network.

The process of calibration led to the determination of a parameter fit, and instead of choosing a manual adjustment for estimation, we preferred the numerical resolution model of the "back analysis"; specifically, the PEST (Parameter Estimation) model, which uses the well-known Levenverg-Mardquardt method. This program runs the SWMM model as many times as necessary until it minimizes the errors, changing the parameter values that are marked as adjustable.

The use of a parameter adjustment model makes it necessary to determine which parameters are to be fixed and which are to be adjusted. A good method of discriminating between parameters that are related to each other is to analyze the correlation matrix. High correlation between parameters bring about adjustment difficulties and indicate, from an optimization point of view, that the parsimony principle is not being respected.

The first attempts at optimization using fifteen or twenty parameters had very poor results: unchangeable parameters which showed several orders of magnitude, parameters which were completely correlated. Improving the adjustment with such a large number of parameters was not an easy task, as the numerical conditioners were poor and there was an absolute dependence on the seed.

The number of parameters was drastically reduced to a total of six, which contained the physical essence of the problem.

The initial performance was merely one of the many possibilities. If the seed is changed, there is a different result, however an acceptable adjustment of the data is obtained as well as some forecasting ability, in both cases, as discussed above: a non-linear problem with six parameters has different relative optimum adjustments, of which some are better than others.

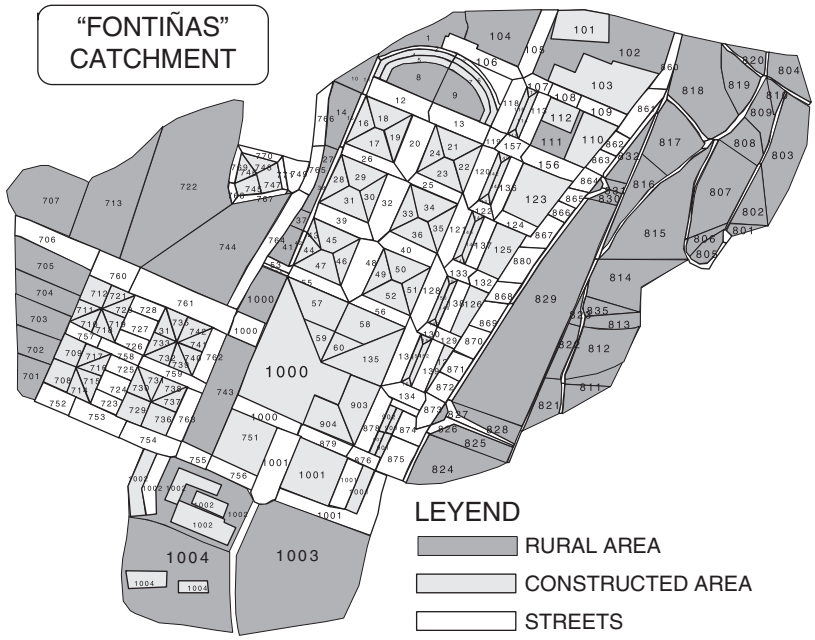


Figure 1. "Fontiñas" catchment and discretization.

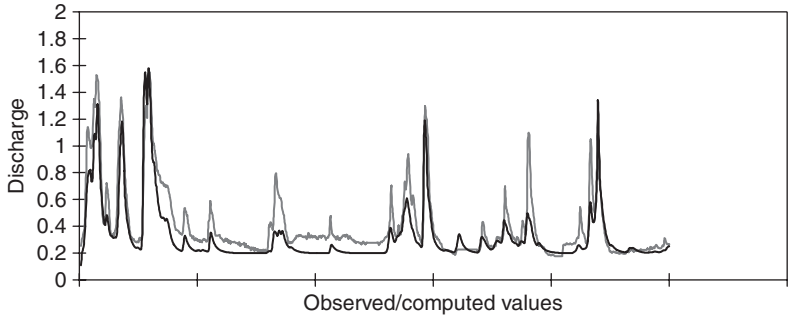


Figure 2. Results of the three parameter adjustment of the Santiago Este network.

The adjusted parameters do not have an intrinsic sense: strict restrictions have not been imposed – wide-ranging variation intervals have been defined – and this gives rise to seeds based on values extracted from a handbook resulting in adjustments that are very different from them. We can see that the two families of variables are different after the adjustment, but that the adjustment error is similar. Is it plausible to ask whether or not the parameters are reasonable?

After analyzing the correlation matrix of the performance with the six parameters, we can see that several of them continue to be correlated. Based on this, we selected partial combinations of between two and three parameters, which successfully and reliably reproduce the hydrograph measured. The rejected parameters should be fixed at a value within the working range (let us avoid the term reasonable) and all the uncertainty is transferred to the parameters that really will be adjusted – two or three parameters for one catchment which is geometrically divided into two hundred subcatchments. Figure 2 shows the adjustment obtained with the first series, whose visual appearance is agreeable. The performance represents an event of over 100 hours long.

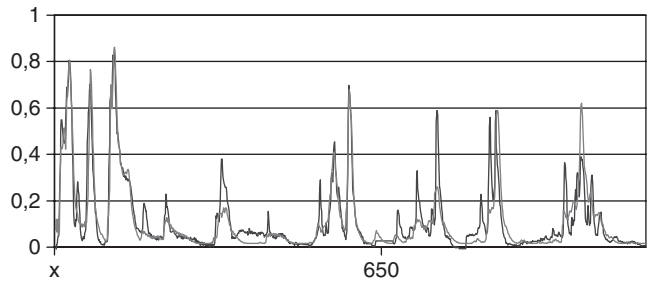


Figure 3. Adjustment and forecasts of discharge of the Santiago Este network and a neural network using 175 data, 650 as training and the remainder for the forecast.

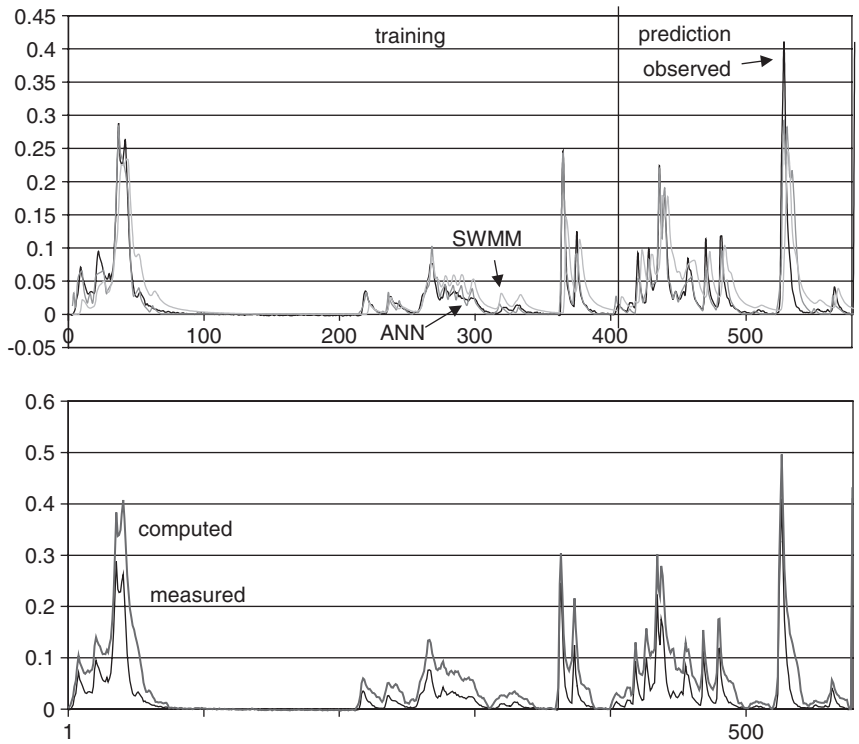


Figure 4. Adjustments and forecasts of the Fontiñas network using the SWMM model, using a neural network (up) and a simple model based on Nash's unit hydrograph (down).

The same data that were used in the calibration of the SWMM model were applied to a totally recursive network with five neurons, as outlined in figure 3. Without handling different types of architectures, and by exploring the different seeds and their results manually (and very rudimentarily), we easily obtained one (undoubtedly not the best) whose forecasting ability was comparable to (if not greater than) the best models obtained using SWMM. The network itself is meaningless, however it is a good forecasting model, as shown in figure 3, which presents the adjustments and predictions of the neural network compared to the data measured. We used 650 preliminary data in training and the rest as a forecast.

The comparison was repeated in the Fontiñas drainage network (Figure 4), which was also modeled using the simplest form of Nash's unit hydrograph (including transport as if it were the

Table 2. Real data obtained in five catchments in Spain. COD, BOD and SS.

Summary (Mean values) City	Madrid	Vitoria	Sevilla	Barcelona	Valencia
Qmax (m^3/s)	3.336	1.337	0.600	9.153	3.374
Qmean (m^3/s)	2.818	0.677	0.426	6.223	2.148
Volume(m^3)					
Total volume	17005	2994	2752	41566	12450
Runoff volume	7874	2303	2083	40653	11765
Runoff v./Total v.	0.485	0.743	0.682	0.966	0.884
Max. concentration (mg/L)					
COD	1344	1396	1717	1212	569
BOD5	792	521	759		341
SS	1189	966	1722	1434	561
EMC (mg/L)					
COD	680	671	834	456	293
BOD5	384	324	389		166
SS	597	421	733	578	229
Surface loads (g/m^2)					
COD	0.201	0.952	1.297	9.406	2.586
BOD5	0.106	0.419	0.546		1.488
SS	0.208	0.674	1.283	16.48	1.879
Mass discharges (kg/s)					
COD	3.616	1.040	0.501	6.46	0.834
BOD5	2.106	0.494	0.206		0.527
SS	3.425	0.724	0.558	10.35	0.632
Runoff mass/Total mass					
COD	0.703	0.793	0.820	0.977	0.887
BOD5	0.675	0.711	0.788		0.846
SS	0.823	0.844	0.896	0.981	0.877
Shape parameter “a”					
COD	0.738	0.685	0.775	0.797	0.864
BOD5	0.727	0.747	0.765		0.847
SS	0.702	0.550	0.743	0.800	0.896

sole natural catchment). Without a doubt, the best model is the neural network, but all three accurately reproduce the aspect of the phenomenon studied using, in the case of the conceptual models, no more than three parameters. The neural network under consideration uses twenty-five, which is obviously not the best according to the parsimony criteria. It is used by way of an example.

We should not make the mistake of thinking that because neural networks have a large number of parameters, that they do not play fair, as, being extremely flexible, they can adjust almost any curve. The calibration stage, where there is a reference that we must adhere to, should not be considered in the comparison. Obviously, with twenty-five parameters, any form can be adjusted. The real interest lies in the forecast, where there is no reference, and where each model produces a result without the least bit of support. It is here that we must determine which one is better, which one works and which one does not, and in this process, it is not an advantage to have either more or fewer parameters.

3.3 *Pollutant loads: COD, SS, metals*

An important number of pollutant parameters have been analyzed during field work campaigns. In order to comment the reliability of the equations, or, better than this, the reliability of the hypotheses in which equations are based on, some real data are to be analyzed, mainly considering COD, suspended solids and metals. The summary of data obtained is presented in table 2.

The main points to be highlighted after an analysis of Table 2 are as follows:

- Even though the Madrid catchment has the largest surface area, higher discharges and volumes are observed in Barcelona and Valencia, due to their Mediterranean climate, with high rainfall intensities.
- The maximum concentration and EMC values for COD, BOD and SS are very similar in Madrid, Sevilla, Vitoria and Barcelona (excluding BOD). Clearly lower values have been measured in Valencia, and this is due to the existence of a considerable amount of parasite inflows from an irrigation system.
- The highest surface load values were found in Barcelona and Valencia. The most likely reason is the removal of pollutants from the sewer system, which, in both cases, is extremely flat with huge sections designed for the evacuation of floods. When low discharges circulate along such large channels (rectangular boxes over 6 m wide), the water velocity is quite slow and sedimentation may occur. High discharges promote the removal of these sediments, and high surface load values are recorded. A separative sewer system would probably avoid most of the pollution in these catchments, as the optimum sizes for a wastewater sewer system and a rainwater sewer system have too many differences to work properly as a combined system. The maximum flux, which is more difficult to analyze, also shows the highest values for Barcelona's catchment.
- Most of the mass measured during a rainfall event is associated with runoff. This pollution may come from the catchment or the sewer system, which would seem to be the main origin of the pollution measured in Barcelona. Catchments with high rainfall intensities seem to have higher runoff percents.
- A slight first-flush was noticed in every catchment, but in quantitative terms, a device which only recorded the first part of the hydrograph would not show a good enough efficiency. The mean value (0.76) means that the pollutant mass which is removed if the first 30% of the runoff is stored (which is quite a large amount of water) is less than 50% of the total mass.

A preliminary statistical analysis of the data has been done, including single plots for every catchment and fully grouped data. Lognormal distribution provided the best approach with our data, in terms of the Anderson-Darling statistic, for both C. max and also EMC values, which will be presented. As pointed out by Butler and Davies (2000), most authors used this probability distribution.

The main conclusions to be drawn from the grouped probability functions are that they offer the most useful percentiles for EMC. As clearly seen here, very high values were obtained for EMC, which represents the mean value of the pollution during the full event. COD amounts of roughly 500 mg/l during the mean event mean that Spanish catchments cannot be designed with standards such as the ATV-128, whose pollution CSO values are around 200 mg/l, in terms of COD. The capacity of the structures designed to avoid CSO events must be able to ensure that very few events take place or there must be an improvement in the cleaning policy of the sewer network and the streets in order to reduce the amount of pollution that can be removed during a rainfall event. In any case, further research is necessary from a practical point of view, in order to establish a standard which would be considered useful for Spanish catchments.

As a conclusion, neither dilution nor diffusion is observed. The pollutant amounts are much bigger than the expected ones. The reasons can be found on the accumulation and reaction processes in the sewer network, which are extremely difficult to model, and the removal of these sediments.

So, removal of solids seems to be more important than diffusion, which is not relevant in this kind of analysis.

It is interesting to consider the behaviour of metals. Metals can be found in both suspended and dissolved solids. When analyzing suspended metals, a clear first flush is observed. When measuring dissolved, no first flush is noticed. That can be observed in the figures below, which represent the normalized volume of runoff (1 means 100% of water, see section 1.2) of an event versus the total mass of Zn reported in the same event. Total Zn, which includes suspended matter, has a certain

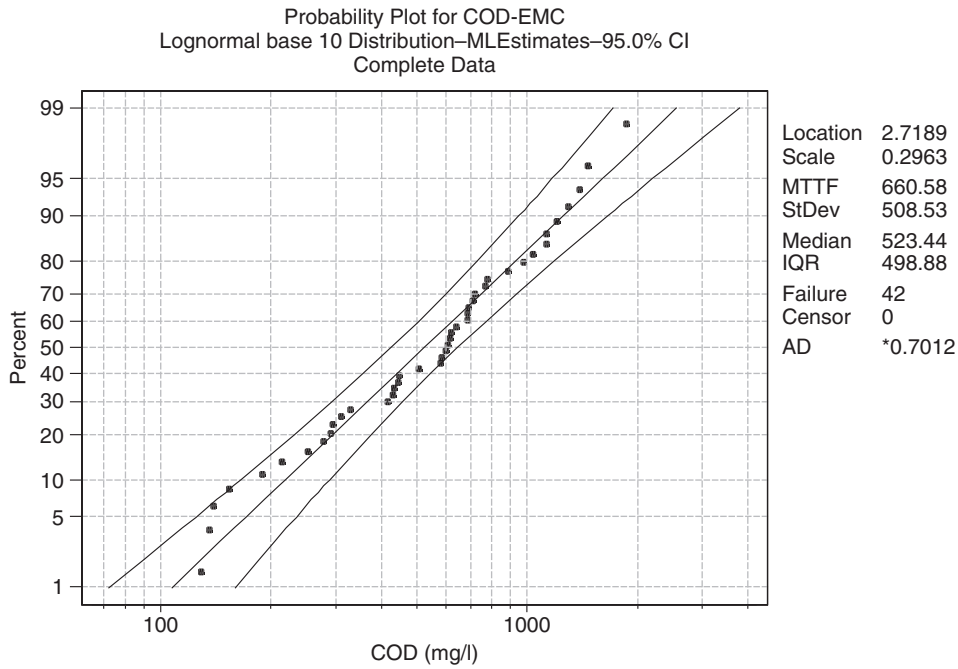


Figure 5. COD-EMC distribution. Integrated curve including five catchments in Spain.

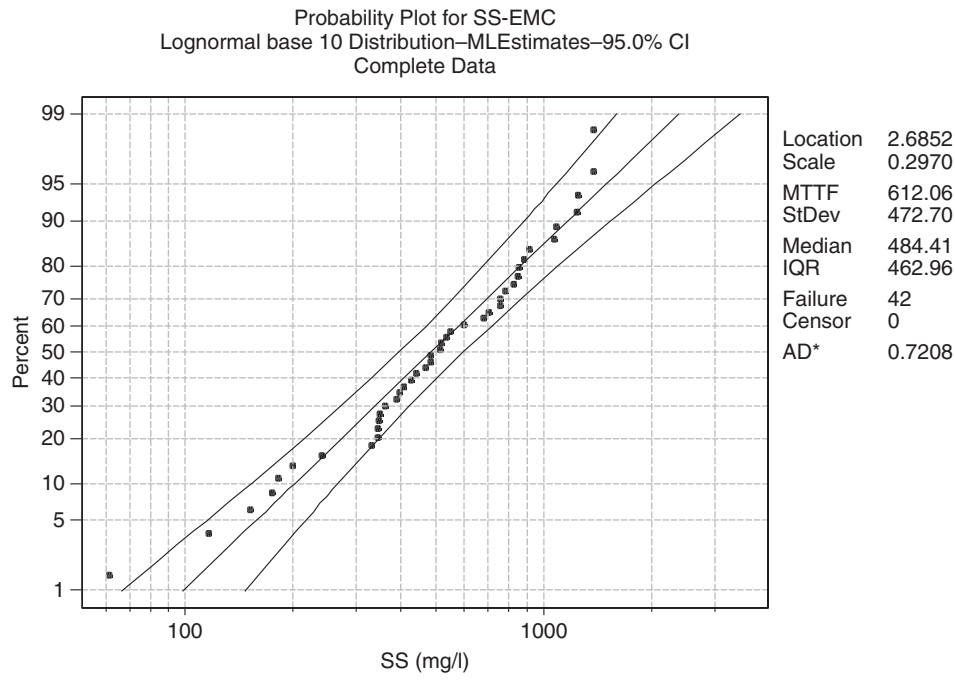


Figure 6. SS-EMC distribution. Integrated curve including five catchments in Spain.

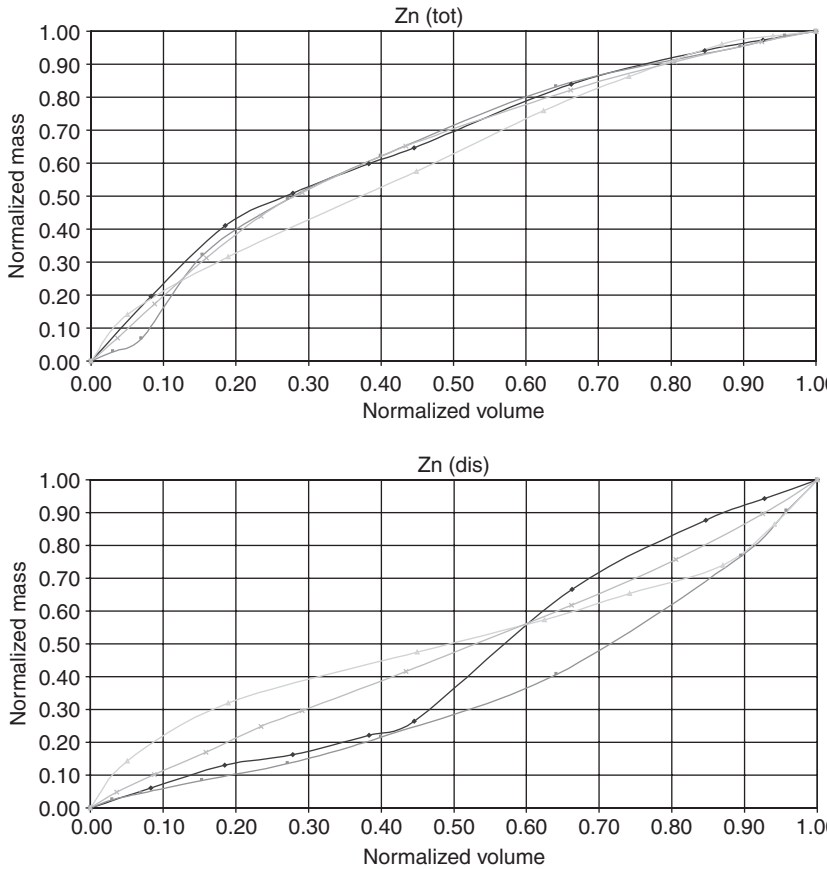


Figure 7. Differences between total (up) and dissolved (down) Zn, in Santiago-Este catchment. Curves represent different rainfall events.

degree of first flush (40% OF Zn is measured into 20% of water). Dissolved Zn does not have this behaviour at all.

All metals have more or less the same origin (traffic, industrial activity, …) and their concentrations are in the same proportion. In figure 8, ten rainfall events in Santiago have been measured, and the specific loads (g/m2) have been analyzed. It can be seen that the amounts of some metals are quite bigger than others (Fe, Al), but there is a general trend. The more solids are removed, the more metals are reported, and in the same proportion. The slopes of all graphs in figure 8 are the same.

3.4 *Solids accumulation and removal*

As it has been already said, solid production in the catchment and solid removal, both in the catchment and in the sewer network, are important phenomena. Production and removal equations are very weak, as they are quite simple, and the phenomena include some questions, as cohesion, which are not well represented until now. So, when trying to calibrate these processes no good agreements are obtained. Figure 9 represents a calibration between real data (Suspended solids) and numerical data, obtained from SWMM model. Concentration peaks, which are not well modeled, correspond with sewer removal events, which are not well modeled in SWMM. MOUSE can do it a bit better, but no good approach is obtained anyway.

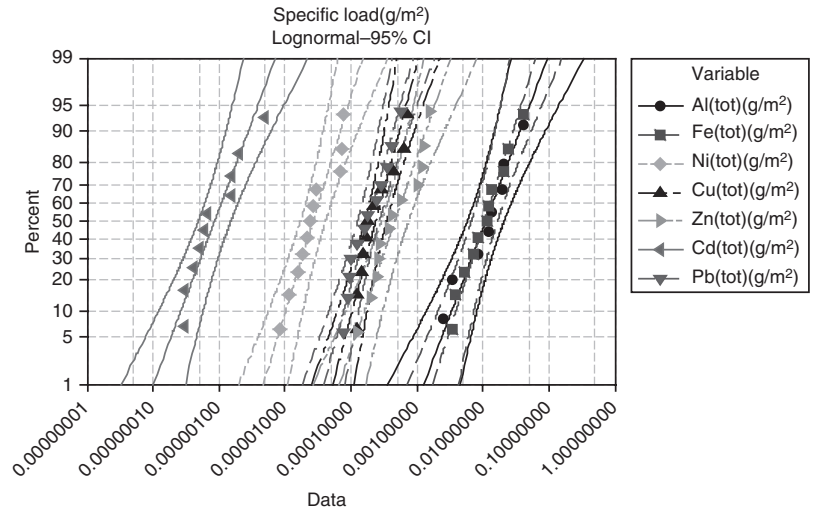


Figure 8. Specific loads of metals in Santiago-Este catchment.

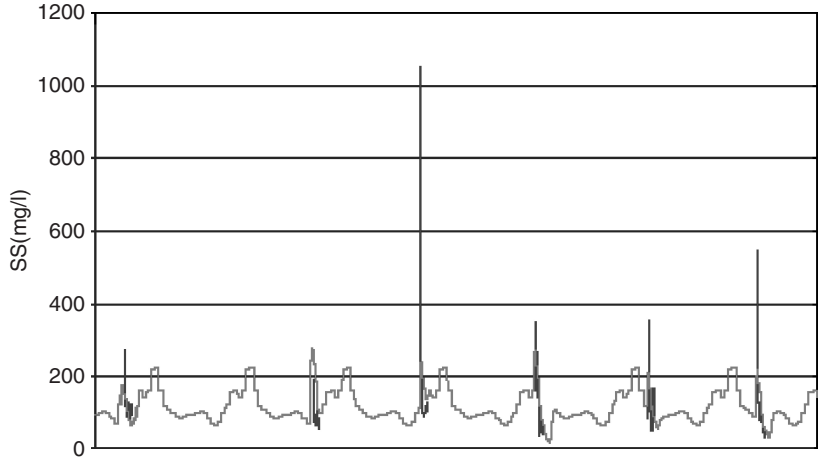


Figure 9. Suspended solids calibration. A poor agreement is achieved.

Besides, figure 9 represents a calibration, with no extrapolation. If parameters obtained in this calibration are applied to other sewer system, the results can be even worse. Figure 10 represents a calibration, including water and suspended solids, and a validation, with the same parameters, in the same catchment, and other rainfall event. As it can be seen, water agreement is more or less good, but solids agreement is really poor.

Some work has to be done to achieve a good agreement between real data and numerical data. Physical processes are more or less known, but there are not good equation packages to model them.

4 CONCLUSIONS

This paper covers a wide range of problems, which can be grouped as urban drainage hydraulics, and urban water pollution. Equations covering the different topics (hydrology, sewer hydraulics,

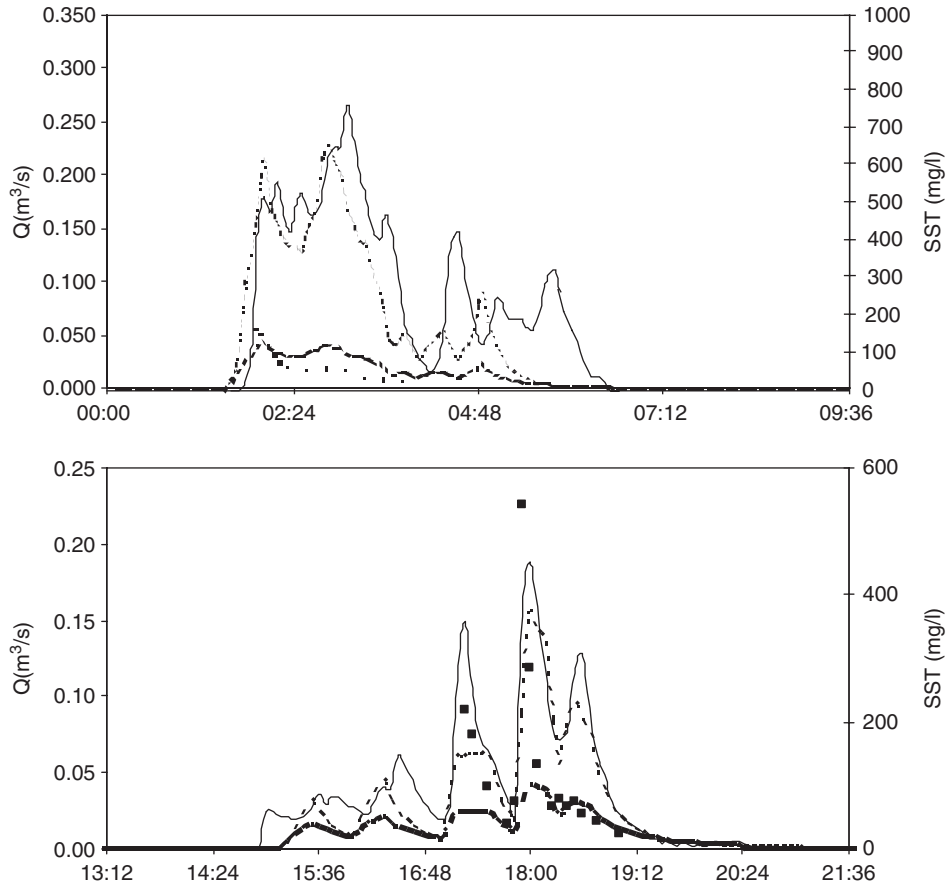


Figure 10. Calibration (up) and validation (down) of water discharge (higher values) and suspended solids (lower values). Solid lines are numerical data. Dots are real data.

pollution transport) do not fit real data in most of cases. So, an important effort has to be done, in order to propose new equation packages, and in order to have more real data to calibrate them.

Sewer hydraulics is probably the more accurate topic. Urban hydrology is full of uncertainties, as models do not represent reality but a simplified geometry, and parameters do not have a clear sense.

Pollution dynamics is very difficult to analyze. Dissolved and suspended fractions have different behaviours, and removal of pollutants from the catchment or the sewer system, which are main topics, have no reliable equations. So, no good agreements can be expected.

In conclusion, plenty of work should be done in this field.

REFERENCES

ATV-128, 1992. Standards for the dimensioning and Design of Stormwater Structures in Combined Sewers.

Bertrand-Krajewski, J.L. Chebbo, G. and Saget, A., 1998. Distribution of pollutant mass vs. volume in stormwater discharges and the first flush phenomenon; Water Research, 32(8): 2341–2350

Butler, D. and Davies, J.W, 2000. Urban Drainage, E&FN SPON, London, 489 pp.

Ellis, J.B., 1989. "Urban Discharges and Receiving Water Quality Impacts" (Adv. Wat. Poll. Control N° 7), Pergamon Press, Oxford.

Ellis, J.B., 1991. Measures for control and treatment of urban runoff quality, Rep.DT3PL /FV/JB, Agence de l´eau Seine-Normandie. Paris.

Gupta, K. and Saul, A.J., 1996. Suspended solids in combined sewer flows. Wat. Sci. Tech., 9: 93–99.

Lager J.A., et al., 1977. "Urban stormwater management and technology: Update and users guide". EPA-600/8-77-014, Cincinnati, OH.

Lijklema, L., and Tyson, J.M. Lesouef, A., 1993 Interactions between sewers, treatment plants and receiving waters in urban areas: a summary of the INTERURBA ´92 WORKSHOP conclusions; Wat. Sci. Tech., 27(12): 1–29.

Metcalf and Eddy., 1991. "Wastewater Engineering. Treatment, Disposal. Reuse"; 3rd Edition; McGraw- Hill International Editions, Civil Engineering Series. ISBN 0-07-100824-1.

MINITAB Statistical Software Release 13, 2000. MINITAB Inc.

NWRW, 1991. Final Report of the 1982–1989 NWRW (National Working Party on Sewerage and Water Quality). Foundation for Applied Waste Water Research, the Dutch Ministry of Housing, Physical Planning and environment, 63 pp.

Saget, A., Chebbo, G. and Bertrand-Krajewski, J.L., 1996. The first flush in sewer systems.; Wat. Sci. Tech., 9: 101–108.

Suárez, J. and Puertas, J., 2005, "Determination of COD, BOD and suspended solids loads during combined sewer overflow (CSO) events in some combined catchments in Spain", Ecological Engineering, 24, pp 201–219.

US-EPA, 1983. Results of the Nationwide Urban Runoff Program. Vol. I, Final Report, NTIS PB84 185552, Water Planning Division, Washington, DC.

Valiron, F. and Tabuchi, J.P., 1992. Maitrise de la pollution urbaine par temps de pluie. État de l´art.; AGH-TM; Tec-doc; París, ISBN 2-85206-863-X.

Water Research Centre, 1987. Sewerage detention tanks – a design guide, WRc, Swindon.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Control of irrigation canals: Why and how?

P.O. Malaterre
Cemagref, UMR G-EAU, Montpellier, France

ABSTRACT: The objective of this survey paper is to review, structure and analyze the main concepts involved in the control of irrigation canals, and to illustrate them on a canal benchmark. After a description of the irrigation and water management contexts, we present the main lines of a classification of the canal control systems. Then, we present a simple benchmark example of an irrigation canal. We compare results obtained with manual operation and some well known control algorithms (PID, LQG and ℓ_1). This illustrates the type of advantages that can be obtained with automatic control and some of the tradeoffs to be addressed in the selection of a solution. This also illustrates the needs for theory, methodologies, design and simulation tools to select and tune these control algorithms. Although a large part of these developments are still at the research stage, more and more of these techniques have successful field implementations.

1 CONTEXT

1.1 *Irrigation*

An irrigation canal is an open-channel hydraulic system whose main objective is to convey water from a source (dam, river) to different users. Cross structures are operated in order to control the water levels, discharges and/or volumes along this canal (Figure 1). These structures are composed of devices that are mainly undershot gates in most countries or overshot weirs in flat regions such as Australia and Netherland where head losses have to be minimized. The users are mostly agricultural lands, accounting for more than 80% of the freshwater withdrawals, in world average. Some users are also, in some cases, industrial or domestic ones. The spread of urban areas over agricultural lands increases the number of situations where a traditional irrigation canal gets several management objectives: irrigation of crops, satisfaction of domestic uses (gardens, swimming pools), sometimes drainage of urban or rural areas in case of heavy rains, etc. Some environmental uses have also to be fulfill for some canals (groundwater recharge, lake inflow-outflow management, fish preservation in upstream or downstream connected rivers) or recreational activities have to be maintained in the

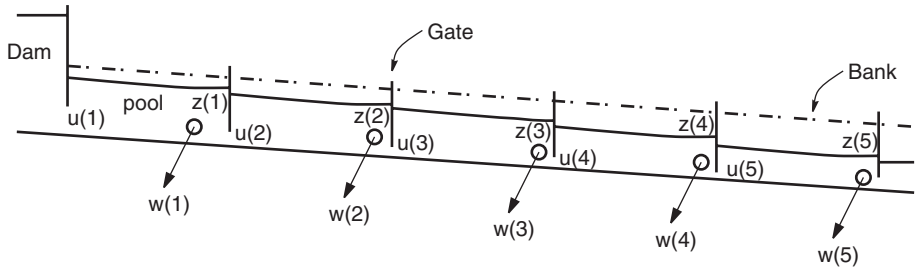


Figure 1. Canal "type 1" from Cemagref Bench Marks.

canal or more often in rivers linked to the canal (canoeing, bathing). Industrial and domestic users have reduced volumetric needs but of strategic importance and have often intermediate storage capacity in order to be able to face distribution constraints or failures (in quantity or in quality).

Such systems can be very large (several hundreds of kilometers), characterized by time delays and non-linear dynamics, strong unknown perturbations, and interactions among subsystems. Varying operational objectives are assigned to their managers. The main general one is to provide water to the different users at the right moment and in the right quantity, reducing losses as much as possible and to guarantee the safety of the infrastructure. In particular, a major concern is to prevent the canals from overtopping, but also from having water levels inside the pools below the supply depths of the gravity offtakes. For unlined canals, the fluctuations of the water levels have to be slow enough in order to limit bank erosion. The cross-structures used as actuators have maximum allowed gate openings, maximum number of operations per hour and minimum gate movements due to structural or management constraints.

1.2 *Modernization of irrigation canals*

The "control" of irrigation canals is also called "hydraulic regulation" in this paper, but we can also understand sometimes the word "control" or "controller" in a narrower meaning, limited to the algorithmic part of the regulation (e.g.: PID, LQG controller). We will see in the following chapters that the control algorithm is an important part, but not the only one (and probably not the most important one) and that the choice of the hydraulic regulation of an entire canal is a selection of several options (variables, logic of control, algorithm, etc.).

Despite many research and development efforts and advances in automatic control applied to open-water hydraulic systems for more than 50 years now, most of the irrigation canals in the world are still controlled following traditional manual operation procedures. Even in France, where the efforts have been particularly early and important, a survey made in 1997 on 180 irrigation canals in the south of France (with design discharges ranging from 0.1 m^3/s to 75 m^3/s), showed that 170 of them were still manually controlled and the remaining 10 ones were indeed partially or fully automated. These 10 canals proved to be the largest ones, which is not surprising since the investments are made in priority on large strategic systems. All together they were conveying almost 40% of the total cumulated discharges (118 m^3/s out of 300 m^3/s).

An important effort is put in the modernization of irrigation canals where the pressure on water conservation is very high such as in Australia (Mareels, Weyer, Ooi, Cantoni, Li, and Nair 2005), Western USA (Clemmens, Strand, and Feuer 2002), Mexico (Lopez-Cantens, Ruiz-Carmona, Arteaga-Tovar, and Ostos-Santos 2007) and many other countries (Pakistan, India, Jordan, Morocco, etc.). Despite recent applications all over the world (new construction or modernization of existing canals), the percentage of automated irrigation canals in the world is still below 5%.

There are several reasons for that. The main one is that it is difficult to change way things are used to be done, sometimes for several hundred of years. Managers, often for good reasons, are reluctant to change the way their system is operated since they cannot afford disruption of delivery during construction work and are afraid of not being able to solve a problem in a modified system using technologies they are not familiar with (sensors, actuators, SCADA systems, RTUs, computers). This introduction of new modern technologies is much easier for new irrigation canals: Narmada canal in India (www.rajirrigation.gov.in/4narmada.htm), Toshka canal in Egypt (www.water-technology.net/projects/mubarak), Pehur canal in Pakistan (www.adrianlaycock.com/irrigate.htm). For these new projects the cost of automation is not a real issue since it represents less than 5% of the total construction project.

New construction projects are not so frequent in Europe, but they are still numerous in developing countries where an efficient irrigation is the only way to increase the agricultural production with a limited water resource. At the moment, the irrigated agricultural lands account for only 18% of the total cultivated areas but they contribute for more than 40% of the total agricultural production. The FAO (Food and Agriculture Organization of the United Nations) indicates that for the year 2030

the agricultural production will have to be increased by +80% to fulfil the food demand, but that will have to be done without the possibility to increase the water withdrawals by more than +12%. The drastic reduction of water spillage required to reach this objective is technically possible since studies indicate that in the Mediterranean region, by the year 2025, the water savings can be of 65% for irrigation, 22% for industrial uses and 13% for domestic uses (Institut Méditerranéen de l'Eau 2007). These savings can be obtained through the implementation of several measures allowing to better manage the water demand at several institutional and technical levels. The survey made in South of France in 1997 confirmed the figures found in other sources (ASCE 1993), (Mareels, Weyer, Ooi, Cantoni, Li, and Nair 2005) showing that the water losses at the distribution level were around 50% in average and could be reduced to less than 10% by the modernization of the canal realtime operation (including automation).

In France, and all over the world, more and more irrigation canals are following modernization processes in order to face the new challenges imposed by the constraints on the reduced water resources and the increasing water demands. These modernization processes usually include an analysis of the existing system operation and maintenance, a diagnosis of the problems and challenges to be faced, a gathering of data (geometry of the canals, water demand, water level or discharge measurements). Then, a modeling of the system is often made in order, latter on, to be able to test the different projected alternatives for hydraulic devices types and locations and for automatic control procedures. The model G_n evoked above (or a family of models for different conditions: weed growth, sediment deposition, etc.) is also often used to confirm or complete the diagnosis of the present situation. A control solution (or several alternatives) is then set up using the experience of the consultant company, the objectives defined with the canal managers and the previous diagnosis. This solution is mainly defined by the type of control devices (gates, weirs, etc), their locations, and the algorithm used for their operations. In some rare cases it can still be manual operation procedures (Malaterre, Rey, and Baume 1993), (Godaliyadda, Renault, Hemakumara, and Makin 1999). In more and more cases it is automatic control algorithms. The most difficult task is then to select and tune these control algorithms so that they fulfil the control objectives. Once they have been tuned, they have to be tested on the simulation model prior to a field implementation in order to verify that the behavior is satisfactory. Once everybody is happy with the proposed solution, the construction and implementation process can start. Of course the previous steps can be iterated.

2 CLASSIFICATION

Several survey papers have been written on the subject: (Malaterre, Rogers, and Schuurmans 1998), (Malaterre, Baume 1998), (Lopez-Cantens, Ruiz-Carmona, Arteaga-Tovar, and Ostos-Santos 2007), (Mareels, Weyer, Ooi, Cantoni, Li and Nair 2005). We summarize and expand the main concepts defined in these papers. Only main and recent references are given in this paper.

The hydraulic regulation (realtime operation) of an entire irrigation canal is processed by one or several independent or interacting "canal control systems" as defined hereafter.

A canal control system is an elementary system (algorithm software + hardware components) in charge of operating one or several canal cross structures, based on information from the canal system. This information may include measured variables, operating conditions (e.g. predicted withdrawals) and objectives (e.g. hydraulic targets). Boundaries of the control system are output of the sensors placed on the canal system, and input to the actuators controlling the cross structures. The following text presents definitions and a classification of canal control algorithms developed or used in the world. Hardware aspects are not addressed in this paper and are presented in other papers (Rogers and Goussard 1998).

The difficulty presented by the classification of canal control system algorithms is due to the different possible ways of characterizing them (e.g. controlled variables, configuration of field implementation, communication management, design technique, data required, location of devices along the canal). Among all possible criteria, one wants to retain those, in minimum number, that

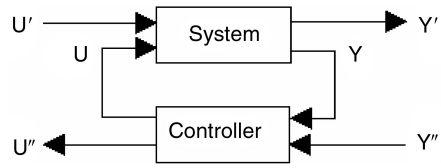


Figure 2. Considered variables.

allow characterization of the hydraulic behavior, the performance, the data requirements and the constraints of the various canal control systems. We choose to retain the following four essential criteria: considered variables, logic of control, modeling technique and controller design technique. Sub-criteria will be defined to refine them.

These different terms are defined, discussed and illustrated in the following sections. The order of the four criteria is not linked to any priority of interest. Depending on technical background and areas of responsibility, different people may have different levels of interest in the four criteria. Civil and hydraulic engineers may be more concerned with the considered variables and modeling technics, while control engineers are more concerned with the logic of control and controller design technique.

2.1 *Considered variables*

The location of the considered variable is given in reference to a pool (a pool is a portion of a canal, situated between two control devices) and not to a structure. This avoids confusion in the case of a multivariable control algorithm, where a variable can be controlled either by an upstream structure, by a downstream structure or by both at the same time. Two types of variables are always considered in control algorithms: the measured variables Y and the control action variables U (Figure 2). In some cases the controlled variables Y' are a subset, or are calculated from a larger set of measured variables. When the controller incorporates a feedforward component, or when the design technic is explicitly taking into account uncertainties such as in the case of robust control, information on some perturbations U' can also be used. When the entire canal is controlled by several decentralized controllers, each of them can use information calculated by all or some of the others. This information is defined as the input control connection Y'' or output control connection U''.

2.1.1 *Controlled variables Y'*

Controlled variables Y' are target variables controlled by the control algorithm. They are important variables that have to be constrained by the controller in order to have a specific behavior, even in the presence of perturbations U'. In fact the controller can be seen as a way to change the natural behavior of the transfer $\Phi : U' \rightarrow Y'$ into a more satisfactory one. Control engineers define this modified transfer as the lower linear fractional transformation of the extended model of the system G by the controller K and is noted $F_l(G, K)$. This modified behavior can be seen in many different ways. It can be a behavior in the time domain or in the frequency domain. It can be equality or inequality constraints. It can be the minimization of some of its norm. In the following chapters we will illustrate some of them. At the moment there is no clear paper on the terms of references that could be imposed on such controllers. This also explains why so many authors are proposing controllers, claiming it is a good one. When the objectives are loose, all stable controllers can be considered as fulfilling them. Progresses in the clarity of the terms of references will certainly be done in the coming years. Before listing the classical controlled variables considered on irrigation canals we can tell that they are not necessarily directly measurable. It is frequent and it is easier if they do, but they can also be functions of some other measurable variables.

Examples of controlled variables are water level at the upstream end of a pool (y_{up}), water level at the downstream end of a pool (y_{dn}), flow rate (Q) at a structure or at a special location, volume (V)

of water in a pool, and weighted water level (e.g. $ay_{up} + by_{dn}$). Control theory speaks of "tracking" when their target values are time dependent.

2.1.1.1 Discharges

The needs of irrigation canal users are defined mainly in terms of discharge. For example, agricultural needs are expressed in terms of given discharges delivered to a plot, to a secondary canal, or to a pumping station; environmental needs as tail end discharge, or minimal discharge; urban needs as discharges delivered to a house or to a city water filtration plant; and industrial needs as discharges delivered to a factory. Natural or artificial storage reservoirs are sometimes available (e.g. soil maximum water storage, lateral or on-line reservoir, basin of a water filtration plant, volume stored into the canal pools). Users' needs can then be defined in a more flexible way, in terms of volume distributed over a time period. In this case, the controlled variable is no longer a given value of discharge, but a volume, which is the integral of a discharge over a given time period. Discharge fluctuations are then authorized, and neutralized by the capacity of the storage reservoirs. However, these reservoirs are expensive and of limited sizes, and constraints of distribution never suppress needs expressed in terms of discharge.

Consequently, all free surface hydraulic systems must be managed, directly or indirectly, in order to satisfy users' demands in discharge. Considering the nature of the physical phenomenon at stake (gravity open channel flow from upstream to downstream), these demands in discharge can be satisfied only from the source situated at the upstream end of the system, by draining the upstream reservoirs.

2.1.1.2 Water levels

Contrary to discharges, water levels can be easily measured in free surface canals and rivers. Furthermore, constraints of feeding gravity turnouts, stability of canal banks, efforts to reduce weed growth, constitution of intermediate water storage volumes, and risks of overflow are expressed in terms of water levels. Controlled water levels "y" can be upstream (y_{up}), downstream (y_{dn}), or intermediate inside the pool (y_{in}). The location of controlled variables in a pool is indicative of the hydraulic behavior of the system (e.g. available storage volume) and structural constraints (e.g. bank slopes). Controlled water level target values are usually equal at null and maximum discharges. This is not always the case (e.g. GEC Alsthom Gates). The corresponding water level difference is called "decrement". Operational characteristics are very different depending on the location of the controlled water levels "y".

One of the advantages of controlling upstream (of the pool) water levels is that a storage volume V is available between the null discharge volume (corresponding to a horizontal water profile) and the maximum discharge volume (corresponding usually to a uniform water profile parallel to the bed). It allows for rapid response to unforeseen demands of turnouts or downstream reaches and for storing water in case of a consumption reduction. But canal banks have to be horizontal, which can be expensive (or even impossible) if the zone is not very flat.

When downstream (of the pool) water levels are controlled, canal banks can follow the field natural slope, which reduces construction costs. But, no storage volume is available between the null discharge volume and the maximum discharge volume. In fact, pool volumes change in the opposite direction from the one that would help satisfying downstream demand changes. Therefore, the system cannot respond rapidly to unforeseen demands. The excess water cannot be stored locally and is "lost" in the downstream pools. Supply flow changes (generated by other control systems further upstream) must overcompensate in order to match downstream demand changes and to establish new pool volumes.

Controlling a particular intermediate water level, close to the middle of the pool, is equivalent to controlling the volume stored in the pool. This water level can be measured directly (no example has been found in the literature), or can be obtained as a linear combination of an upstream and a downstream water level (e.g. BIVAL). Controlling an intermediate water level is a compromise between the two previous options, in terms of construction cost and availability of storage volume V.

Indeed, banks have to be horizontal only downstream from the controlled intermediate water level. But, one or several distant water levels have to be measured, which increases telemetry requirements.

2.1.1.3 Volumes

When volumes are controlled, controllers are less sensitive to perturbations, but response times are increased. These methods are applicable to irrigation canals with important storage volumes, and equipped with turnouts whose feeding is not dependent on water levels in the main canal (e.g. pumping stations).

2.1.2 *Measured variables Y*

Measured variables Y, also called inputs of the control algorithm, are the variables measured on the canal system and provided to the controller. Usually the controlled variables are a subset of the measured variables. The extra variables can be used, for example, in an observer to reconstruct state variables used by the controller. Examples of measured variables are water level at the upstream end of a pool (y_{up}), water level at the downstream end of a pool (y_{dn}), water level at an intermediate point of a pool (y_{in}), flow rate at a structure (Q), and setting of a structure (G).

Measured variables on irrigation canals are generally water levels. In some cases, measured variables can be discharges. Discharge can be measured with flow meters (based in general on the measurement of one or several flow velocities, with a propeller, an ultrasonic, electromagnetic or acoustic doppler device), measurement flumes using water level measurement to compute $Q(z)$, through a cross structure equation $Q(z_1, z_2, G)$, or a local control section rating curve $Q(z)$ with a sufficient precision. When such an equation exists, it is assumed that a discharge Q is really measured, whatever the process used to obtain it, even if it is calculated from one or several water level measurements. Finally, measured variables can be volumes, evaluated by measuring several water levels along the canal, or by evaluating input – output discharge balance.

The quality of the measurement process is important since the controller will use these measurements and the control devices will be manipulated consequently. If a sensor has a default, it can be crucial to detect this and take corrective action. If the number of sensors is limited, a human supervisor can be in charge of this detection. But, when the number of sensors becomes very high, automatic fault detection procedures can be very useful in terms of detection quality and rapidity of action. Recently, some works have been done on this subject (Canivet 2002), (Deltour, Sanfilippo, Canivet, and Sau 2005), (Koenig, Bedjaoui, and Litrico 2005), (Bedjaoui, Litrico, Koening, and Malaterre 2006), (Bedjaoui 2007), (Choy and Weyer 2005), (Choy and Weyer 2006).

The measured variables as defined here, being the input of the control algorithm, can also be estimated, instead of being effectively "measured". This can be done in several occasions such as for example in the case the control algorithm has a feedforward component (or is fully feedforward) using prediction of future perturbations, or when a default has occurred on some measured variables and therefore when a reconstruction of this variable is used instead.

2.1.3 *Input control connection Y"*

In case the global hydraulic regulation of an irrigation canal is done using several control systems, these control system can either be completely decentralized or can communicate through the exchange of some data. In most cases ones prefer completely decentralized controllers for simplicity reasons (tuning and communication). But, some others use interconnected partially decentralized controllers with good performance improvements. The dynamic regulation of SCP is using such technic adding a portion (e.g.: 80%, (Deltour 1992)) of the discharge control action at gate n to the next upstream discharge control action at gate $n - -1$. Studies using non-linear optimization (Malaterre and Baume 1999) proved that such connection improves a lot the performance of simple PI controllers. The rational for these interconnections is to improve the decoupling of the decentralized controllers (Malaterre 1994), (Schuurmans 1992). Studies showed that, although a fully centralized controller performs better than a series of local controllers, reasonably good

performance can be obtained for partially decentralized PI controllers that pass information to one additional check structure upstream and downstream (Clemmens and Schuurmans 2004).

2.1.4 *Perturbation variables U'*

The perturbation variables U' are the non-controllable input influencing the system to be controlled. They are for example known or unknown offtake discharges, rain inflows, seepage and evaporation, etc. They can also be anything occurring on the real system but that is not modeled by the selected model G_n such as failures, breakdowns, un-modeled dynamics, parametric uncertainties, etc. Even if these perturbations are non-controllable, any information on them (in the frequency or in the time domain, deterministic or stochastic, advanced model or just bounds on some of their norms) can be used to improve a feedback controller or to design an openloop controller.

The robust controller design is based on a modeling of some of these perturbations, on an evaluation of the norm of the transfer from these perturbations to some controlled variables, and a design of the controller providing specific performance even in the presence of these perturbations.

2.1.5 *Control action variables U*

Control action variables (U), also called outputs of the control algorithm, are issued from the control algorithm and supplied to the cross structures' actuators in order to move the controlled variables towards their established target values. They are usually either gate or weir positions (G) or flow rates (Q). In this latter case, another algorithm transforms the flow rate into a gate position. This latter algorithm is important from hydraulic and control points of view, and is considered as a separate control algorithm, from those presented in this classification. This conversion can be done through the inversion of the device static equation $Q(z_1, z_2, G)$, or by a local dynamic controller (e.g. PID controller). The inversion of the device static equation can be done using several alternatives depending on the way the evolution of the upstream z_1 and downstream z_2 water levels are taken into account (Malaterre and Baume 1999), (Litrico, Malaterre, Baume, and Ribot-Bruno 2007). Control action variables G or Q can be considered as absolute values, relative values (relative to a reference state) or as incremental values (to be added to the value of the previous time step).

Considering the gate position G has the advantage of taking into account the complex dynamics linking this position with the local discharge and upstream and downstream water levels (granted the model used is modeling this correctly). These dynamics are important and it can be hazardous not to take them into account. Considering discharge Q as the control action variable allows for decoupling of the different subsystems. This is interesting when monovariable controllers are used in series (e.g. Dynamic Regulation, PIR).

Some additional algorithms can handle specific features such as anti-windup technics for actuator saturation (Zaccarian, Li, Weyer, Cantoni, and Teel 2007).

2.1.6 *Output control connection U"*

The output control connection variables U" are the variables calculated by this control system that will be provided to another control system through its input connection variable Y" as defined above.

2.1.7 *Impact of the choice of the variables*

The choice of the above variables is extremely important. Most of the papers dealing with automatic control applied to irrigation canals are dealing mainly with the control technic (PID, LQG, Predictive control, $\mathcal{H}_\infty$, etc), assuming it is the most important choice to be made. Few papers are dealing with the choice of the variables and their impact, although the obtained performances can change by a factor 2 to 10 (Malaterre and Baume 1999). One of the reasons for that is the decoupling effect introduced by the selection of the discharge as the control action instead of the gate position. Another reason is that the model G often required for the design and tuning of the control algorithm is depending a lot on this choice. For example, if the control action variable U is the gate opening at cross structure n and if the controlled variable Y' is the water level downstream of the pool n, then the model G linking these two variables can be something like a second order delayed transfer function. Now, if the control action variable U is the discharge at cross structure n and

if the controlled variable Y' is the volume of the pool n, then the model G linking these two variables can be something like a pure integrator. In addition to the structure of the model that can be very different, in the first case, the model parameters depend a lot on the flow, water levels and gate position, whereas in the second case the model is always the same and linear if the canal is rectangular. These differences have an impact on the type of algorithms that can be designed and on their nominal and robust performances.

2.2 *Logics of control*

The logic of control refers to the "type" and "direction" of the links between controlled variables and control action variables above defined.

2.2.1 *Type*

The control algorithm uses either feedback control (FB, also called closed-loop control), feedforward control (FF, also called open-loop control) or a combination (FB + FF). These concepts are further detailed in section 2.6.

2.2.2 *Direction*

A structure can be operated to control a variable located further downstream, which is called downstream control. All variables (discharge, level or volume) can be controlled with downstream control. A structure can also be operated to control a variable located further upstream, which is called upstream control. Only levels or volumes can be controlled with upstream control, when flow conditions are subcritical and under the limitations of the backwater effects. The discharge can be modified in transient phases, but it is not controllable without drying or overtopping the upstream pool. This limitation explains why downstream control is a very interesting method compared to upstream control. Upstream control presents the advantage of simplicity.

Some control methods combine upstream and downstream control logics. They are sometimes called mixed controls. Since they benefit from the main advantages of downstream control methods, they are often called, to simplify, downstream control methods. Examples and references are given in (Malaterre, Rogers, and Schuurmans 1998).

2.3 *Modeling techniques*

When defining the different variables involved in the problem addressed in this paper, we have seen that the objective of the design of a control system is to modify the transfer between perturbation variables U' and controlled variables Y'. This modification is based on the manipulation of control action variables U calculated by the control algorithm K from measured variables Y. For a PID controller, this design and tuning can be done directly on the process to control (Litrico and Malaterre 2007), (Litrico, Malaterre, Baume, Vion, and Ribot-Bruno 2007) using the ATV and Ziegler-Nichols methods. The AMIL, AVIS, AVIO and Mixte hydrodynamic gates are also tuned directly in situ following a precise tuning procedure. But, to my knowledge, there are the only methods allowing this, and all other methods require a model of the process to control. Let's call this model G_n.

$$\begin{pmatrix} Y' \\ Y \end{pmatrix} = G_n \begin{pmatrix} U' \\ U \end{pmatrix} \tag{1}$$

Any model is always a simplification of the real system. There is a large range of modeling approaches, from quite complete ones to very simplified ones. The link between the type of model and its possible uses is very tight. The quite complete models can be used for simulation purposes, and will be described in the next section "Simulation models". Unfortunately, these models are too complex to be used for the design of controllers. Simplified models described in the next section "Models for control" are required. Hopefully, as described in the next sections, the robustness of the control algorithm can be a posteriori evaluated or a priori imposed, and we can therefore prove that the automatic control algorithms will have some precise stability and performance characteristics

even though the nominal model used is not a "perfect" one. The only conditions for that is to be able to give a bound for all the possible models G around the selected one(s) G_n.

For example, $G = G_n + \Delta$ with $|\Delta| < \alpha$ for an additive uncertainty or $G = G_n(1 + \Delta)$ with $|\Delta| < \alpha$ for a multiplicative uncertainty, where G is the "real" system, G_n is the selected model(s), Δ is the uncertainty including all sorts of features (unknown parameters, neglected high frequency dynamics, etc.) (Skogestad and Postlethwaite 1998).

The evaluation of this uncertainty Δ is possible using different technics developed by automatic control engineers (e.g.: Linear Fractional Transformation form obtained from parametric uncertainty propagation). Due to the number of parameters implied in these models, this task is quite difficult, time consuming, and no hydraulic software includes this calculation for the moment. But this will be certainly the case in the future and it will be a very interesting feature allowing to give model outputs with an indication on the uncertainty and to give a bound on the model uncertainty that can be used for robust control design.

2.3.1 *Simulation models*

At the beginning of the last century, when the computers were not so present on our desks, the models used were scale physical models. For example, the hydrodynamic AMIL, AVIS, AVIO and Mixte gates were designed, tuned and tested on such physical models by the Neyrpic company (called latter on Neyrtec and Gec-Alsthom) in Grenoble (France) and in Alger (Algeria) in the years 1930–1990.

Nowadays, for irrigation canals, the most precise and complete hydraulic models are 1D hydrodynamic models based on Saint-Venant's equations (Barré de Saint-Venant 1871), (Cunge, Holly, and Verwey 1980) completed with internal and external boundary conditions (Malaterre and Baume 1997), (Malaterre and Baume 2007). This type of model is usually sufficient for the objectives of testing the different projected alternatives for hydraulic devices and automatic control procedures, and eventually for confirming or completing the diagnosis of the present situation for an existing canal to be modernized. A list of numerical models based on such approach and adapted to irrigation canals was set up by the ICID commission (Goussard 2000). Among the listed software, only a very limited number is proposing a control library allowing to test feedback controllers at some devices. The SIC software is certainly the most adapted to this use (Malaterre and Baume 1997), (Baume, Malaterre, Belaud, and Le Guennec 2005), (Litrico, Malaterre, Baume, Vion, and Ribot-Bruno 2007), (Malaterre and Chateau 2007). Before being used on an existing canal, these models have to be calibrated and validated on measured field data, in steady flow and unsteady flow conditions. The unsteady flow validation is an important issue, which is seldom carried out correctly for lack of data reasons. But this should be done since the dynamical features such as time lags and signal attenuations have an impact on the behavior of the control algorithms.

2.3.2 *Models for control*

Another challenge is the selection of a model to be used for the design of the automatic control algorithms and/or procedures. As we will see in the following sections, the control technics allowing using a complete high order (using many state variables) and non-linear models are not numerous, difficult to design, with limited performance guaranties. Simplified models, with less parameters and lower orders, can be used with much more design technics developed by control engineers and used in many other industrial applications.

These models can be written in the time domain or in the frequency domain, in continuous or in discrete form, with mathematical procedures to shift from one form to the others (Skogestad and Postlethwaite 1998). Different models are adapted to different design technics. A classical state space discrete form can be:

$$\begin{cases} x(k+1) = Ax(k) + B_1 u'(k) + B_2 u(k) \\ y'(k) = C_1 x(k) + D_{11} u'(k) + D_{12} u(k) \\ y(k) = C_2 x(k) + D_{21} u'(k) + D_{22} u(k) \end{cases} \tag{2}$$

where x is the state vector, u' the perturbation vector, u the control action vector, y' the controlled variable vector, y the measured variable vector, k the time instant and A, B, C and D matrices of

appropriate dimensions. For the specially interesting family of LTI (Linear Time Invariant) models, these matrices are constant. For a LTV (Linear Time Varying) model they can be time dependent. For a non-linear model they can be state and time dependent.

Complete non-linear model: The Saint-Venant's equations and the complementary internal and external boundary conditions can directly be used to design a controller as we will see in the next chapter (De Halleux, Prieur, Coron, D'Andrea-Novel, and Bastin 2003). But this is possible only on limited systems making simplification assumptions (e.g.: homogeneous rectangular section, no slope, no friction).

Complete linearized model: A linearized version of these equations can also be used (Xu and Sallet 1999), (Bounit, Hammouri, and Sau 1997). This approach and the former non-linear one are powerful since they use the most complete version of the system model. But, the mathematical techniques required to design the controller are more complex, and can be used, for the moment, only on a very simple canal reach.

Infinite order linear transfer function: Since the above described models are not very easy to use when designing a controller, some simplified models have been derived from the previous ones through some simplifications. A linearization around uniform flow, Laplace transform and integration of the above Saint-Venant's equations lead to a linear infinite order model (Baume, Sau, and Malaterre 1998), (Baume and Sau 1997). This model has the advantage of keeping the distributed parameter system characteristics and therefore the infinite state space dimension. This approach has been extended to the non-uniform case (Litrico and Fromion 2002), (Litrico and Fromion 2004a).

Finite order non-linear model: A discretized version of Saint-Venant's equations (in space, or time and space) can also be used (Liu, Feyen, and Berlamont 1994). A numerical approach is therefore used instead of an analytical one. The main advantage is to simplify the control design and to allow this approach for almost any type of canal system, while keeping the non-linear features of the system. The main limitation is that these models, based on numerical schemes such as the Preissmann scheme, as used in the literature so far, require subcritical flow. The same approach can maybe be extended to supercritical flow, using other schemes, but this was never tested according to the authors' knowledge. The numerical scheme introduces a discrepancy in the modelization, but this type of model can still be considered as very precise.

Finite order linear model (state space model): The non-linear or the infinite order feature of the previous models reduce the spectrum of control theories that can be used. In particular, all the methodologies developed in the field of LQ optimal controllers cannot be used. To allow this, a linear finite-order state-space model is required and can be obtained (lumped model obtained from linearization and discretization of the Saint-Venant's equations). The general form of this model is given above (Eq. 2). Here, the A, B, C and D matrices are constant (they do not depend on the time t nor on the state x). A similar model can also be constructed from the aggregation of different SISO transfer functions, between the different perturbation-control action variables and controlled-measured variables (Kosuth 1994), (Litrico, Malaterre, Georges, and Trouvat 1998).

Finite order linear model (transfer function): The advantage of the previous model is to be simple enough and to cope with the multivariable feature of the considered systems. But, on large systems, it can be costly in terms of required data, memory space, computation time (although progresses in computer and numerical tools decrease this burden). In order to overcome this difficulty, some simpler linear MIMO or SISO transfer function models can be used. These models can be first order, second order or second order with delay, depending on the size of the system and hydraulic conditions (Malaterre 1994), (Schuurmans 1997), (Schuurmans, Bosgra, and Brouwer 1995), (Sanfilippo 1997). The model parameters can even be obtained, in some cases, directly from geometric and hydraulic data (Rey 1990), (Malaterre 1994), (Litrico and Fromion 2004b). One advantage of such models is to be adapted to parametric identification, when field data is available. This can be done through openloop identification (Weyer 2000), (Weyer 2001), (Eurén and Weyer 2005), (Ooi, Krutzen, and Weyer 2005) or through closed loop identification: (Ooi and Weyer 2001).

Other original but limited modeling approaches have been studied and applied on irrigation canals: Neural network, Fuzzy and Petri Net models.

2.4 *Controller design techniques*

2.4.1 *Definitions*

Control theory implies a three step process: 1) system modeling (i.e. the definition of a model), 2) system analysis (i.e. the study of the model behavior), and 3) controller design. The design technique is the algorithm or methodology used within the control algorithm in order to generate the control action variables from the measured variables. The design techniques can be split into many different categories linear / non-linear, classical / robust, frequency / time domain, model based or not, feedback or closed loop / feedforward or open loop, centralized / decentralized, offline / on-line, monovariable SISO / multivariable MIMO methods. A main technique can benefit from additional components that may improve control algorithm performance by accounting for canal system features. Examples are filter, antiwindup, decoupler, observer, gain scheduling or autoadaptative tuning.

The controllers are also dynamical systems and therefore all the details indicated above for the models are also valid for the controllers. Nowadays they are mainly implemented in numerical form in Remote Terminal Units (RTU) or in SCADA softwares. Previously they were mainly implemented in electrical analogical circuits or in mechanical devices. All three options still exist for irrigation canals but with a strong tendency in favor of the numerical form. We can imagine other ways of implementing control loops. The domain where we can find the highest number of control loops is the living world, where each living organism encompasses from tenth to billions of loops triggered by biochemical reactions (Mitra and Khammash 2006).

A feedback controller K_n is able to calculate control action variables U at some devices using variables Y measured on the system, and control parameters p.

$$U(t) = K_n(Y(t'), p(t'), t') \tag{3}$$

As for models, for feedback controllers, the causality principle imposes $t' \leq t$. We can have only one controller K or a family of controllers K_n for different functioning conditions (e.g.: high flows or low flows, shortage context or flood context, etc). The situation where one shifts between several controller is called gain scheduling. Again, we have an interesting subset of controllers called LTI, having the same possible forms as above for models (e.g.: discrete state space as in Eq. 2).

Most of the technics that have been used so far, for the automation of irrigation canals, and used routinely for the real time control of canals are based on PID controllers. They are certainly the simplest and more widely used algorithms in all industrial applications. They belong to the LTI family of controllers since the control action variable U can be calculated from the measured (equal to the controlled) variable Y using the following linear equation:

$$U = -K_p \left(1 + \frac{1}{T_i s} + T_d s\right) Y \tag{4}$$

where K_p, T_i and T_d are the PID parameters (called respectively Proportional gain, Integral time, Derivative time), and s is the Laplace variable since the above equation is written in the Laplace frequency domain. It can have an equivalent form in the discrete state space format. It may seem a little bit restrictive that most of the control algorithms used in industry and in particular on irrigation canals are based on such simple mathematics. But, this PID controller family is a huge one, at least for 2 reasons:

- the PID algorithm is just the link between measured variables Y and control action variables U, but there are several options for the choice of these variables. As indicated previously, this choice of variables is at least as important as the algorithm itself, as far as the behavior and performances of the canal operation is concerned.
- the PID structure is simple as described in Eq. 4, but the selection of the K_p, T_i and T_d is the difficult part and there are plenty of methods for doing that having a great impact on the behavior and performances of the canal operation (e.g.: non-linear optimization, linear optimization, gain and phase margin specification, Ziegler-Nichols time response method, etc.).

Due to the time delays inherent to the hydraulic systems, to non-linear features, and to interactions among subsystems, many other methods have been tested and are still under development. Some have been used routinely on some irrigation canals (Internal Model, Predictive Control, Backward Computation and Fuzzy Control), or experimentally on some real canals or scale models (LQG, $\mathcal{H}_\infty$) (Malaterre, Rogers, and Schuurmans 1998). A large part of these algorithms can be written in state space form or in RST Form (Ruiz-Carmona, Clemmens, and Schuurmans 1998). Their resulting mathematical forms are therefore not so different. The main difference is the order of the controller (number of states) and the way the control parameters are tuned.

Due to the time delays and the strong perturbations occurring on the systems, a feedforward controller can advantageously be added to the feedback controller. Several ways of combining these 2 controllers are possible: separate design and addition (Deltour 1988), (Bautista and Clemmens 1998), (Bautista and Clemmens 1999) or combined design (Malaterre 1998), (Duviella, Charbonnaud, Chiron, and Carrillo 2005).

2.4.2 *Examples*

A list of methods is given hereafter, with the basic and recent references. For complete literature references, refer to (ASCE 1998), (Malaterre, Baume, Pourciel, Dahhou, Aguilar, and Benhammou 1997), (Malaterre, Rogers, and Schuurmans 1998).

Heuristic monovariable methods have been developed based on hydraulics and not on control theory (e.g.: Zimbelman, CARDD). Although quoted in the literature they are hardly operational and too site specific. LittleMan is an empirical method based on a three position controller (Rogers, Ehler, Falvey, Serfozo, Voorheis, Johansen, Arrington, and Rossi 1995). These methods have to be tuned on a complete simulation model, or on the real system, since no mathematical tool can enable this, nor prove their performances. This is one main drawback of heuristic methods. Very few of them are still used.

Examples of PID related methods are: P: AMIL, AVIS, AVIO; PI: ELFLO, BIVAL, Dynamic Regulation, (Weyer 2002), DACL (Casangcapan and Chilcott 1993; Casangcapan, Bright, and Chilcott 1993), PI in RST form (Prado-Hernández, De León-Mojarro, Ruiz-Carmona, Exebio-García, and Mejía- Sáenz 2003). Some are tuned using a simplified model of the process (e.g.: by pole placement on a SISO transfer function), some are tuned directly on the real process or on a full non-linear simulation model using the ATV method (Litrico and Malaterre 2007), (Litrico, Malaterre, Baume, Vion, and Ribot-Bruno 2007). When the global performance of the PI controller is not sufficient, several PI coefficients can be optimized for different flows conditions (gain scheduling) (Overloop, Schuurmans, Brouwer, and Burt 2005). Decentralized PI can be optimized offline using linear-quadratic optimization (Clemmens and Schuurmans 2004) or non-linear optimization (Baume, Malaterre, and Sau 1999).

Smith Predictor: Although very efficient in most cases, PID controllers do not explicitly take into account the characteristic canal time delays. (Shand 1971) prospected the possibility to use a Smith Predictor in order to overcome this problem, when studying the automation of Corning Canal, California, USA. Developing an analog dead time model raised technological difficulties, at this time. Therefore, though less efficient, ELFLO method was eventually selected. Recently, the combination of a PI controller with a Smith Predictor was further developed (Deltour 1992), (Sanfilippo 1997). This controller is called PIR. Modern digital technology has solved problems faced by analogical circuits.

Internal model: The dynamics of the process to control is modeled and used in the controller (Duviella, Charbonnaud, Chiron, and Carrillo 2005). It can be seen as an extension of the Smith Predictor technic.

Pole placement: Other linear controllers have been used on river systems with long time delays by CACG, (Piquereau, Tardieu, Verdier, and Villocel 1984). High order transfer functions are used, and tuned with the pole placement technique. It has been used routinely for more than 20 years on the CACG Neste system, France.

Predictive control: The predictive control method, an usually monovariable optimization method, has been applied to canal systems by several authors (Sawadogo 1992), (Rodellar, Gomez, and

Bonet 1993), (Ruiz-Carmona and Ramirez Luna 1998), (Pagès, Compas, and Sau 1998), (Akouz, Benhammou, Malaterre, Dahhou, and Roux 1998), (Tadeo, Alvarez, del Valle, and Rivas 2001), (Overloop 2006). It is not based on the desired closed-loop behavior, but on the minimization of a criterion J, weighting the control action variable and the error between the controlled variable and its targeted value. Predictive control method uses transfer function or state space models (Malaterre and Rodellar 1997). It can naturally incorporate an open-loop and a closed-loop. It is used routinely on the CNR systems (Compagnie Nationale du Rhne, France).

Methods based on fuzzy control (Bouillot 1994), (Voron and Bouillot 1997), (Stringam and Merkley 1997b), (Stringam and Merkley 1997a), expert systems, or neural networks (Schaalje and Manz 1993), (Toudeft 1996) have been developed. Genetic Algorithms have also been tested on such systems. At least 3 successful field applications can be reported for fuzzy control, on SISO systems (T2, CPBS, and Roosevelt canals). In some cases a fuzzy approach is used to switch between several classical controllers (LQG in (Begovich, Salinas, and Ruiz-Carmona 2003)) in order to adapt the controller to the field conditions (this is called gain scheduling).

Model inversion: Different model inversion methods (also called Backward Computation) are described in the literature, leading generally to open-loop controllers (Falvey 1987), (O'Loughlin 1972), (Chevereau 1991), (Liu, Feyen, and Berlamont 1992), (Bautista and Clemmens 1998), and more rarely to closed-loop controllers (Liu, Feyen, and Berlamont 1994). They are based on a finite-order non-linear model. Inverting such model including diffusive dynamics can generate oscillatory behavior. In order to prevent this, damper coefficients have to be introduced. They have to be tuned by try and error procedures since no mathematical tool can enable this.

Optimization methods have also been developed. These methods are, in essence, multivariable. Different methods exist: linear optimization (Sabet, Coe, Ramirez, and Ford 1985), non-linear optimization (Tomicic 1989), (Khaladi 1992), (Zihui and Manz 1992), (Soler i Guitart 2003), LQR-LQG (Corriga, Sanna, and Usai 1983), (Balogun, Hubbard, and DeVries 1988), (Filipovic and Milosevic 1989), (Malaterre 1998), (Clemmens 1999), (Weyer 2003). The classical non-linear optimization leads solely to an open-loop, sensitive to errors and perturbations. In order to introduce a closed-loop, the optimization has to be processed periodically (for example at each time step). This complicates the method and limits its applications due to real-time constraints. Furthermore, the determination of real initial conditions, required for the optimization, is not easy. On the other hand, *LQR – LQG* methods, based on a state space representation, can incorporate in a straightforward manner an open-loop and a closed-loop (Malaterre 1998). $\mathcal{H}_2$ norm minimization has been tested on canal systems. This approach has the advantage to allow the choice of the structure of the multivariable controller, and in particular to design a decentralized controller (Seatzu, Giua, and Usai 1998), (Schuurmans 1997), (Seatzu 2000). When the order of the system is too large, or in order to design a decentralized controller, a decomposition-coordination approach can be used (Kosuth 1994), (El Fawal, Georges, and Bornard 1998).

Robust control: since some important characteristics of the considered systems are the strong unknown perturbations, and the model errors partly due to non-linear effects and neglected dynamics, robust control approaches are interesting and have been tested by several authors. Robust control design can be based on several induced norms such as the $\mathcal{H}_\infty$ norm (Schuurmans 1997), (Litrico, Georges, and Trouvat 1998), (Pognant-Gros 2000), (Pognant-Gros 2003), (Li, Cantoni, and Weyer 2004) or the ℓ_1 norm (Malaterre and Khammash 2000), (Malaterre and Khammash 2003). Multi-norm optimization is also possible (Qi, Khammash, and Salapaka 2001). The first approach ($\mathcal{H}_\infty$ norm) has the advantage of having a limited order state space solution. The second approach (ℓ_1 norm) has the advantage of having a time domain interpretation and, nowadays, a solution through an efficient linear optimization problem.

Adaptive control: Non-linear features of the model can be taken into account by using adaptive linear controllers (Cardona, Gomez, and Rodellar 1997) or by using gain scheduling (Sanfilippo 1997).

Non-linear control: A direct non-linear approach can also be used (Coron, d'Andrea Novel, and Bastin 1999), (De Halleux, Bastin, d'Andrea Novel, and Coron 2001), but on simple systems, reducing the interest of such methods for the moment. Other non-linear methods use collocation

models (Georges, Dulhoste, and Besanç on 2000), (Dulhoste, Besançon, and Georges 2001), (Dulhoste, Georges, and Besancon 2004).

3 COMPARISON OF CONTROLLERS

3.1 *Benchmark canals*

Until recently, it was very difficult to have a proof of the advantages and drawbacks of the different above methods, and on their performances on canal systems, except those claimed by their designers. This was due to the fact that each author designed a new method on a new canal system, without comparing its method to others. Some synthesis work started to fill the gap (Zimbelman 1987), (Goussard 1993), (Malaterre 1995b), (Malaterre 1995a), (Malaterre, Rogers, and Schuurmans 1998). A task committee of the ASCE proposed to push this effort further, and defined benchmark Test Cases. Some authors started to compare their methods on the proposed Test Cases, (ASCE 1998). So far 3 methods have been fully tested on them: CLIS based on backward computation (Liu, Feyen, Malaterre, Baume, and Kosuth 1998), PILOTE based on LQG (Malaterre 1998) and PIR based on PI + Smith predictor (Sanfilippo 1997).

Some other authors have used only part of the benchmark (Akouz, Benhammou, Malaterre, Baume, Sawadogo, and Dahhou 1999), (Wahlin and Clemmens 2002), (Wahlin 2004), (Bautista and Clemmens 2005). But these results are hardly useful in a benchmarking process since some important features of the benchmark have been skipped (e.g.: minimum gate movements, untuned scenarios, only one canal out of two). Some people complained about the complexity of the benchmark. It is true they could be simplified, but they had the merit to exist and could play an important role.

Even though these results have been obtained from simulation on the same canal and scenarios, it is still difficult to make a definitive comparison. First of all, each author used its own simulation model, with some differences, since the Test Cases left some minor parameters undefined (such as gate equations and gate sizes). Then, the objectives of the different authors were maybe different during the design and tuning process. We can observe that PILOTE obtains the best performances on the controlled variables (up to 50% improvement), but with more aggressive use of control action variables (Malaterre and Baume 1998). The ability of a multivariable controller to allow aggressive actions (in a reasonable domain), while staying stable and non-oscillatory is one of the proofs of its quality. But maybe all authors did not try to get this type of performance. This means that these Test Cases have probably to be improved, and the terms of references to be clarified, in order to provide a better basis for comparison, which seems to be an interesting objective.

Based on dimensionless analysis, Cemagref defined other benchmark canals (Baume, Sau, and Malaterre 1998), available at the Internet web site http://canari.malaterre.net. These canals have been defined from two dimensionless coefficients, in order to cover all kinds of hydraulic behaviors.

3.2 *Manual control of an irrigation canal*

A game on manual control has been played at 3 French Engineering Schools for many years. Students perform manual control using the SIC simulation model, which solves Saint-Venant equations and enables interaction with users by the means of a regulation module. This experience is carried out as an interactive game where the students have to make decisions (gate operations whenever they want during the 12 h scenario), can ask questions and have a final score that can be compared to other student classes and automatic controllers. The duration of the game is about 2 h00. The canal to be controlled during this game is the canal 1 of the Cemagref Benchmark (Baume, Malaterre, and Sau 1999).

The student class is divided into 5 groups, each of them having the responsibility of managing one gated cross regulator in order to stabilize the water level at the downstream end of its downstream pool. The type of regulation we choose for this canal therefore belongs to the "distant downstream control" class (Malaterre, Rogers, and Schuurmans 1998) which is the more interesting but challenging one. The students do not know when the offtake discharges change nor in which

quantity, being in the same position as the manager of a canal functioning on a on-demand basis. The performance indicator used to give a mark to the students is the absolute sum of the maximum water level deviations around the targeted value (which is the initial steady state value) obtained at the 5 controlled locations. At the end of the game we observe the marks obtained, compare them to other controllers and make a summary of the concepts that have been addressed and understood by the students during the game (Litrico, Malaterre, Harran, and Alquier 2006).

We can observe, even on this simple canal and scenario, and allowing the student gate keepers to operate every 5 minutes (which is not very realistic), that most automatic controllers perform better than the manual operators 3. The best results are obtained on this example with a centralized ℓ_1 controller 5(Malaterre and Khammash 2003). Even though the optimized decentralized PI controllers (Baume, Malaterre, and Sau 1999) obtain reasonable results we can observe an oscillatory behavior that should be corrected 4. The best students (Enseeiht 2006) could achieve these results 6 making many operations and making guesses and calculations using rough models for reaches and gates. It is certain that they could not compete with the automatic controllers on a more complex scenario. To prevent misinterpretation of the comparison between classes, I must say that the performance they obtain is also linked to the help and information I provide them (the only information I never give, even if asked, is about the offtake perturbations).

4 CONCLUSION

This paper describes the context of irrigation canals and the increasing incentive to modernize their design and regulation procedures. It then presents a clear structure for characterizing each control algorithm according to a series of defined criteria (Considered variables, Logic of control, Models and Controllers Design techniques). The advantage of this structure is to allow the comparison of different algorithms' characteristics and to make possible the classification of them according to a selected criterion.

Canal control algorithms detailed in the literature are listed according to their main design technique (e.g. heuristic, PID, LQG, etc.). Only the most recent references are given in this paper. Previous references can be found in (Zimbelman 1987), (Goussard 1993) or (Malaterre, Rogers,

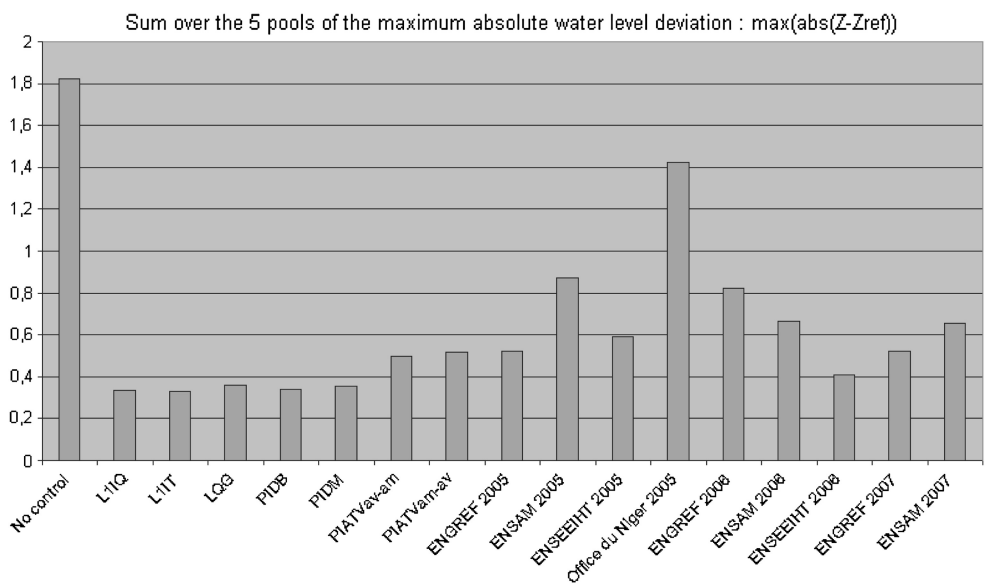
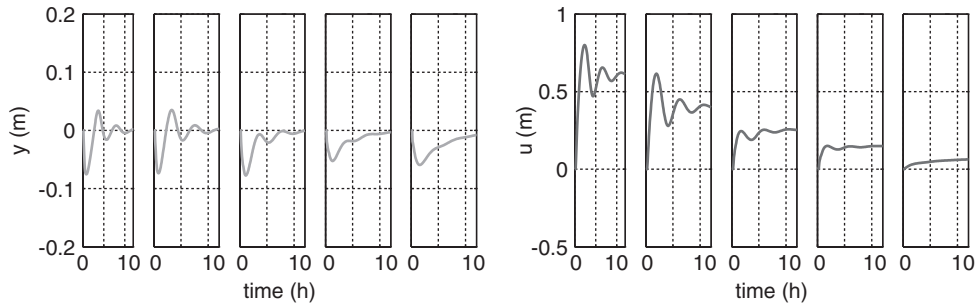
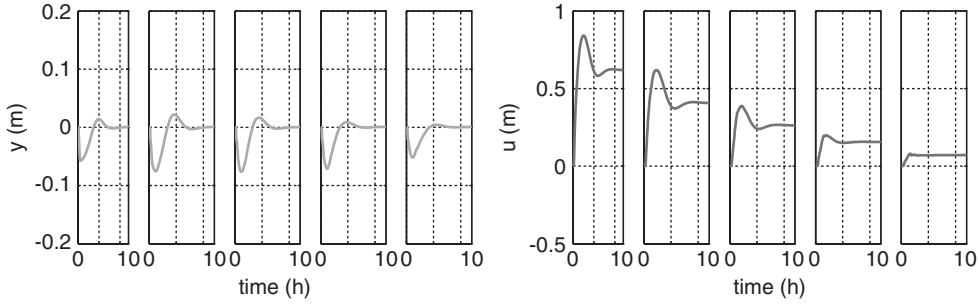


Figure 3. Performance indicator on water level deviations.



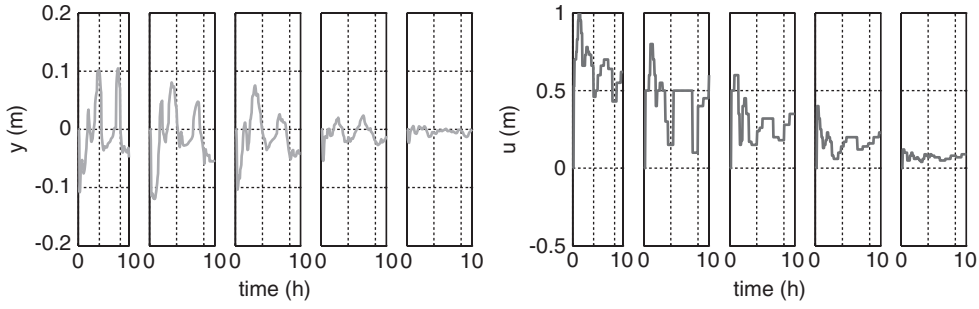
(a) Controlled variables y_1 to y_5 (b) Control action variables u_1 to u_5

Figure 4. Controlled and control action variables for PID_B.



(a) Controlled variables y_1 to y_5 (b) Control action variables u_1 to u_5

Figure 5. Controlled and control action variables for ℓ_{1iq}.



(a) Controlled variables y_1 to y_5 (b) Control action variables u_1 to u_5

Figure 6. Controlled and control action variables for Enseeiht06.

and Schuurmans 1998). Additional and updated information on canal regulation is available on the CANARI web site: http://canari.malaterre.net.

Such characterizations and classifications are useful to get a better understanding of the features and properties of each regulation method. Indeed, the characteristics of each regulation method will have corresponding advantages, disadvantages, performance and constraints. Canal operators and engineers should find these classifications useful to determine appropriate regulation methods for specific installations.

The difficulty is then to find the good compromise between a simple controller with (maybe) limited performances, and a complex controller with (maybe) better performances. The authors think that the solution lies, for the moment, in the group of robust multivariable linear time invariant controllers, such as partially decentralized robust PI, LQG, $\mathcal{H}_\infty$ or ℓ_1. Still important efforts have to be made to improve the robustness of the algorithms (robust control, adaptive control, non-linear control), to improve the performance of monovariable controllers (decoupler, predictive controllers, internal model controllers), to reduce the computational efforts of the multivariable controllers (decentralized controllers), and to take into account possible defaults on sensors or actuators. Important efforts have also to be made to implement the methodologies into numerical tools (e.g. simulation models) so that engineers and canal managers can use these methods. A close interaction of the different people working in this field, with different approaches, and the definition of a common basis of comparison of the developed methods such as the above benchmarks will facilitate this work. The managers and engineers will then be able to select the controller providing the good compromise performance/complexity they want, depending on their specific context and constraints.

REFERENCES

Akouz, K., A. Benhammou, P.-O. Malaterre, J.-P. Baume, S. Sawadogo, and B. Dahhou (1999, August 31 – September 3). Predictive control of a portion of asce canal 2. *European Control Conference (ECC99), Karlsruhe, Germany*.

Akouz, K., A. Benhammou, P.-O. Malaterre, B. Dahhou, and G. Roux (1998, October 11 to 14). Predictive control applied to asce canal 2. *IEEE Int. Conference on Systems, Man and Cybernetics (SMC'98), San Diego, California 1*, 3920–3924.

ASCE (1993, July/August). Unsteady-flow modeling of irrigation canals. *Journal of Irrigation and Drainage Engineering 119*(4), 615–630.

ASCE (1998, Jan./Feb). Special issue on canal automation. *Journal of Irrigation and Drainage Engineering 124*(1), 1–62.

Balogun, O., M. Hubbard, and J. DeVries (1988, January). Automatic control of canal flow using linear quadratic regulator theory. *Journal of Hydraulic Engineering 114*(1), 75–102.

Barréde Saint-Venant, A. J. C. (1871). Théorie du mouvement non-permanent des eaux avec application aux crues des rivières et à l'introduction des marées dans leur lit. *Comptes rendus des séances de l'Académie des Sciences 73*, 148–154.

Baume, J., P. Malaterre, G. Belaud, and B. LeGuennec (2005). Sic: a 1D hydrodynamic model for river and irrigation canal modeling and regulation. In *Métodos Numéricos em Recursos Hidricos 7*, pp. 1–81. Associacao Brasileira de Recursos Hidricos (ABRH).

Baume, J.-P., P.-O. Malaterre, and J. Sau (1999, October 18–21). Tuning of PI to control an irrigation canal using optimization tools. *ASCE-ICID Workshop on Modernization of Irrigation Water Delivery Systems, in Phoenix, Arizona, USA 1*, 1–16.

Baume, J.-P. and J. Sau (1997, April 22–24). Study of irrigation canal dynamics for control purpose. *International Workshop on the Regulation of Irrigation Canals: State of the Art of Research and Applications, RIC97 1*, 3–12.

Baume, J.-P., J. Sau, and P.-O. Malaterre (1998, October 11–14). Modeling of irrigation channel dynamics for controller design. *IEEE International Conference on Systems, Man and Cybernetics (SMC'98), San Diego, California October 11–14*, 3856–3861.

Bautista, E. and A.J. Clemmens (1998, August 3–7). An open-loop control system for open-channel flows. *ASCE Int. Water Resources Engineering Conference, Memphis August 3–7*, 1852–1857.

Bautista, E. and A.J. Clemmens (1999, July/August). Response of asce task committee test cases to open-loop control measures. *Journal of Irrigation and Drainage Engineering 125*(4), 179–188.

Bautista, E. and A.J. Clemmens (2005, November/December). Volume compensation method for routing irrigation canal demand changes. *ASCE, J. Irrig. Drain. Eng. 131*(6), 494–503.

Bedjaoui, N. (2007). *Supervision dynamique pour la gestion oprationnelle d'un canal d'irrigation*. Ph. D. thesis, LAG Grenoble.

Bedjaoui, N., X. Litrico, D. Koening, and P.-O. Malaterre (2006). H_∞ observer for time-delay systems. application to fdi for irrigation canals. In *IEEE Conference on Decision and Control (CDC), 2006, San Diego, USA*.

Begovich, O., J. Salinas, and V. Ruiz-Carmona (2003, October 6–8). Real-time implementation of a fuzzy gain scheduling control in a multi-pool open irrigation canal prototype. In *Proceedings of the 2003 IEEE International Symposium on Intelligent Control, Houston, Texas USA*, pp. 304–309.

Bouillot, A.-P. (1994, October). Application of the fuzzy set theory to the control of a large canal. *ASCE – AFEID Meeting, Aix-en-Provence, France 1*, 25.

Bounit, H., H. Hammouri, and J. Sau (1997, April 22–24). sRegulation of an irrigation canal through the semigroup approach. *International Workshop on the Regulation of Irrigation Canals: State of the Art of Research and Applications, RIC97 1*, 261–267.

Canivet, E. (2002, Février). *Réconciliation et validation des mesures sur un système hydraulique complexe : Le Canal de Provence*. Ph. D. thesis, Université Claude Bernard – Lyon 1.

Cardona, J., M. Gomez, and J. Rodellar (1997, April 22–24). A decentralized adaptive predictive controller for irrigation canals. *International Workshop on the Regulation of Irrigation Canals: State of the Art of Research and Applications, RIC97, Marrakech (Morocco) April 22–24*, 215–219.

Casangcapan, M., J. Bright, and R. Chilcott (1993). DACL system dynamic performance. II : Laboratory model testing. *Journal of irrigation and drainage engineering 119*(1), 64–73.

Casangcapan, M. and R. Chilcott (1993). DACL system dynamic performance. I : Response-prediction analysis. *Journal of irrigation and drainage engineering 119*(1), 50–63.

Chevereau, G. (1991). *Contribution à l'étude de la régulation dans les systèmes hydrauliques à surface libre*. Ph. D. thesis, Institut National Polytechnique de Grenoble.

Choy, S. and E. Weyer (2005, 12–15 Dec). Reconfiguration scheme to mitigate faults in automated irrigation channels. In *44th IEEE Conference on Decision and Control*, pp. 1875 – 1880.

Choy, S. and E. Weyer (2006, December 13–15). Reconfiguration schemes to mitigate faults in automated irrigation channels: Experimental results. In *45th IEEE Conference on Decision and Control*, Manchester Grand Hyatt Hotel, San Diego, CA, USA. IEEE.

Clemmens, A. and J. Schuurmans (2004). Simple optimal downstream feedback canal controllers: Theory. *J. Irrig. Drain. Eng. 130*(1), 26–34.

Clemmens, A.J. (1999, June). Real-time operation of water resources systems: On- versus off-line optimization. In *26th Annual conference of the Water Resources Planning and Management Division of the American Society of Civil Engineers.*, pp. 10. Tempe, Arizona June 6–9, 1999.

Clemmens, A. J., R. Strand, and L. Feuer (2002, May 15). Application of canal automation in central arizona. *United States Committee Of Irrigation And Drainage Engineering Conference*, 453–464.

Coron, J., B. d'Andrea Novel, and G. Bastin (1999). A Lyapunov approach to control irrigation canals modeled by Saint-Venant's equations. In *International European Control Conference ECC'99*, Karlsruhe, Allemagne.

Corriga, G., S. Sanna, and G. Usai (1983). Sub-optimal constant-volume control for open channel networks. *Applied Mathematical Modelling 7*, 262–267.

Cunge, J. A., F. Holly, and A. Verwey (1980). *Practical aspects of computational river hydraulics*. Pitman Advanced Publishing Program.

De Halleux, J., G. Bastin, B. d'Andrea Novel, and J.-M. Coron (2001, July). A lyapunov approach for the control of multi-reach channels modelled by saint venant equations. *CD-Rom IFAC NOLCOS 01 Conference, St Petersburg July*, 1515–1520.

De Halleux, J., C. Prieur, J.-M. Coron, B. D'Andrea-Novel, and G.Bastin (2003). Boundary feedback control in networks of open channels. *Automatica 39*(8), 1365–1376.

Deltour, J.-L. (1988). La régulation des systèmes d'irrigation. Master's thesis, Ecole Nationale Supérieure d'Hydraulique de Grenoble.

Deltour, J.-L. (1992). *Application de l'automatique numérique à la régulation des canaux ; Proposition d'une méthodologie d'étude*. Ph. D. thesis, Institut National Polytechnique de Grenoble. in French.

Deltour, J.-L., F. Sanfilippo, E. Canivet, and J. Sau (2005). Data reconciliation on the complex hydraulic system of canal de provence. *Journal of Irrigation and Drainage Engineering 131*(3), 291–297.

Dulhoste, J.-F., G. Besancon, and D. Georges (2001, septiembre). Non-linear control of water flow dynamics by input-output linearization based on a collocation model. In *European Control Conference Ecc'01 Porto, Portugal*.

Dulhoste, J.-F., D. Georges, and G. Besancon (2004). Nonlinear control of open-channel water flow based on collocation control model. *Journal of Hydraulic Engineering 130*(3), 254–266.

Duviella, E., P. Charbonnaud, P. Chiron, and F. Carrillo (2005). Supervised internal multi-model control of a dam-gallery open-channel system. In *44th IEEE Conference on Decision and Control, and the European Control Conference, CDC-ECC '05*, Volume 2005, Seville, pp. 1887–1892.

El Fawal, H., D. Georges, and G. Bornard (1998, October 11–14). Optimal control of complex irrigation systems via decomposition-coordination and the use of augmented lagrangian. *IEEE International Conference on Systems, Man and Cybernetics (SMC'98) October 11–14*, 3874–3879.

Eurén, K. and E. Weyer (2005). System identification of open water channels with undershot and overshot gates. In *16th IFAC World Congress*, Prague, Czech Republic, pp. –.

Falvey, H. (1987). Philosophy and implementation of gate stroking. *Proceedings ASCE Portland, Zimbelman D.D. (ed.) 1*, 176–179.

Filipovic, V. and Z. Milosevic (1989). Dyn^2 method for optimal control of water flow in open channels. *Journal of irrigation and drainage engineering 115*(6), 973–981.

Georges, D., J.-F. Dulhoste, and G. Besançon (2000, June 19 – 23). Modelling and control of water flow dynamics via a collocation method. *Proceedings of the Fourteenth International Symposium of Mathematical Theory of Networks and Systems, MTNS 2000, Perpignan, France CD*, 10.

Godaliyadda, G., D. Renault, H. Hemakumara, and I. Makin (1999). Strategies to improve manual operation of irrigation systems in sri lanka. *Irrigation and Drainage Systems*, 33–54.

Goussard, J. (1993). *L'automatisation des réseaux d'irrigation en canaux*. Commission Internationale des Irrigations et du Drainage, Groupe de travail sur la Construction, la Réhabilitation et la Modernisation des Projets d'Irrigation, ICID.

Goussard, J. (2000, September). Canal operation simulation model. Technical report, ICID.

Institut Méditerranéen de l'Eau (2007, 19–21 Mars). 3éme atelier régional. In *La demande en eau en Méditerranée*, Saragosse (Espagne), pp. 18.

Khaladi, A. (1992). *Gestion automatique des transferts d'éau en réseaux hydrauliques maillés à surface libre; Application au réseau des Wateringues*. Ph. D. thesis, Ecole Centrale de Lyon, CEMAGREF.

Koenig, D., N. Bedjaoui, and X. Litrico (2005). Unknown input observers design for time-delay systems. application to an open-channel. In *44th Conference on Decision and Control and European Control Conference ECC'05, Sevilla, Spain, Decembre 12–15, 2005*, Volume 2005, Seville, pp. 5794–5799.

Kosuth, P. (1994). *Techniques de régulation automatique des systèmes complexes: application aux systèmes hydrauliques à surface libre*. Ph. D. thesis, Institut National Polytechnique de Toulouse – CEMAGREF – LAAS CNRS. 334 p.

Li, Y., M. Cantoni, and E. Weyer (2004). Design of a centralized controller for an irrigation channel using hinf loop-shaping. In *Proc UKACC Control Conference*, University of Bath, UK, pp. –.

Litrico, X. and V. Fromion (2002, December). Infinite dimensional modeling of open-channel hydraulic systems for control purposes. In *41th IEEE Conference on Decision and Control*, Volume1, Las Vegas, Nevada.

Litrico, X. and V. Fromion (2004a). Frequency modeling of open-channel flow. *Journal of Hydraulic Engineering*, 806–815.

Litrico, X. and V. Fromion (2004b). Simplified modeling of irrigation canals for controller design. *Journal of Irrigation and Drainage Engineering*, 373–383.

Litrico, X., D. Georges, and J.-L. Trouvat (1998, October 11 to 14). Modelling and robust control of a dam-river system. *IEEE International Conference on Systems, Man and Cybernetics (SMC'98), San Diego, California 4*, 3862–3867.

Litrico, X. and P.-O. Malaterre (2007, June). Test of auto-tuned automatic downstream controllers on gignac canal. In *USCID conference*, Denver.

Litrico, X., P.-O. Malaterre, J.-P. Baume, and J. Ribot-Bruno (2007). Conversion from discharge to gate opening for the control of irrigation canals. *J. Irrig. Drain. Eng.*.

Litrico, X., P.-O. Malaterre, J.-P. Baume, P.-Y. Vion, and J. Ribot-Bruno (2007). Automatic tuning of pi controllers for an irrigation canal pool. *Journal of Irrigation and Drainage Engineering 133*(1), 27–37.

Litrico, X., P.-O. Malaterre, D. Georges, and J.-L. Trouvat (1998, Septembre). Commande optimale d'un système barrage-rivière avec points de mesure intermédiaires. In *Colloque SHF, 25èmes Journées de l'Hydraulique*, Volume 2, Chambéry.

Litrico, X., P.-O. Malaterre, G. Harran, and J.-F. Alquier (2006). Hands on teaching for real-time control of open-channels. *7th International Conference on Hydroinformatics 2006 Acropolis – Nice, FRANCE 4th – 8th September 2006 1*.

Liu, F., J. Feyen, and J. Berlamont (1992). Computation method for regulating unsteady flow in open channels. *Journal of irrigation and drainage engineering 118*(10), 674–689.

Liu, F., J. Feyen, and J. Berlamont (1994). Downstream control algorithm for irrigation canals. *Journal of irrigation and drainage engineering 120*(3), 468–483.

Liu, F., J. Feyen, P.-O. Malaterre, J.-P. Baume, and P. Kosuth (1998, January/February). Development and evaluation of canal automation algorithm CLIS. *Journal of Irrigation and Drainage Engineering 124*(1), 40–46. ISSN 0733–9437.

Lopez-Cantens, G., V. Ruiz-Carmona, E. Arteaga-Tovar, and J. Ostos-Santos (2007). Orientations of the control theory in the automation of canals. *Ing. Hidraul. Mex. 22*(1), 117–123.

Malaterre, P.-O. (1994). *Modélisation, analyse et commande optimale LQR d'un canal d'irrigation*. Ph.D. thesis, ISBN 2-85362-368-8, Etude EEE n°. 14, LAAS – CNRS – ENGREF – Cemagref.

Malaterre, P.-O. (1995a). La régulation des canaux d'irrigation: caractérisation et classification. *La Houille Blanche 5*(6), 17–35.

Malaterre, P.-O. (1995b, November). Regulation of irrigation canals: Characterisation and classification. *International Journal of Irrigation and Drainage Systems 9*(4), 297–327.

Malaterre, P.-O. (1998, July/August). PILOTE: linear quadratic optimal controller for irrigation canals. *ASCE Journal of Irrigation and Drainage Engineering 124*(4), 187–194. ISSN 0733–9437.

Malaterre, P.-O. and J.-P. Baume (1997, April 22–24). Sic 3.0, a simulation model for canal automation design. *International Workshop on the Regulation of Irrigation Canals: State of the Art of Research and Applications, RIC97, Marrakech (Morocco) 1*, 68–75.

Malaterre, P.-O. and J.-P. Baume (1998, October 11 to 14). Modeling and regulation of irrigation canals: existing applications and ongoing researches. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC'98)*, Volume 4, San Diego, California, pp. 3850–3855.

Malaterre, P.-O. and J.-P. Baume (1999, October 18–21). Optimum choice of control action variables and linked algorithms. Comparison of different alternatives. *ASCE-ICID Workshop on Modernization of Irrigation Water Delivery Systems, in Phoenix, Arizona, USA*.

Malaterre, P.-O. and J.-P. Baume (2007, March). *SIC : theoretical concepts and user's guide. V4.25*. 361 rue Jean-Francois Breton – BP 5095, 34196 MONTPELLIER cedex 5, FRANCE: UMR G-EAU "Gestion de l'Eau, Acteurs et Usages" – Cemagref – Cirad – ENGREF – IRD – TR Transcan (modélisation et régulation des canaux et rivières).

Malaterre, P.-O., J.-P. Baume, J.-B. Pourciel, B. Dahhou, J. Aguilar, and A. Benhammou (1997, April 22–24). *Proceedings of the International Workshop on the Regulation of Irrigation Canals: State of the Art of Research and Applications, RIC97, Marrakech (Morocco)*. Cemagref, Laas, University of Marrakech. in English.

Malaterre, P.-O. and C. Chateau (2007, June). Scada interface of the sic software for easy real time application of advanced regulation algorithms. In *USCID conference*, Denver.

Malaterre, P.-O. and M. Khammash (2000, December). ℓ_1 controller design for a high-order 5-pool irrigation canal system. In *39th IEEE-CDC conference*, Sydney, Australia.

Malaterre, P.-O. and M. Khammash (2003, December). ℓ_1 controller design for a high-order 5-pool irrigation canal system. *ASME Journal of Dynamic Systems, Measurement, and Control 125*(4), 639–645.

Malaterre, P.-O., J. Rey, and J.-P. Baume (1993, May). Use of a simulation model to improve the manual operation procedures of an irrigation canal (Kirindi Oya – Sri Lanka). *Asian Regional Symposium on "Maintenance and Operation of Irrigation/Drainage Schemes for Improved Performances", Beijing, China, HR Wallingford*, session 3, 27–40.

Malaterre, P.-O. and J. Rodellar (1997, April 22–24). Multivariable predictive control of irrigation canals. design and evaluation on a 2-pool model. *International Workshop on the Regulation of Irrigation Canals: State of the Art of Research and Applications, RIC97 1*, 230–238.

Malaterre, P.-O., D.C. Rogers, and J. Schuurmans (1998, January/February). Classification of canal control algorithms. *ASCE Journal of Irrigation and Drainage Engineering 124*(1), 3–10. ISSN 0733–9437.

Mareels, I., E.Weyer, S. Ooi, M. Cantoni, Y.Li, and G. Nair (2005). Systems engineering for irrigation systems: Successes and challenges. *Annual Reviews in Control 29*(2), 191–204.

Mitra, P.P. and M. Khammash (2006, September). Special issue on "control biology – emerging field of life science that connects biology and control". *Asian Journal of Control 8*(III–IV).

O'Loughlin, E. (1972). Application of unsteady flow analysis to operation decisions in long aqueducts. *VIIIth ICID Congress, Bulgaria R. 16, Q. 28–2*, 207–223.

Ooi, S. and E. Weyer (2001). Closed loop identification of an irrigation channel. In *Proceedings of the IEEE Conference on Decision and Control*, Volume 5, CSSIP Group, Dept. of Elec. and Electron. Eng., The University of Melbourne, Parkville, Vic. 3010, Australia, pp. 4338–4343.

Ooi, S.K., M.P.M. Krutzen, and E. Weyer (2005, April). On physical and data driven modelling of irrigation channels. *Control Engineering Practice 13*(4), 461–471.

Overloop, P. J.V., J. Schuurmans, R. Brouwer, and C.M. Burt (2005). Multiple-model optimization of proportional integral controllers on canals. *Journal of Irrigation and Drainage Engineering 131*(2), 190–196.

Overloop, P.V. (2006, June). *Model Predictive Control on Open Water Systems*. Ph. D. thesis, Delft. 192 pp.

Pagès, J., J. Compas, and J. Sau (1998, October 11–14). MIMO predictive control with constraints by using an embedded knowledge based model. *IEEE International Conference on Systems, Man and Cybernetics (SMC'98) October 11–14*, 3902–3907.

Piquereau, A., H. Tardieu, J. Verdier, and A. Villocel (1984). Réduction des pertes en eau par l'automatisation de la gestion d'un barrage realimentant une rivière en vue de l'irrigation. *Communication au XIIème congrès de la CIID Question 38, Chap 1, Paragraphe 5*, 16 p.

Pognant-Gros, P. (2000). Modélisation et commande H_∞ d'un canal d'irrigation. M. Sc. thesis, Université de Caen, Laboratoire d'Automatique et de Commande des Processus de l'ISMRA, INRA Biométrie, Cemagref.

Pognant-Gros, P. (2003). *Réduction de modèle, application à la commande d'un canal d'irrigation*. Ph. D. thesis, Université de Caen, Laboratoire d'Automatique et de Commande des Processus de l'ISMRA.

Prado-Hernández, J.V., B. De León-Mojarro, V.M. Ruiz-Carmona, A.Exebio-García, and E.Mejía-Sáenz (2003). Evaluación experimental de algoritmos de control de canales de riego. *Agrociencia 37*, 393–403.

Qi, X., M.H. Khammash, and M.V. Salapaka (2001, December). A matlab package for multiobjective control synthesis. *40th IEEE conference on Decision and Control 4*, 3991–3996. Orlando, Florida.

Rey, J. (1990). Contribution à la modélisation et la régulation des transferts d'éau sur des systèmes de type rivière/baches intermédiaires. Master's thesis, CEMAGREF, ENGREF, DEA National d'Hydrologie USTL Montpellier. 80 p.

Rodellar, J., M. Gomez, and L. Bonet (1993). Control method for on-demand operation of open-channel flow. *Journal of Irrigation and Drainage Engineering 119*(2), 225–241.

Rogers, D., D. Ehler, H. Falvey, E. Serfozo, P. Voorheis, R. Johansen, R. Arrington, and L. Rossi (1995). Canal systems automation manual, volume 2. Technical report, A Water Resources Technical Publication, U.S. Department of Interior, Bureau of Reclamation.

Rogers, D.C. and J. Goussard (1998, January/February). Canal control algorithms currently in use. *Journal of Irrigation and Drainage Engineering 124*(1), 11–15.

Ruiz-Carmona, V. and J. RamirezLuna (1998, October 11–14). Predictive control in irrigation canal operation. *IEEE International Conference on Systems, Man and Cybernetics (SMC'98) 1*, 3897–3901.

Ruiz-Carmona, V.M., A.J. Clemmens, and J. Schuurmans (1998, January/February). Canal control algorithms formulations. *Journal of Irrigation and Drainage Engineering 124*(1), 31–39.

Sabet, M., J. Coe, H. Ramirez, and D. Ford (1985). Optimal operation of california aqueduct. *Journal of Water Ressources Planning and Management 111*(2), 222–237.

Sanfilippo, F. (1997, Octobre). *Application de l'automatique multivariable à la régulation des canaux. Comparaison avec une approche monovariable*. Ph. D. thesis, Université Claude Bernard-Lyon 1. 197 p.

Sawadogo, S. (1992). *Modélisation, commande prédictive et supervision d'un système d'irrigation*. Ph. D. thesis, LAAS-CNRS Toulouse. 152 p.

Schaalje, M. and D.Manz (1993). Ann controller. Personal communication.

Schuurmans, J. (1992). Controller design for a regional downstream controlled canal. Master's thesis, Delft University of Technology, Laboratory for Measurement and Control. 100 p.

Schuurmans, J. (1997). *Control of water levels in open-channels*. Ph. D thesis, ISBN 90-9010995-1, Delft University of Technology.

Schuurmans, J., O. Bosgra, and R. Brouwer (1995, September). Open-channel flow model approximation for controller design. *Applied Mathematical Modelling 19*, 525–530.

Seatzu, C. (2000). Design of decentralized constant-volume controllers for open-channels by solving a least squares problem. *International Journal of Systems Science 31*(6), 759–770.

Seatzu, C., A. Giua, and G. Usai (1998, October 11–14). Decentralized volume control of open-channels using H_2 norm minimization. *IEEE International Conference on Systems, Man and Cybernetics (SMC'98), San Diego, California 1*, 3891–3896.

Shand, M. (1971). *Automatic downstream control systems for irrigation canals*. Ph.D. thesis, Université de Californie, Berkeley. 159 p.

Skogestad, S. and I. Postlethwaite (1998, November). *Multivariable feedback control, Analysis and Design*. Baffins lane, Chichester, West Sussex PO19 1UD, England: John Wiley and Sons Ltd.

Soleri Guitart, J. (2003, 24 Octobre). *Contribucio a léstudi del control de canals de reg per mitja de metodes numerics de programacio no lineal. Aplicacio al calcul de les trajectories de referencia*. Ph. D. thesis, UPC.

Stringam, B. and G. Merkley (1997a, April 22–24). Field application of a fuzzy controller for an irrigation canal in roosevelt utah. *International Workshop on the Regulation of Irrigation Canals: State of the Art of Research and Applications, RIC97, Marrakech (Morocco) April 22–24*, 349–354.

Stringam, B. and G. Merkley (1997b, April 22–24). Fuzzy controller simulation for local downstream water level control in canals. *International Workshop on the Regulation of Irrigation Canals: State of the Art of Research and Applications, RIC97, Marrakech (Morocco) April 22–24*, 342–348.

Tadeo, F., T. Alvarez, F. del Valle, and R. Rivas (2001, September). Decentralized constrained predictive control of irrigation canals using on-line linearized models. *SAAEI 2001 1*, 5. 5 p.

Tomicic, B. (1989). *A general optimization module for real-time control of surface water resources*. Ph. D. thesis, The International Institute for Hydraulic and Environmental Engineering, Delft University of Technology. 174 p.

Toudeft, A. (1996, November 12–13). Combining neural and adaptive controllers for a non-minimum phase varying time delay system. In *Journées Hispano-Françaises: Systèmes intelligents et contrôle avancé*, Barcelona (Spain), Université Polytechnique de Catalogne.

Voron, B. and A. Bouillot (1997, April 22–24). Application of the fuzzy set theory to the control of a large canal. *International Workshop on the Regulation of Irrigation Canals: State of the Art of Research and Applications, RIC97 April 22–24*, 317–331.

Wahlin, B.T. (2004). Performance of model predictive control on asce test canal 1. *Journal of Irrigation and Drainage Engineering*, 227–238.

Wahlin, B.T. and A.J. Clemmens (2002, November/Decembe). Performance of historic downstream canal control algorithms on asce test canal 1. *Journal of Irrigation and Drainage Engineering 128*(6), 365–375.

Weyer, E. (2000, June). System identification of an open water channel. In *IFAC Symposium on system identification*, Santa Barbara, pp. 6.

Weyer, E. (2001). System identification of an open water channel. *Control engineering practice 9*(12), 1289–1299.

Weyer, E. (2002). Decentralised PI control of an open water channel. In *15th IFAC World Congress, Barcelona, Spain*.

Weyer, E. (2003). Linear quadratic control of an irrigation channel. In *Proceedings of the IEEE Conference on Decision and Control, Hawaii, USA*, Volume 1, CSSIP, Department of Electrical Engineering, University of Melbourne, Parkville Vic. 3010, Australia, pp.750–755.

Xu, C. and G. Sallet (1999). Proportional and integral regulation of irrigation canal systems governed by the St Venant equation. In *14th Triennal World Congress, IFAC 99*, Volume E-2c-10-2, Beijing, China, pp.147–152.

Zaccarian, L., Y. Li, E. Weyer, M. Cantoni, and A.R. Teel (2007, February). Anti-windup for marginally stable plants and its application to open water channel control systems. *Control Engineering Practice 15*(2), 261–272.

Zihui, L. and D. Manz (1992, October). Optimal operation of irrigation canal systems using nonlinear programming – dynamic simulation model. *CEMAGREF-IIMI International Workshop, Montpellier 1*, 297–306.

Zimbelman, D. D.e. (1987, 28–30 July). *Planning, operation, rehabilitation and automation of irrigation water delivery systems*. Proceedings of a symposium ASCE, Portland, Oregon, USA. New York: ASCE. 391 pp.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

A computer model for pipe flow irrigation problems

R. Aliod & C. González
Escuela Politécnica Superior, Huesca, Spain

ABSTRACT: The need for hydraulic simulation tools which deal adequately with the peculiarities of irrigation systems is the reason behind the setting up of a Research and Development Project in Spain, with the collaboration of various local and national organizations. Its aim is to develop advanced computational hydraulic software, GESTAR, specially adapted for the design, planning and management of pressurized irrigation networks. A summary is given of the most significant characteristics of GESTAR's recent versions. The software has been validated by extensive fieldwork and, since it first appeared, has been used in the successful design and management of numerous irrigation systems.

1 INTRODUCTION

Pressurized water distribution systems for irrigation and drinking water share many characteristics but also present substantially different features. The former show, compared to the latter:

1. A simpler branched topology with a sparser layout.
2. A reduced number of independent users but with highly intensive water demands.
3. A discontinuous demand pattern, different for each demand node.

The simple branched topology having only one supply point with a known hydraulic grade is the most common in these systems. For a long time this has made systematic use of analytical tools and hydraulic simulation, used extensively in other contexts, unnecessary. In contrast to this, minimum cost sizing techniques, which are especially suited to branched systems, are extensively used in irrigation projects to the point where a design is considered to be complete once the conduits are sized.

However, it is essential to make use of general hydraulic analysis methodology in the design stage of irrigation systems in both the general distribution network and the network within individual plots. This recourse to hydraulic analysis methodology enables ever more frequent situations to be tackled, such as: the existence of numerous points within the system with a known hydraulic grade (multiple simultaneous feed, intermediate regulating tanks, pressure dependent emitters, leakage simulation, etc.), the presence of loops in the network (redundant parallel pipes in the network's supply, the interconnection of branches to correct shortages or equalize pressure) and the addition of automatic devices to regulate pressure and flow.

In addition, the use of hydraulic models in the context of pressurized irrigation networks permits the introduction of analytical instruments in the planning stages as well as in the day to day running of the network, whether it be for routine or exceptional circumstances.

In this way a wide range of problems, which up until now were either not considered or dealt with by means of over-simplified calculations, can be diagnosed and solved satisfactorily.

The use of predictive hydraulic tools in pressurized networks will probably be more effective and predictions more accurate in irrigation systems rather than in drinking water networks, where they were first developed, given the relative simplicity of these systems. In such cases, in which low uncertainties are present, detailed hydraulic modelling can provide accurate predictions, even

without any model calibration, supplying an effective interactive decision tool in the management of day to day use or complex scenarios (overdemand, failures, locating obstructions and leaks, multiple sources, etc.).

However, computer assisted modelling tools originally developed for drinking water supply systems are not necessarily suitable for satisfying the design demands and hydraulic management of irrigation systems, nor for reproducing the particular behaviour of such systems.

Due to the lack of software specifically applicable to the context of irrigation systems, and the absence of well documented field experiences, the Agriculture Department of the Regional Government of Aragón (Spain), in co-ordination with Regional (CONSID) and National (CICYT) research agencies, has promoted several applied research projects since 1996, whose aims have been:

1. to develop an advanced computational hydraulic software tool, GESTAR, specially adapted for the optimum design, hydraulic analysis, planning and management of pressurized irrigation networks.
2. to demonstrate in practice the benefits of using such tools from the early stages of engineering to daily operations.
3. to evaluate and improve the code capabilities for engineering and management through experiments and systematic validation in applications to real problems.

The major achievement of the project, with respect to R&D software, is GESTAR, a public domain software package for Windows 95/98/XP O.S. specifically conceived for the design and analysis of stationary hydraulic simulation in generalised pressurized irrigation networks. GESTAR (in Spanish) can be downloaded from the website www.gestarcad.com.

Extensions and improvements have been added continuously since the early releases in 1998.

Due to its unique characteristics GESTAR has become the reference engineering tool, widely used in the context of the extensive Spanish National Plans for Irrigation Development and Streamlining.

More than 1000 private or corporation users are registered, and the Zaragoza University research group supporting the code has worked on the design and modelling of networks covering more than 130,000 ha.

At this moment the GESTAR integrated package, interactive help and manuals are only available in Spanish, which limits its widespread diffusion overseas, but hopefully it will be translated to English in the future.

This paper summarises the description of the main tools available in the GESTAR package and some of the advantages of the computational techniques therein.

2 GESTAR 2007 GENERAL DESCRIPTION

2.1 *GESTAR User Interface*

GESTAR 2007 consists of two computational main modules, NETCAL and OPTI linked by means of a Graphical User Interface (GUI) (Fig. 1). The GESTAR GUI not only supports a fully graphic set of input/output operations and tools, but incorporates numerous supplementary capabilities, satellite modules, auxiliary tools and utilities.

The GUI supplies sufficiently efficient performances for stand-alone execution, and offers simple integration with CAD/GIS standard packages by means of Visual Basic for Applications support and MS-ACCESS database exchange formats.

Network loading, modification of parameters, handling options, presentation and analysis of results are all carried out in GESTAR by means of icons, dialogue boxes and menus which are all extremely user friendly (Fig. 2). Its intuitive nature, which requires no complex preliminary study, permits all types of user, from project managers to users of irrigation systems, to become familiar with the program very rapidly. GESTAR shares current software philosophy in that it aims to be

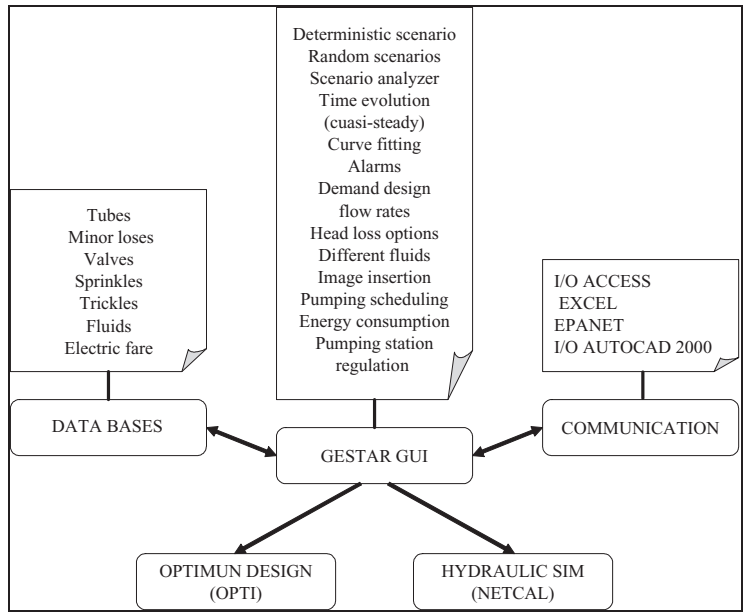


Figure 1. GESTAR structure.

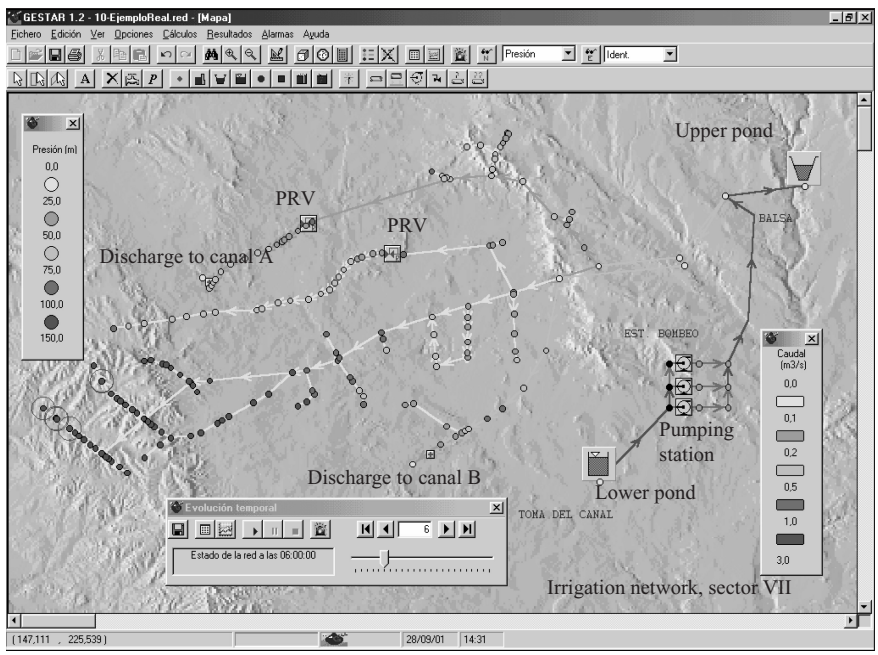


Figure 2. GESTAR screen for a typical irrigation network.

as user friendly as possible and it includes verification of data entered, automatic checking for compatibility of boundary conditions, auxiliary utilities, and user defined default values. There is also a wide range of tools which can be used for the creation, manipulation and modification of drawings; numerical and colour coded visualisation of data and results as well as a whole series

of options, user defined defaults and utilities. Different kind of tables, filters, graphics, numerical listing and MS-Office documents can be generated in the context of the GUI.

2.2 *GESTAR main computational routines*

OPTI is a MS Visual Basic code dedicated to optimum design of branched networks in a given layout, the design of flow rates in lines and the required minimum pressure at nodes. OPTI implements an optimization algorithm for series of conducts which are a combination of continuous Lagrange method and discontinuous Laybe (Laybe, 1988)–type method (González & Aliod, 2003). OPTI provides an effective solution for minimum investment in the network tubes or total annual expenses (including energy cost when pumping stations are present), with a very close approach to absolute minimum cost, spending only a few seconds of CPU time.

NETCAL is GESTAR's nucleus for quasi-steady hydraulic simulation of generalized pressurized networks. Fortran 90 coded, NETCAL incorporates diverse innovations related to the treatment of low resistance elements (LRE), inverse analysis and pressure dependent emitters lines. It also provides refinements in numerical techniques which increase calculation speed, guarantee convergence and realistically reproduce the hydraulic behaviour of equipment encountered in pressurized irrigation systems. This allows it to overcome some limitations and problems found in other computational tools (like EPANET hydraulic computation engine, a code for drinking water networks hydraulic and water quality simulation by Rossman, 1993) when they face the unique characteristics of irrigation networks.

Apart from the classical types of nodes and elements usually found in codes for drinking water networks (demand nodes, reservoirs, tanks, pipes, pumps, regulating valves, check valves, stroke valves, etc.), new kinds of combined elements and innovative treatments for some of them are introduced to account for the idiosyncrasy of agriculture equipment and mixed (looped-branched) topologies.

The wide range of different kinds of elements and operational nodes introduced make accurate modelling feasible, not only of the collective distribution system, but also networks of sprinkler and drip water application in irrigated plots.

Dedicated generalized modelling of any kind of point emitters (sprinklers, free discharges, leaks, etc.) and sophisticated modelling for conduits with continuous pressure dependent fluid emitters (useful for modelling trickle irrigation, porous tubing and distributed leaks) are included, for those purposes.

2.3 *GESTAR utilities*

Scenario generation tools allow the easy configuration of deterministic and random demand states. Any specific irrigation demand configuration, with a certain distribution of open and closed hydrants, can be established by means of time dependent tables of state, by mouse-driven point and click open/close orders, and by means of logical control queries.

Random scenarios can be generated by giving the percentage of node demands to be opened or the percentage of total flow rate demanded. In random scenarios, hydrants can be optionally considered to be opened with equal or different probabilities. In the latter case individual probabilities are computed in terms of the hydrant nominal flow rate, irrigated surface and agronomic parameters such as the design's daily irrigation requirement and the network efficiency.

Automatically-nested multiple random simulation of opened hydrants can be programmed, indicating the number of trials for each demand percentage assumed. Statistical analysis is then performed on the results, for each variable at any node or element, giving average, root mean square, maximum and minimum values.

Quasi-steady time evolutions (extended period simulation) can be performed, introducing time operational schedules for hydrants and pumps and logical operations conditioned by the state of components and levels of node and link variables.

What ever the kind of scenario analyzed, (deterministic or random, instantaneous or time dependent) alarms can be defined for every kind of variable, detecting the failure of appropriate flow conditions. Visual and detailed text report alarms are obtained. Alarm scenarios can be kept individually to inspect the origin, causes and repeatability of service failures.

Several other satellite software applications can be launched from the GUI:

Module for computation of accumulated or stochastic (Clement & Galand, 1986) flow rates: supplies the lines design flow rates to be used in the OPTI optimization routine.

Module for pumping station schedule curve: by means of multiple random scenario analysis, the routine automatically creates the recommended envelope curves of the pressure to supply according to demand flow rate.

Module for estimation of Probability Density Function of the pumping station flow rates: Computes the stochastic frequency of flow rate in every monthly period along the irrigation campaign after specifying the monthly average needs.

Module for computing the energy consumed at pumping station in a given period: integrates the product of the energy demanded (depending on the pump head and efficiency curves, type of regulation and pumping station schedule curve) by the PDF over all possible flow rates.

2.4 *Software communication*

Complete and flexible data bases for commercial pipes, stroke valves, minor losses, emitters, trickle irrigation pipes, fluid properties, and even electric fares are linked to the elements edition forms. They are open to users and can be updated and/or expanded with the data base tools supplied.

Network construction parameters, scenario specifications and computational results can be exported/imported in data base formats MS-ACCESS, allowing complex search, filtering or data manipulation and full input/output exchange with any comercial CAD or GIS tool.

Node and pipe data can be exported to EPANET v.1.e, input file, and data exchange with MS-EXCEL.

In the past a beta GESTAR version for AUTOCAD 2000 was available, but at this moment this version is no longer offered. Nevertheless tools exist to convert AUTOCAD drawings to GESTAR network files and vice versa.

3 GESTAR HYDRAULIC SIMULATION MODULE

3.1 *Computational techniques for analysis of pressurized stationary networks*

In order to consider any kind of branched, nested or mixed network the hydraulic computational kernel of GESTAR, NETCAL, makes use of a suitable adaptation of the nodal type analysis matrix methodology, Todini & Pilati (1987), for steady flows.

Water distribution networks topology is defined in terms of links (conducts and other hydraulic elements i.e. pumps, valves. . .) limited at extremes by nodes connected to other links or to boundary conditions (Fig. 3). The flow rate of all elements, Q_k, and the Head, H_j, and water demand, C_j, at any network node are the fundamental variables that need to be defined to know the state of the network. At first instance either H_j or C_j is a datum known in every node.

Any link – connecting initial node j and final node i- has a non-linear hydraulic constitutive equation that relate head-loss, $H_j - H_i$, and flow rate, Q_k for stationary incompressible flow regimes.

$$H_j - H_i = F_k(Q_k)Q_k^2 \tag{1}$$

For example, in a pipe element of constant diameter the exponential type head loss formulae yields to:

$$\Delta H = R\,|Q|^{n-2}Q^2; F(Q) = R\,|Q|^{n-2} \tag{2}$$

If the Darcy-Weisbach head loss expression is used:

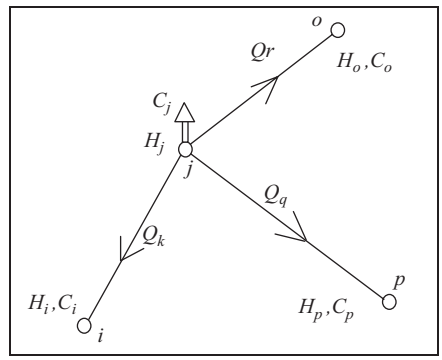


Figure 3. Definition of node and link variables.

$$\Delta H = f(Re, \varepsilon)\ \frac{L}{2gA^2D}\ Q^2; F(Q) = f(Re, \varepsilon)\ \frac{L}{2gA^2D} \tag{3}$$

Element constitutive equation (1) can be expressed in resistance like pseudo-linear form:

$$H_j - H_i = KR_k \cdot Q_k;\ KR_k = F_k(Q_k)Q_k \tag{4}$$

or in an alternative conductivity form:

$$Q_k = KC_k \lfloor H_j - H_i \rfloor; KC_k = \frac{1}{KR_k} \tag{5}$$

In Low Resistance Elements (LRE), *KR* values are very small, and thus *KC* goes to infinite.

Being *ne* the number of elements at network, and *nnd* the number of nodes with unknown, H_j, head, the topological incidence *ne* x *nnd* matrix A_{12}, with components $A_{12}(k,j)$ is defined as:

– $+1$ if the positive flow rate of the element k comes from the node j.
– 0 If the element k and the node j are not connected.
– -1 if the positive flow rate of the element k goes to the node j.

The transposed matrice $\boldsymbol{A_{12}^T}$ and $\boldsymbol{A_{10}^T}$ are denoted $\boldsymbol{A_{21}}$ and $\boldsymbol{A_{01}}$.

Being *nnc* the nodes with known head, H_{0j}, $\boldsymbol{A_{10}}$ is a *ne x nnc* matrix defined identical to $\boldsymbol{A_{12}}$ but involving the *nnc* unknown head nodes.

Finally the diagonal resistance link matrix, $\boldsymbol{A_{11}}$ is defined

$$\boldsymbol{A_{11}} = \begin{pmatrix} -KR_1 & & & \\ & -KR_2 & & \\ & & \ddots & \\ & & & -KR_{ne} \end{pmatrix} \tag{6}$$

with the above definitions, the mass flow conservation equations at each node and the constitutive equation of every element can be formally written.

$$\begin{pmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{Q} \\ \boldsymbol{H_*} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{A_{10}H_0} \\ -\boldsymbol{C_0} \end{pmatrix} \tag{7}$$

where $\boldsymbol{H}_*$ is the *nnd* vector containing the node unknown heads, $\boldsymbol{H_0}$ is the *nnc* vector defined with the node known heads, $\boldsymbol{Q}$ is the *ne* vector containing the unknown flow rates at every element

and $\boldsymbol{C_0}$ is the *nnd* vector of node known demands. The *ne*+*nnd* non linear equation system (7) is independent as long as $\boldsymbol{H_0}$ is not empty. The solution of the "hybrid" vector ($\boldsymbol{H_*}$, $\boldsymbol{Q}$) of *ne*+*nnd* unknowns determines the complete hydraulic state of the network. The *nnc* vector $\boldsymbol{C_*}$ of unknown node demands at nodes $\boldsymbol{H_0}$ is decoupled from (7) and it can be explicitly found after (7) solution applying the nodal continuity equations (8).

$$\boldsymbol{A_{01}} \cdot \boldsymbol{Q} = -\boldsymbol{C_*} \tag{8}$$

The non linear system (7) can be efficiently solved by means of Newton-Raphson iterative method, yielding to:

$$\begin{pmatrix} \boldsymbol{NA_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\delta Q} \\ \boldsymbol{\delta H_*} \end{pmatrix} = \begin{pmatrix} \boldsymbol{dE} \\ \boldsymbol{dq} \end{pmatrix} \tag{9}$$

where $\boldsymbol{N}$ matrix is constructed as

$$\boldsymbol{N} = \begin{pmatrix} 2-P_1 & & & \\ & 2-P_2 & & \\ & & \ddots & \\ & & & 2-P_{ne} \end{pmatrix}$$

And the "laminar index", P_k (P is 1 for laminar flow, and 0 for fully turbulent flow, in exponential head loss formulae $P = 2 - n$, so $2 - P = n$), for k element with constitutive equation (1) is defined as:

$$P_k = -\frac{Q_k}{F_k}\frac{dF_k}{dQ_k}$$

The equation system solution (9) supply the corrections for element flow rate and node head vectors, neccesary to compute values at iteration $k+1$ knowing the values at iteration k:

$$\boldsymbol{\delta Q} = \boldsymbol{Q}^{k+1} - \boldsymbol{Q}^k \quad ; \quad \boldsymbol{\delta H_*} = \boldsymbol{H_*}^{k+1} - \boldsymbol{H_*}^k$$

and $\boldsymbol{dE}$ and $\boldsymbol{dq}$ are the residues ("error") of the equation system (7) at the iteration level k

$$\begin{aligned} \boldsymbol{dE} &= -\left(\boldsymbol{A_{11}Q}^k + \boldsymbol{A_{12}H_*}^k + \boldsymbol{A_{10}H_0}\right) \\ \boldsymbol{dq} &= -\left(\boldsymbol{A_{21}Q}^k + \boldsymbol{C_0}\right) \end{aligned}$$

Despite the convenient mathematical properties of the system (7) where coefficient matrix are simple and symmetric, LRE does not induce any singularity and all the hydraulic "hybrid unknowns" are solved simultaneously and coupled, different families of procedures have been deduced from (8,9). These alternative methods were forced to be adopted in past, due to computing limitations, in order to minimize the number of equations to be solved, since the expression (9) involves the maximum possible. Loop Analysis Methods (LAP) or Nodal Analysis Methods (NAM), each one with different subvariants, can be obtained by manipulation of (8,9).

To deduce the Nodal Analysis methods, first the $\boldsymbol{dQ}$ vector form is extracted (9)

$$\boldsymbol{\delta Q} = -(\boldsymbol{NA_{11}})^{-1}(\boldsymbol{A_{11}Q}^k + \boldsymbol{A_{12}H_*}^k + \boldsymbol{A_{10}H_0} + \boldsymbol{A_{12}\delta H_*}) \tag{10}$$

Then (10) is introduced in the (9) second row block, yielding, after regrouping of terms to:

$$\begin{aligned} -\boldsymbol{A_{21}}(\boldsymbol{NA_{11}})^{-1}\boldsymbol{A_{12}}\,\boldsymbol{\delta H_*} = \boldsymbol{A_{21}}\,&(\boldsymbol{NA_{11}})^{-1}(\boldsymbol{A_{11}Q}^k + \boldsymbol{A_{12}}\,\boldsymbol{H_*}^k + \boldsymbol{A_{10}}\,\boldsymbol{H_0}) \\ &- (\boldsymbol{A_{21}Q}^k + \boldsymbol{C_0}) \end{aligned} \tag{11}$$

In the simpler NAM versions the linearized constitutive equations at iteration k are supposed to be fulfilled, so:

$$\boldsymbol{A}_{11}^{k}\boldsymbol{Q}^{k} + \boldsymbol{A}_{12}\boldsymbol{H}_{*}^{k} + \boldsymbol{A}_{10}\boldsymbol{H}_{0} = \boldsymbol{0} \tag{12}$$

And the system of equations to solve correction vector $\boldsymbol{\delta H}_{*}$ is:

$$-\boldsymbol{A}_{21}(\boldsymbol{NA}_{11})^{-1}\boldsymbol{A}_{12}\delta H_{*} = -(\boldsymbol{A}_{21}\boldsymbol{Q}^{k} + \boldsymbol{C}_{0}) \tag{13}$$

Let us note that in any NAM matrix $(\boldsymbol{NA}_{11})^{-1}$ is a diagonal matrix, with the following structure.

$$(\boldsymbol{NA}_{11})^{-1} = \begin{pmatrix} \frac{-1}{(2-P_1)KR_1} & & & \\ & \frac{-1}{(2-P_2)KR_2} & & \\ & & \ddots & \\ & & & \frac{-1}{(2-P_{ne})KR_{ne}} \end{pmatrix}$$

That matrix is ill conditioned when low resistance ($KR = > 0$) elements are present, The assembling of the eq. (11) *LHS* tangent matrix

$$\boldsymbol{AD} = -\boldsymbol{A}_{21}(\boldsymbol{NA}_{11})^{-1}\boldsymbol{A}_{12}$$

is straightforward.

Denoting:

$$KD = \frac{1}{(2-P)KR}$$

If link m connect the nodes r *and* s the matrix non diagonal components are $AD_{rs} = AD_{sr} = -KD_m$. If nodes r s are disconnected $AD_{rs} = AD_{sr}. = 0$. The diagonal elements AD_{ss} consist of the sum of KD for all elements which are connected to s node (13), as sketched in:

$$\boldsymbol{AD} = \begin{pmatrix} & \vdots & & \vdots & & \vdots & \\ \cdots & \sum_{r} + KD & \cdots & 0 & \cdots & -KD_m & \cdots \\ & \vdots & & \vdots & & \vdots & \\ \cdots & \cdots & \cdots & +\sum_{i} + KD & \cdots & \cdots & \cdots \\ & \vdots & & \vdots & & \vdots & \\ \cdots & -KD_m & \cdots & 0 & \cdots & \sum_{s} + KD & \cdots \\ & \vdots & & \vdots & & \vdots & \end{pmatrix} \tag{14}$$

The method described in Todini & Pilati (1987), claimed to be implemented in EPANET, and referred to by the authors as a "hybrid method", is not in fact a "hybrid method" but a branch of the NAM family because computing expression (11) involves the nodal matrix $(\boldsymbol{NA}_{11})^{-1}$. Complete equations (11) differ from the more classical nodal method equations (13) in the RHS structure, where it can be shown that the RHS at (11) represents a relaxation of the nodal flow

balance residues $\boldsymbol{dq}$ using values of iteration k , $k+1$,weighted with $1/(2-\mathrm{P})$ factor. On the one hand, the nodal balance residue relaxation slows down the convergence rate, but on the other hand it introduces a positive damping effect of instabilities that can arise in the iteration process. In any case, the loss of convergence and inconsistencies attributed to the presence of low resistance elements in a network still are present at (11). Despite the tricky truncation limit proposed by Todini & Pilati (1987), when $(\boldsymbol{NA_{11}})^{-1}$ elements tend to grow, the formulation is not satisfactory.

3.2 *Low resistance links treatment*

Since *KR* elements coefficients depend not only on dimensional characteristics but also converged flow rates, (or pressure head loss) elements with very small *KR* (or head loss) values can not be identified in advance, inducing in the past unpredictable failures in the convergence of the iterative non-linear solvers. GESTAR development had to pay special attention to such low resistance elements, since zero head-loss conditions are systematically found in any branched irrigation network from the last opened demand node in a branch, to its end. As NAM is rather convenient thanks to the reduced number of equations to be solved, simplicity of matrix assembling, and ability to introduce any kind of hydraulic elements, appropriate strategies to overcome the LRE should be included.

The technique of "conditional decoupling" of hydraulic constitutive equations proposed by Campos (1993) for LRE was adopted. In this strategy the k LRE (or zero flow line) constitutive equations (1) are kept in the original presentation (4) and added to the nodal system (11), maintaining the line flow rate Q_k as an extra variable to balance the number of equations and unknowns. Although this procedure increase the system dimension (11) by one for each low resistance link, it has the advantage of being able to work quite robustly and accurate with either null or infinite resistance conduits. Conditional decoupling recovers the "hybrid method" scheme, but only for the necessary LRE and, in the limit, if all the elements are decoupled, the original (9) fully "hybrid method" is found.

As far as a low resistance element k exists connecting i,j nodes, the updated product of tangent matrix and vector of corrections for unknowns on the *LHS* in (11) is:

$$\begin{pmatrix} & \vdots & & \vdots & & \vdots \\ \cdots & +0 & \cdots & +0 & \cdots & +1 \\ & \vdots & & \vdots & & \vdots \\ \cdots & +0 & \cdots & +0 & \cdots & -1 \\ & \vdots & & \vdots & & \vdots \\ \cdots & +1 & \cdots & -1 & \cdots & -\frac{1}{KD_k} \end{pmatrix} \begin{pmatrix} \vdots \\ \delta H_j \\ \vdots \\ \delta H_i \\ \vdots \\ \delta Q_k \end{pmatrix} \tag{15}$$

Since the low resistance condition can appear at any stage of the iterative solution procedure, maximum threshold value *KDo* is compared with *KD* values of every element at the beginning of each iteration, reassembling properly the equation system and tangent matrix when any link becomes catalogued as LRE.

3.3 *Flow rate computing at emitters and nodes with known head*

In the systems with a great number of nodes with imposed head, like the networks with many emitters present in irrigation systems, it may be convenient to perform the computation of their resulting consumption in the same step and coupled with the rest of the unknowns in order to gain in consistency and accuracy. It means including the equations (8) in the system formulation, being $\boldsymbol{C}_*$ new unknowns.

The introduction at node k of demand C_k^* as unknowns implies a extended product of tangent matrix on the LHS of (11), with the following structure:

$$\begin{pmatrix} & \vdots & & \vdots & & \vdots & \\ \cdots & \sum_r +KD_r & \cdots & 0 & \cdots & -KD_m & \cdots \\ & \vdots & & \vdots & & \vdots & \\ \cdots & \cdots & \cdots & +1 & \cdots & \cdots & \cdots \\ & \vdots & & \vdots & & \vdots & \\ \cdots & -KD_m & \cdots & 0 & \cdots & \sum_s +KD_s & \cdots \\ & \vdots & & \vdots & & \vdots & \end{pmatrix} \begin{pmatrix} \vdots \\ \delta H_i \\ \vdots \\ \delta C_k^* \\ \vdots \\ \delta H_j \\ \vdots \end{pmatrix} \tag{16}$$

Initial symmetry properties of matrices (14) (15) are no more sustained in (16), forcing the use of non symmetric solvers and operators that require some more memory and computing time than the symmetric alternatives.

Nevertheless, other convenient modelling strategies and options also induce lack of symmetry in the equation systems resulting:

a) Regulating valves' constitutive equations incorporated in the equations system (11) and simultaneously solved with the rest of the nodal equations, improve the stability and ability to find solution in complex configurations, but are associated with non symmetric matrices.
b) Inverse analysis techniques defined in paragraph *3.4* cause non symmetry at the equations system to be solved.

All these are added reasons overcome the inconvenience of dealing with non symmetric matrices. Therefore NETCAL is powered with matrix direct inversion routines and matrix operation libraries specially adapted for non symmetric matrix, and optimized for sparse matrices, and compact storage, Duff et al. (1986). The object-oriented capabilities of FORTAN 90 language are exploited in the code in order to define dynamic data structures appropriate for compact dealing with systems of arbitrary dimensions.

The non symmetric matrix operation and inversion routines, the augmented number of equations when LRE are decoupled, and the matrix reassembling at every iteration when necessary, increase the computational time in comparison to other alternatives, but in any case it stays at acceptable levels, i.e. 0.3 s CPU time in standard 2006 INTEL M 750 1,86 Ghz processor for a 2600 equations system in a fully looped network requiring 19 iterations.

The NETCAL code results have been systematically benchmarked and compared with other reference network solvers, offering quite similar results under identical configuration and boundary conditions.

Figure 4 presents correlation of NETCAL pressure predictions at nodes in respect EPANET for the same scenario of collective distribution irrigation network shown in Figure 2. Identical pressures can be found with irrelevant maximum discrepancies of 0.044 m.

3.4 *Inverse analysis*

Implementation of inverse analysis capabilities, which allow the direct computation of unknown control parameters by solving a modified system of equations, was proved to be suitable both for designing individual components and for controlling and regulating day to day operation of the network in order to obtain optimum service.

Whereas deterministic inverse design tools have been documented for loop analysis techniques of networks (Boulos & Wood 1990), few references are found in the context of nodal analysis. Therefore, inverse design strategies have been researched and implemented.

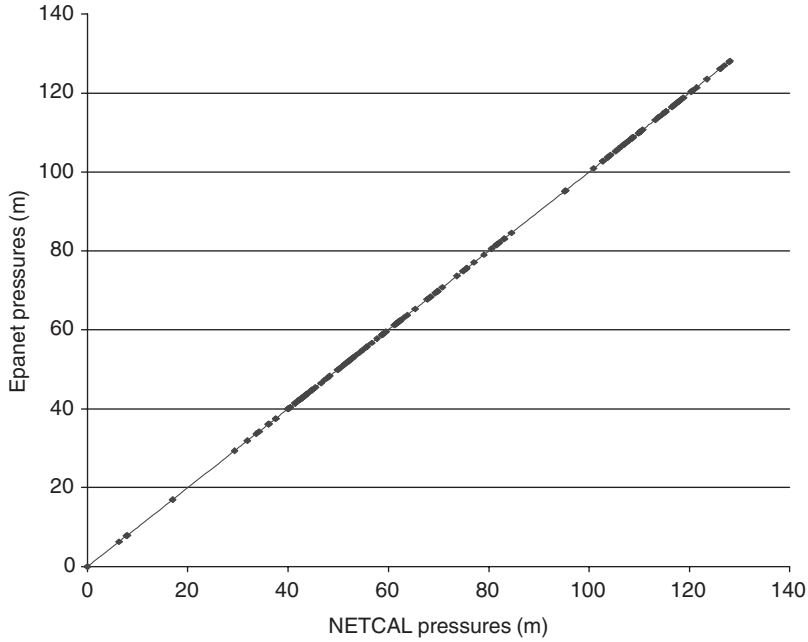


Figure 4. GESTAR vs EPANET comparison of pressure predictions for network in Figure 3.

Most of the usual deterministic operational goals to be fulfilled in a network can be translated in such a way that both hydraulic grade, *Ho*, and demand, *Co*, are imposed at some "target" nodes (Fig. 1). Each of such duplicated node conditions requires the introduction of a new parameter as a freedom degree to balance unknowns and equations at (9,11).

Depending on the nature of the freedom degree available to fit functional objectives, it will be associated to a single node, p, or a single link, q, or to a common parameter, Y, for a set of lines. In the case of conditions regulated by a single node, p, both H_p and C_p will be unknowns that can be located, after rearrangement of the system (7), at the $\boldsymbol{H}_*$ vector, whereas target related values, H_o and C_o, are placed on the vector $\boldsymbol{Co}$ in (7) RHS. If the regulation is performed by changing a parameter at a single link, q, the line q is then substituted by an unknown line flow rate, Q^q, with opposite signs allocated at the link extreme nodes. This transformation leads to an approach similar to that presented in the first case. Finally, when the fitting parameter is a property Y, associated to a set of lines (e.g., roughness, diameter), the implicit dependence of the solution on Y should be extracted from constitutive equations (1). Taking Y as a variable which replaces the unknown initially associated to the target node, H_o, in the correction equations of the Newton-Raphson solver for the node o, an extra term is obtained:

$$\left[\sum_j \sum_p \sum_\beta A_{i\beta} \frac{\partial K C^\beta}{\partial Y} \delta_{\beta p} A_{jp} H_j\right] \Delta Y \tag{17}$$

In (17) dY is the correction of the Y unknown and the rest of the term represents the coefficient (i, o) of the tangent matrix, which is accommodated to substitute dH_o by dY.

3.5 *Control device modeling*

The proper implementation of dives as Flow Control Valves (FCV) and Pressure Sustaining Valves (PSV) in a non-looped branch is quite a subtle matter since the simple modelling algorithms for

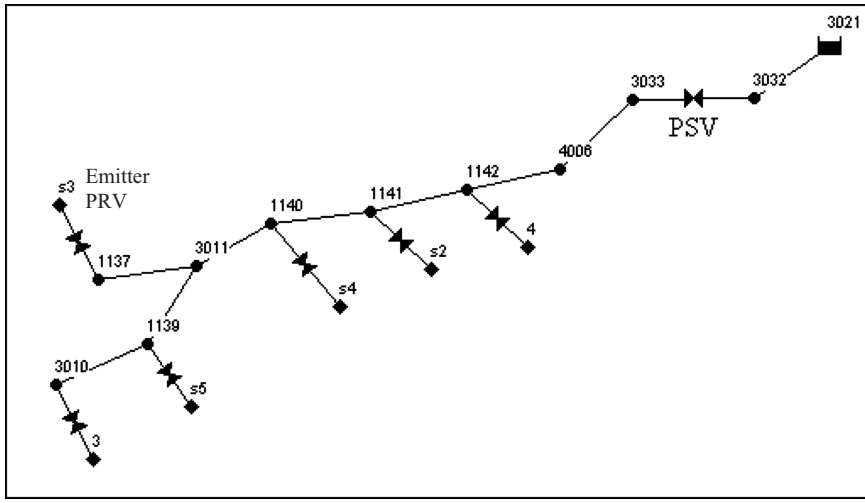


Figure 5. Study case for testing hydrants modeling.

looped lines fail (due to the isolation of the nodes from the reference pressure) when such devices reach an active regulation state in a branch where non-pressure dependent demand nodes are present downstream. Research, connected with the hydrants behaviour at low upstream pressure, has shown that the computational strategies implemented in GESTAR (where equations for the control devices are introduced and solved altogether and simultaneously with the rest of network variables) for reproducing the behaviour of PSV and FCV in their regulating actions seem to be more robust and ensure better convergence that other alternatives.

For example, irrigation hydrants, can either be working as a pressure independent flow demand or as pressure dependent emitters. This is referred to as "dual modelling". When the hydrants flow supply in the distribution network is controlled with pressure reducing valves (PRV) and flow control valves (FCV), and the pressure level at the hydrant is above the setting level, the node demand will be fixed by the FCV flow setting. If the pressure level goes below the PRV setting, the PRV will be fully opened and the demand at the node will be switched to pressure dependent behaviour characterised by a function

$$H = K_s \cdot C^N \tag{18}$$

Where C is the flow rate supplied to the farm when the hydraulic devices installed are fed with the pressure head H. Function (18) synthesises the flow response of the plot's inner irrigation network to the distribution network pressure. It can be easily obtained by means field experiments, in existing facilities, and/or by previous simulation with GESTAR of the plot's inner hydraulic network. K_s and N values are introduced optionally at the hydrant window definition.

In EPANET, such a hybrid node type is not available, but it can be indirectly configured (for steady states) substituting each hydrant junction by the upstream node of a PRV with an emitter connected to the downstream node of the PRV (Figure 5). This introduces in EPANET an extra node (emitter) and an element (PRV) to reproduce each "Hydrant" with hybrid regulating characteristics.

In the "hydrant" node (PRV plus emitter) modeled in EPANET, the EPANET hydraulic solver structure remains the same whatever range of pressure exists in the network, and coefficient K_{sp} and exponent N must be given exactly as (19) in order to achieve the imposed pressure independent demand when PRV is operating (network pressure exceeds setting level). Emitter characterization in EPANET for every emitter is given by:

$$Q = \text{Coeff} \cdot \Delta H^{1/N}; \text{Coeff} = K_{sp}^{-1/N} \tag{19}$$

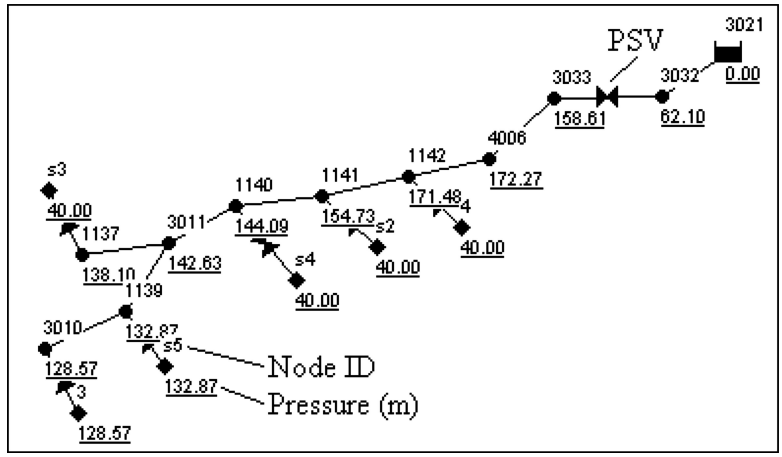


Figure 6. Unbalanced EPANET results for pressure heads (m) and node demands (l/s) after 200 iterations for the test case. Note that main line pressure heads are much larger than is possible.

Default value for exponent N in EPANET is 2.

"Hydrant" components in GESTAR have always converged to feasible solution. Nevertheless, in some cases EPANET has failed to find realistic solutions as the one documented below. In any case, when EPANET achieves proper convergence, GESTAR and EPANET solvers offer the same results.

One case of study in which GESTAR obtains proper results but EPANET is not able to reach a converged solution is offered in the Figure 5. This network is a non-looped branch extracted from another more complex network, where available total head at node 3021 is known (417.17 m).

In this configuration the PSV (with an upstream pressure setting of 62.1 m) should be operating, because full downstream node demands will reduce node 3032 pressure head below the setting.

In this configuration it has been observed that the introduction of the PSV in EPANET results in a non realistic flow distribution in the downstream "Hydrants". This non-convergence is detected verifying the state of the PSV, which is found being in fully open state, contradicting the upstream pressure setting, showing control action. Moreover, absurd pressure head values appear (Figure 6), that lead to a total head larger than available head at 3021 reservoir, and unbalanced flow rates.

Configuration values of case are shown in Appendix I, Tables 1, 2 and 3.

GESTAR obtains a coherent solution in which PSV is controlling the setting of the upstream pressure, and downstream hybrid nodes pressure heads are lower than 40 m (Figure 7). Consequently, hybrid nodes demands are lower than base demands, as shown in Figure 7.

The authors can not identify the causes of this unexpected misfunction of EPANET, but it may be due either to some bug, or related to insufficiencies in the PSV and PRV modelling at branched topologies.

Let us note that GESTAR introduces changes in the equation system structure when any valve or node modifies its state, but these changes do not take place in EPANET.

3.6 *Pressure dependent emission at trickle and porous pipes*

The plot's inner irrigation facilities are modelled mostly in terms of pressure dependent devices: sprinkle and trickle emitters. Both of them can be modelled as individual emitters, Setp by Step (SBS) method, using their particular pressure/flow rate characteristic curve, given by the manufacturer or fitted to experimental pressure/flow rate data.

Nevertheless, as the number of drip emitters is usually quite large, trickle tubing can be better modelled as continuous-like emission, in the same way that porous pipes or pipe distributed leaks

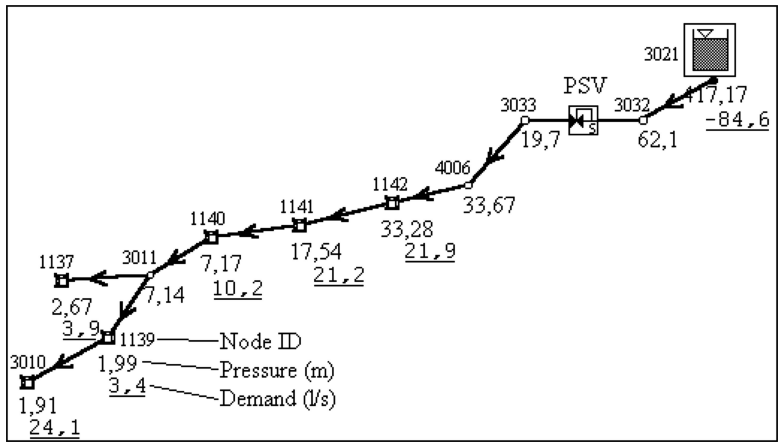


Figure 7. GESTAR converged (14 iterations) results. Pressure head (m) and calculated nodes demands (l/s).

in any water distribution system. Advanced modelling for pressure dependent continuous emission pipes have been implemented in GESTAR 2007 introducing innovative strategies that improve and generalize the integral- differential approach (Warrick & Yitayew, 1988).

The local emission of flow rate per unit lengh, q_p, in a position of a trickle tube, of constant section A and diameter D, with emitters spacing e, at a point where the pressure is p_p can be expresed:

$$q_p = \left(\frac{c}{e}\right) \cdot p_p^y$$

And the local tube velocity, v_p, is such that

$$A \left.\frac{dv_p}{dx}\right|_p = -q_p \tag{20}$$

The energy equation for a segment with slope s,

$$\frac{d}{dx}\left(p_p + \frac{v_p^2}{2g}\right) + \frac{f v_p^2}{2gD} + s = 0 \tag{21}$$

Combining the expressions (20) and (21) neglecting the kinetic energy term, (21) becomes:

$$\frac{d}{dX_p}\left(-\frac{dV_p}{dX_p}\right)^{\frac{1}{y}} + V_p^m + S_p = 0 \tag{22}$$

Where the dimensionless variables appearing in (22) are defined:

$$\begin{aligned} X_p &= \frac{x_p}{x_0}; V_p = \frac{v_p}{v};\ S_p = \frac{2gDs}{f_0 v^m};\ x_0^{\,1+\frac{1}{y}} \\ &= \frac{2^{1-\frac{2}{y}} \pi^{\frac{1}{y}} g D^{1+\frac{2}{y}} e^{\frac{1}{y}}}{c^{\frac{1}{y}} f_0 v^{m-\frac{1}{y}}} \end{aligned} \tag{23}$$

Being v the fluid velocity at the tube intake and m, f_o parameters related with the head losses formulation (Appendix II).

The integration of the equation (22) yields to

$$\left.\frac{dV}{dX}\right|_p = \frac{-1}{F(V_p)};\ X_p = \int_{V_p}^{1} F(V) dV \tag{24}$$

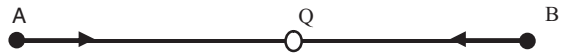


Figure 8. Trickle irrigation (or porous) tube fed by both extremes.

where

$$F(V_p) = \left((1+y) \cdot \left(\frac{V_p^{m+1}}{m+1} + SV_p \right) + C \right)^{-\frac{y}{y+1}} \tag{25}$$

Combining dimensionless velocity of (23) with (24)

$$\left. \frac{dv}{dx} \right|_p = \frac{-v}{x_0 F(V_p)} \tag{26}$$

This equation can be introduced in (20), and after regrouping of terms, the resulting condition that is satisfied at a location where pressure p_p is.

$$\frac{vA}{x_0 \left((1+y) \cdot \left(\frac{V_p^{m+1}}{m+1} + S \cdot V_p \right) + C \right)^{-\frac{y}{y+1}}} - \left(\frac{c}{e} \right) p_p^y = 0 \tag{27}$$

And the condition (24) applied to the end of the segment of length L gives:

$$\int_{V_f}^{1} \left((1+y) \cdot \left(\frac{V^{m+1}}{m+1} + S \cdot V \right) + C \right)^{-\frac{y}{y+1}} dV - \frac{L}{x_0} = 0 \tag{28}$$

By means of (27) and (28), Estrada (2000) developed and implemented at GESTAR a solution method (ECR, "elemento de consumo en ruta", on course demand element) for the complex system of coupled non linear equations that model a series of trickle (or porous) tubes composed by segments with different characteristics. Lines can be linked to regular network nodes and elements with several kind of boundary conditions. González & Aliod (2005) and González (2006) expanded the model capabilities to consider the more general situation where the tube is fed by both extremes A and B (Fig. 8), where pressure is given, with a null velocity (equilibrium) intermediate point, Q.

Applying pressure condition (24) at extremes A and B, expressions f_1, f_2 are obtained:

$$f_1 = \frac{v_A A}{x_{A,0} \left((1+y) \cdot \left(\frac{1}{m+1} + S_A \right) + C_A \right)^{-\frac{y}{y+1}}} - \left(\frac{c}{e} \right) p_A^y = 0$$

$$f_2 = \frac{v_B A}{x_{B,0} \left((1+y) \cdot \left(\frac{1}{m+1} + S_B \right) + C_B \right)^{-\frac{y}{y+1}}} - \left(\frac{c}{e} \right) p_B^y = 0$$

Applying equation (26) to segments AQ and BQ, with respective lengths, L_A, L- L_A, conditions f_3, f_4 are deduced:

$$f_3 = \int_0^1 \left((1+y) \cdot \left(\frac{V^{m+1}}{m+1} + S_A \cdot V \right) + C_A \right)^{-\frac{y}{y+1}} dV - \frac{L_A}{x_{A,0}} = 0$$

$$f_4 = \int_0^1 \left((1+y) \cdot \left(\frac{V^{m+1}}{m+1} + S_B \cdot V \right) + C_B \right)^{-\frac{y}{y+1}} dV - \frac{L - L_A}{x_{B,0}} = 0$$

Finally imposing the pressure continuity at Q point:

$$f_5 = \left(\frac{v_B A}{\frac{c}{e} x_{B,0} C_B^{-y/y+1}} \right)^{1/y} - \left(\frac{v_A A}{\frac{c}{e} x_{A,0} C_A^{-y/y+1}} \right)^{1/y} = 0$$

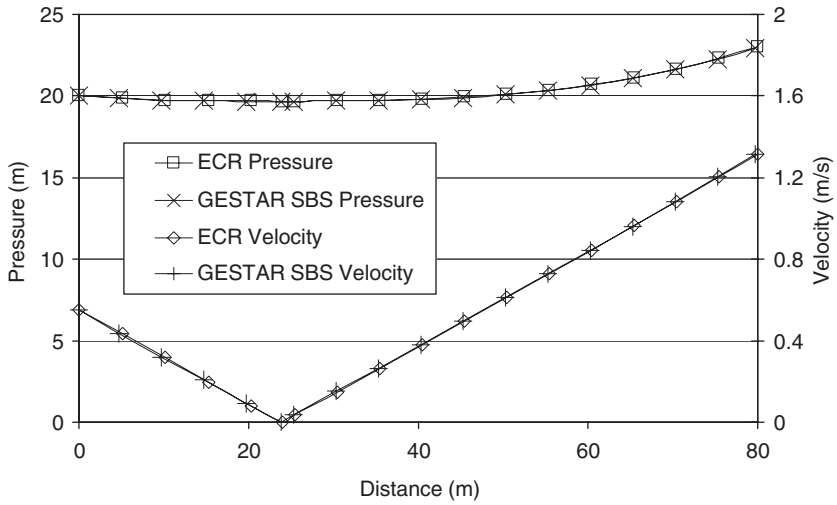


Figure 9. ECR and SBS GESTAR results for a 80 m trickle irrigation pipe.

The expressions $f_1, f_2\, f_3, f_4\, f_5$, configure a system of 5 eqs. well suited to obtain the solution on the unknowns ($v_A, v_B, C_A, C_B y L_A$). The equation systems is solved by means a Newton – Raphson algorithm.

Smart strategies of initialization of the iterative process have been introduced to obtain proper convergence, because results are sensitive initial guess values and pure numerical parameters C_A, C_B are not easy to estimate.

Once the system is solved the pressure and velocity distribution along of the porous conduit can be deduced.

The comparison of these integral-differential solution (ECR) with Step by Step (SBS) for multi point emitter simulation (trickle tubing) gives quite satisfactory results as shown in Figure 9 where the pressure and velocity distribution computed are shown. The case corresponds to a 80 m trickle conduit, with 14,4 mm inner diameter, and left and right extremes pressures imposed 20 m and 23 m respectively. Drip emitters are equally spaced 0.5 m, (160 emitters) with a emission characteristic curve:

$$q_p = 1,350 \cdot p_p^{0.539}, (q \text{ l/h}, p \text{ m})$$

Let us note that the despite the complex integral- differential modeling of pressure dependent emission of porous conduit with ECR technique the computing time for this single component is more than 60 times lower than the equivalent 160 emitters, 160 junctions and 320 pipe elements computation involved in the SBS method, and of course much more straightforward to configure in comparison with the lengthy set up of the individual emitters. Therefore ECR modelling for emitters tubes is quite suited to be integrated as a new specific kind of element to be linked in general networks computed by NETCAL hydraulic solver.

In the next example, the ECR approach is compared with SBS simulation, not only using GESTAR-NETCAL software, but also EPANET code, and RIEGOLOC (Rodrigo et al., 1992), a simple software tool to perform isolated drip pipe irrigation calculations. To allow include then RLOC predictions the drip irrigation tube is fed only by one side, as this is a restriction of RLOC limited capabilities. The emitters spacing is 0.5 m along 90 m long conduit of diameter 13,5 mm, with pressure-flow rate response given by:

$$q_p = 1.242 \cdot p_p^{0.487}, (q \text{ l/h}, p \text{ m})$$

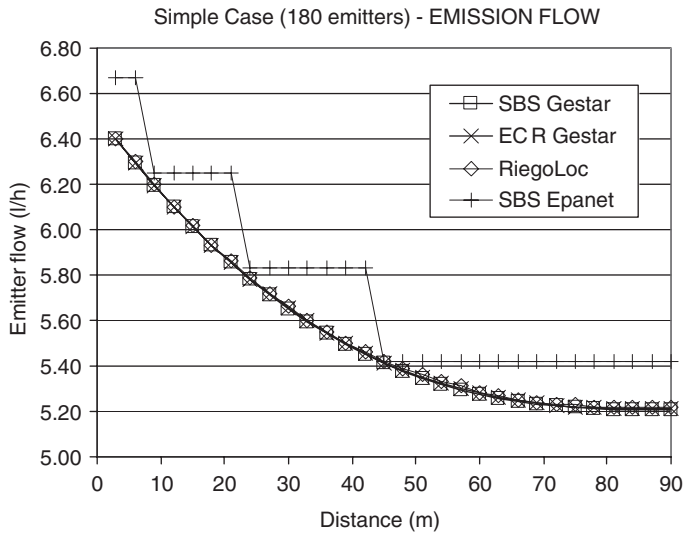


Figure 10. GESTAR, EPANET, RIEGOLOC results comparison for a 90 m trickle irrigation pipe.

The Figure 10 tells that computed flow rate emission are equivalent in GESTAR ECR, GESTAR SBS and RLOC, but EPANET SBS fails to predict accurately enough the small flow rates emitted. The errors are attributed to either a flow rate round off or truncation of coefficients for low resistance connections between emitters.

4 GESTAR'S FIELD VALIDATION

In relation to the demonstration of the code's capabilities for engineering and management through existing facilities, a systematic validation was performed over an extensive irrigation zone, covering 9000 ha located in the course of the river Cinca (Huesca, Spain).

The main focus of interest in this validation study was not only centered on testing whether algorithms calculations have been correctly implemented, (this being verified through comparison with other reference programs), but rather on research into how the joint effects of various uncertainty factors inherent in the construction of the model affect the final results thus establishing rigorously checked viability margins for prediction. The uncertainties are generally associated with inaccuracy in layout plans and altitudes, quality of construction; lack of precision in the flow from outlets; in the calculation of distributed and located head losses; the roughness of pipes, etc.

The results of the code have been compared with field experimental pressure measures in 125 testing points, involving five different independent hydraulic sectors.

In order to do this data about the five distribution networks relating to construction, topology and hydraulic equipment was first entered into GESTAR using available plans, maps and other documents. Layout plans were exhaustively checked against the reality in the field and were found on occasion to differ.

A minimum of 19 measurement points and a maximum of 36 were placed in each sector. Measurement points were selected so that there was at least one control point in each branch of the network. At each of these points a previously checked standard manometer with a range of 0 to 10 kg/cm^2 and accuracy of 0.1 kg/cm^2 was installed. The distribution of open hydrants and the volume each one of them extracted from the network was precisely given at the same instant as pressure readings were taken. The flow rate from the hydrants, which were equipped with mechanical type PRVs and FCVs, was estimated with a 5% error margin. As a means of contrasting and verifying flow rate, readings were taken of the total water volume supply to the main arteries of each network. This was done by means of permanently installed ultrasonic flow meters with a range

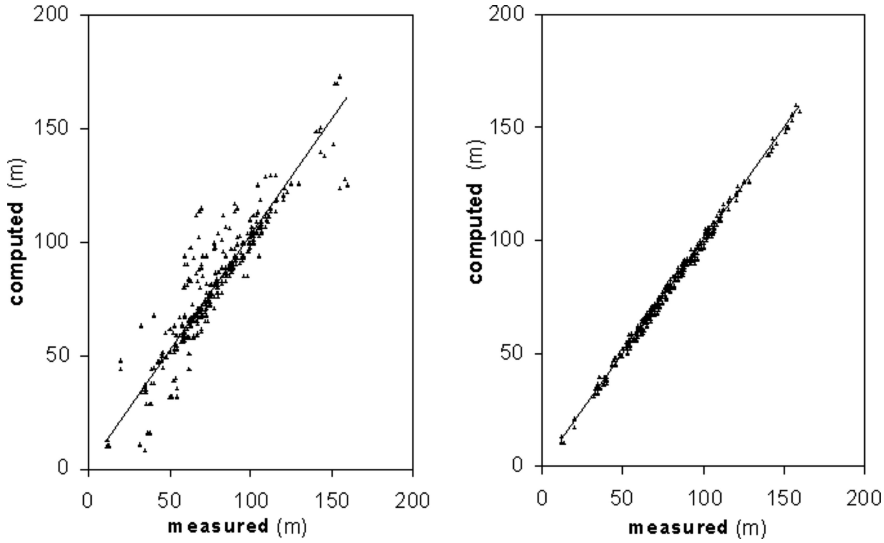


Figure 11. Measured and computed pressure previous to data checking and calibration and measured and computed pressure after correction of input error and calibration.

of 0 to 1500 l/s and a resolution of 10 l/s. Discrepancies were found of less than 6% with respect to the total volume of the open outlets.

Readings were taken, with each hydraulic sector functioning independently, over a period of between four and six consecutive days. Bearing in mind that at least 50% of the working hydrants were different ones on different days, the range of operating conditions tested was sufficiently varied. A total of 405 valid readings were taken in varying controlled flow conditions.

Once a model had been generated GESTAR's scenario tools were used to reproduce the exact aperture of the hydrants at the moment when readings were taken. The program came up with the pressure values calculated for these configurations. The calculated values were compared to the readings and the differences between them were observed. First graph in Figure 11 shows the correlation of experiments and computations found in the first rough comparison. This initial step helped to detect previously unnoticed mistakes in the data entry.

After correcting errors, the calculated values and the read values were found to be of the same order in most cases. However, notable differences appeared systematically in some locations. These were found to be the result of errors in the altitudes given in plans for those particular nodes. Once these were corrected by means of careful data checking and "ad-hoc" topographic measurements, the values obtained from the model and the pressure readings from the field tests were similar in all cases, with differences of 3.3 m on average (second graph in Fig. 11).

Finally, with the aim of refining the predictions and calibrating the model, located head losses were meticulously considered in every element (bends, valves, bifurcations, etc.) by means of suitable head-loss coefficient. Some of the values for the roughness of the pipes were increased with respect to commercial values. Owing to this, calculated and measured values became even more similar with a very satisfactory agreement, without further calibrations. Absolute discrepancies of 1.67 m on average (rms. equal to 0.98 m, maximum difference 3 m) were obtained.

The good fit between predicted and measured values have not required the complex adjustment or exhaustive calibration procedures used in the hydraulic modelling of drinking water systems. It shows that the use of predictive techniques and generalised hydraulic analysis is very reliable in pressurized irrigation systems. This fact was predictable given the following reasons:

a) In pressurized irrigation systems plans, layout and constructive details are available for consultation, given that construction dates are relatively recent.

b) The complete conduit network, with its individual demand outlets and all its singularities, can be incorporated into the model with no need for simplifications or synthesis.
c) The presence of flow control mechanisms and/or meters at all outlets allows sufficiently accurate, moment by moment monitoring of consumption.

In the irrigation areas where the code was validated, GESTAR is being used to plan daily demand; to allocate turns in periods of maximum consumption as well as to propose changes in network design to correct malfunctions that may occur.

5 CONCLUSIONS

The engineering and management of pressurized irrigation networks is a field which lends itself very well to the application of simulation systems and DSS. In particular, detailed hydraulic simulation of trickle and drip irrigation networks within plots show numerous advantages over more traditional, less accurate procedures of an explicit type. Irrigation systems have a large number of characteristics which differentiate them from drinking water networks. These differences are especially suited to analysis and diagnosis by software adapted to their specific idiosyncrasies. However, scarcely any advanced codes with these characteristics are available at the moment.

GESTAR is just such a software package. It is being developed in the Department of Fluid Mechanics at the Zaragoza University in collaboration with those public entities responsible for irrigation system construction, with engineering and with communities using the networks. The efficient computational nucleus incorporates options for carrying out inverse analysis and reproduces the actual behaviour of the regulating devices in the irrigation network, in both routine and exceptional circumstances. A GUI containing a variety of tools enabling network loading, configuration and irrigation demand scenario analysis facilitates the use of GESTAR in both the design and use of pressurized irrigation systems.

From the exhaustive validations, carried out by means of comparisons between calculated pressures and those obtained from experimental readings in a set of five operational networks, viability and accuracy margins have been obtained which fit real, routine conditions and make extensive calibration unnecessary.

The software is being used to tackle an important number of very varied experiences which will in turn permit the development of new tools and the stabilising of systematic user protocols. This has introduced new concepts in the design and management of pressurized irrigation networks.

The use of GESTAR in innumerable projects of new irrigation systems design as well as in reforms and the day to day running of existing networks, is leading to numerous new applications and strategies for the program which, once they are systematised, have introduced a radically different new approach to hydraulic engineering in irrigation systems.

GESTAR is a solid a versatile tool wich incorporate, as seen in this paper, complex and advanced numerical techniques for hydraulic simulation of pressurized irrigation networks, but its design target is to give an easy and friendly experience to the user.

ACKNOWLEDGEMENTS

The research presented here is being undertaken with the support of: the Department of Agriculture and the Environment of the Aragonese government, under contracts 95/0106, 2000/0287, 2003/0280 and 2006/0280, the Aragonese Council of Research and Development (CONSID), project P-43/967, P-316-2001, and the Spanish government's Inter-Ministry Commission for Science and Technology (CICYT), as part of their Programme for Research into Hydraulic Resources, projects HID97/0999, HID98-0341-C03-C02, AGL2001-1180-CO-02.

APPENDIX I

Table 1. Nodes values of studied network.

ID	Type	Total head (m)	Elevation (m)	Emitter Coeff. (l/s/m$^{1/2}$)	GESTAR K_S (s^2/m^5)	N exponent	Base demand (l/s)
1	Reservoir	417.17					
3032	Junction		355				
3033	Junction		355				
4006	Junction		341				
1142	Junction		341.36				
4	Junction		341.36	3.7947	69444.45	2	24
1141	Junction		357.04				
s2	Junction		357.04	5.0596	39062.5	2	32
1140	Junction		367.41				
s4	Junction		367.41	3.7947	69444.45	2	24
3011	Junction		367.41				
1139	Junction		372.47				
s5	Junction		372.47	2.3717	177777.8	2	15
3010	Junction		372.47				
3	Junction		372.47	17.3925	3305.78	2	110
1137	Junction		371.87				
s3	Junction		371.87	2.3717	177777.8	2	15

Table 2. Pipe values of studied network.

ID	Start node	End node	Length (m)	Diameter (mm)	Roughness (mm)
1	1	3032	1132.39	700	0.025
4006	3033	4006	455.89	700	0.025
1142	4006	1142	269.61	600	0.025
1141	1142	1141	319.28	500	0.025
1140	1141	1140	59.1	450	0.025
3011	1140	3011	372.47	450	0.025
1139	3011	1139	387.34	350	0.025
3010	1139	3010	452.49	350	0.025
1137	3011	1137	440.99	300	0.025

Table 3. Valves values of studied network.

ID	Start node	End node	Diameter (mm)	Type	Setting (m)	Loss Coeff. (dimensionless)	Fixed status
4	3032	3033	700	PSV	62.1	1.306264	None
18	1142	4	600	PRV	40	0	None
8	1141	s2	500	PRV	40	0	None
9	1140	s4	450	PRV	40	0	None
10	1139	s5	350	PRV	40	0	None
17	3010	3	350	PRV	40	0	None
12	1137	s3	300	PRV	40	0	None

APPENDIX II

In Darcy-Weisbach head losses equation, friction factor f can be expressed by

$$f = f_0 v^{m-2}$$

where v is velocity and f_0 and m are constants values.

This strategy is derivated of Moddy diagram's form, and it is necessary to distinguish three intervals. For example:

If $10^5 < Re$. Complete turbulent flow. Watters & Keller (1978) expression is used:

$$f_0 = 0,130 \left(\frac{\mu}{\rho D}\right)^{0,172} ; \quad m = 1,828$$

REFERENCES

Aliod, R., Eizaguerri, A., Estrada, C. & Perna, E. 1997. Design and Hydraulic Analysis of Pressurized Irrigation Networks. Application of GESTAR. (In Spanish). *Riegos y Drenajes Siglo XXI* 92: 22–38.

Boulos, P. F. & Wood, J. 1990. Explicit Calculation of Pipe-Network Parameters. *Journal of Hydraulic Engineering* 116(11): 1321–1344.

Campos, A. 1993. Steady-State Fluid Network Analysis. *Journal of Hydraulic Engineering* 119(3): 431–436.

Clement, R. & Galand, A. 1986. *El riego por aspersión y las redes colectivas de distribución a presión*. Barcelona: Editores Técnicos Asociados.

Duff, I.F., Erissman, A.M. & Reid, J.K. 1986. *Direct Methods for Sparse Matrices*. Oxford Science Publications.

Estrada Collado, C. 2000. *Técnicas Robustas y Generalizadas para la Simulación Hidráulica Estacionaria de Redes de Riego a Presión con Aplicaciones a Diseño y Gestión*. PhD Thesis. Zaragoza (Spain): Zaragoza University.

González, C. & Aliod, R. 2003. *Mejoras en el método de la serie económica para el dimensionado de redes ramificadas*. Actas del XXI Congreso Nacional de Riegos (Mérida, Badajoz). Madrid (Spain): Asociación Española de Riegos y Drenajes.

González, C. & Aliod, R. 2005. *Modelización integrodiferencial de la hidráulica de un elemento con consumo en ruta dependiente de la presión y alimentado por los dos extremos*. Actas del XXIII Congreso Nacional de Riegos (Elche, Alicante). Madrid (Spain): Asociación Española de Riegos y Drenajes.

González, C. 2006. *Diseño hidráulico avanzado de sistemas de riego presurizado mediante técnicas de simulación y optimización*. PhD Thesis. Zaragoza (Spain): Zaragoza University.

Guillén, J., Bescós, M., Doz, J.R. Marzal, A. & Royo, J. 1997. *New developments and aplications of GESTAR. Validation by field experiments in the Sector XXXIII of the Cinca Canal* (In Spanish). XIII Congreso de Ingeniería de Proyectos. Sevilla (Spain): Asociación Española de Ingeniería de Proyectos.

Labye, Y., Olson, M.A., Galand, A. & Tsiourtis, N. 1988. *Design and optimization of irrigation distribution networks.* FAO Irrigation and Drainage paper. Food and Agriculture Organization of the United Nations.

Rodrigo, J., Hernández, J.M., Pérez, A. & González, J.F. 1992. *Riego Localizado*. Madrid (Spain): Ediciones Mundi-Prensa.

Rossman, L. 2000. *Epanet 2 User's Manual*. Cincinnati, USA: Environmental Protection Agency.

Todini, E. & Pilati, S. 1987. *A Gradient Algorithm for the Analysis of Pipe Networks*. International Conference on Computer Applications for Water Supply and Distribution: 1–20. Leicester (UK).

Watters, G. Z. & Keller, J. 1978. Trickle irrigation tubing hydraulics. *Summer Meeting of ASAE*.Paper 78–2015, Utah State University, USA.

Technical Contributions

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Parallel computation of unsteady inundation shallow-water flow

B. Latorre, J. Murillo & P. García-Navarro
Fluid Mechanics, CPS, University of Zaragoza, Zaragoza, Spain

ABSTRACT: This paper presents a parallel solver for the simulation of transient shallow water flows in the presence of wetting/drying fronts. A first order explicit finite volume numerical scheme on unstructured triangular grids has been adopted to solve the two-dimensional shallow water equations. The parallel implementation relies on a dynamic domain decomposition technique intended to ensure the balance of the computational load during the simulation. The strategy developed allows the parallel computation of surface geophysical flows, such as floods or dam-break events in real scenarios, where the wet domain is dynamically changing and a static domain decomposition technique is inefficient. The inter-processor data communication is incorporated by the message passing interface (MPI) protocol. A practical application is presented to validate the performance of the implementation.

1 INTRODUCTION

Shallow water model is widely used to simulate surface geophysical flows. These situations usually involve large size domains and long time scales. Practical applications require a compromise between spatial accuracy and computational efficiency. In order to achieve the necessary spatial resolution, rather fine grids become necessary in many cases so requiring more data storage, increasing proportionally the number of operations and reducing the allowable time step size for explicit calculations. When, at the same time, a reasonable computational time is desirable, the use of parallel codes in distributed memory systems is one of the options. It allows computing realistically large size and long time problems in a controllable computation time.

To construct an efficient parallel implementation there are two requirements: the computational load assigned to each processor must be balanced and the communication between processors must be minimized. In the case of explicit numerical schemes this can be easily done using domain decomposition strategies (Castro et al., 2005). When dealing with transient inundation flows, the computational domain in our model changes during the simulation since only wet cells are included. This makes necessary a dynamic work assignment to the processors in order to keep balanced their work load. At the same time, the method used to control the correct numerical modeling of the wetting/drying fronts involves a correction step that is not fully local. This is an additional difficulty for the parallelization structure.

2 GOVERNING EQUATIONS AND NUMERICAL SCHEME

The two-dimensional shallow water model is formed by the unsteady equations of mass and momentum conservation, obtained by depth averaging Navier-Stokes equations:

$$\frac{\partial h}{\partial t} + \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} = 0 \tag{1}$$

$$\frac{\partial q_x}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q_x^2}{h} + \frac{gh^2}{2}\right) + \frac{\partial}{\partial y}\left(\frac{q_x q_y}{h}\right) = gh\left(S_{ox} - S_{fx}\right) \tag{2}$$

$$\frac{\partial q_y}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q_x q_y}{h}\right) + \frac{\partial}{\partial y}\left(\frac{q_y^2}{h} + \frac{gh^2}{2}\right) = gh\left(S_{oy} - S_{fy}\right) \tag{3}$$

where, h represents the water depth, $\mathbf{q} = (q_x, q_y)$ is the unit discharge vector and g is the acceleration of the gravity. The source terms are composed by the bed slope terms and the friction losses, expressed in terms of the Manning's roughness.

The numerical scheme adopted to solve the sallow water equations is first order accurate in space and time and is based on an upwind finite volume method. A detailed description of the numerical method can be found in (Murillo et al., 2005). The resulting updating expression at a generic cell 'i' can be expressed as a sum of incoming contributions from the edges of the cell:

$$A_i \frac{\Delta \mathbf{U}_i}{\Delta t} = \sum_k \left[l_k \cdot \mathbf{F}\left(\mathbf{U}_i, \mathbf{U}_j, \mathbf{n}_k\right)\right] + \mathbf{S}\left(\mathbf{U}_i\right) \tag{4}$$

where the index j represents the cell that limits the cell i through the wall k. A represents the area of the cell, $\mathbf{U}$ is the vector of conserved variables, l is the length of the wall and $\mathbf{n}$ is the unit normal vector to the cell edge outward to cell i. On the other hand, $\mathbf{F}$ is the part of the normal flux difference vector across the edge that participates in the updating of the variables at cell i. It only depends on the values $\mathbf{U}_i$ and $\mathbf{U}_j$ sharing the edge. $\mathbf{S}_i$ represents the average cell value of the source term vector. Finally, Δt represents the global time step chosen as the most restrictive among all the cells in the domain. At each cell, the maximum time step size compatible with numerical stability is calculated according to the CFL condition.

A specific modification of the solver was described in (Murillo et al., 2006) to overcome the generation of negative values of depth that can appear as a consequence of wetting/drying fronts over variable bed levels. The algorithm of the explicit scheme is based on a complete run over the grid edges so that the information travelling into every cell is accumulated for updating purposes. Afterwards, a second run is performed over the grid cells defining the wet/dry front to make the necessary corrections to the solution so that it is positive, monotone and conservative. These corrections affect to both the cells getting and giving water at the front hence requiring a communication between processors when the front crosses from one sub-domain to the other.

3 PARALLEL IMPLEMENTATION

The strategy proposed for the efficient parallel implementation of the numerical method is of the dynamic domain decomposition type. The computational cell domain is partitioned, independently of the number of processors, into a high number of non-overlapping portions. The partition is made in a one-dimensional way so that portion i limits only to portions $i+1$ and $i-1$. During the simulation, connected sets of portions are dynamically assigned to each single processor, as sketched in Figure 1, to address load balancing. The number of portions determines the optimization of work load assignment to the processors.

To minimize the amount of communication data, the partition must ensure a minimum number of virtual cells and walls. This is a very difficult task when the computational domain is unknown and transient that can become case dependent. In our model, the basic option is to establish portions perpendicular to the local flow main direction that can be based on the bottom level isolines. As a second suggestion, given an approximate estimation of the wet domain, a decomposition based on software tools could be used, always taking into account that these tools are designed for the optimum parallelization of steady domains.

The flowchart of the parallel algorithm is sketched in Figure 3. The main algorithm can be summarized as follows. Edges shared by two computational domains are called virtual edges

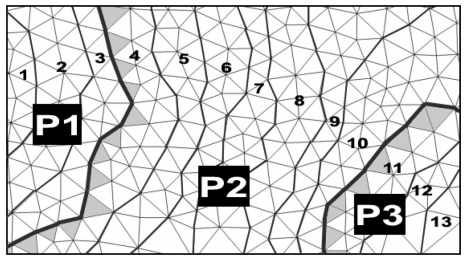


Figure 1. Mesh partition and load assignment to processors example. Virtual cells are grey colored and virtual edges are drawn in thick lines.

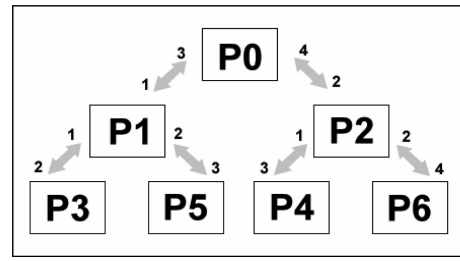


Figure 2. Communication tree diagram. The numbers near the arrows set the step number, for the up-tree or down-tree process, in which the communication occurs.

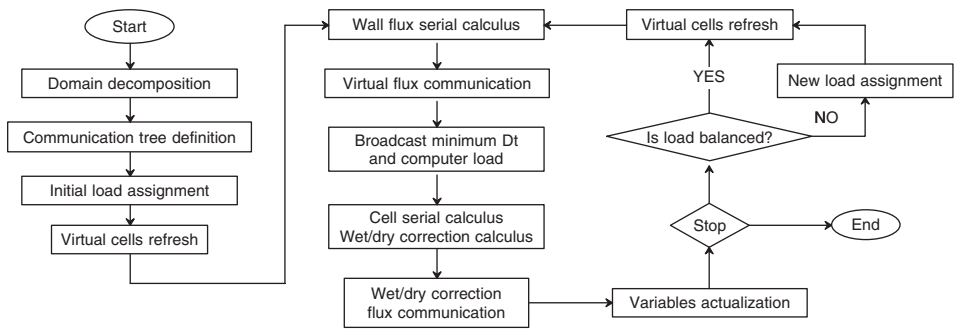


Figure 3. Flowchart of the parallel algorithm.

(Fig. 1). Only one processor calculates the fluxes at these edges (the one with a lower label) and requires the information of the variables at the cells on both sides; some of these belong to the limiting domain (virtual cells). The flux contributions updating the virtual cells are sent to the corresponding processor. Once all the flux contributions are accumulated at the cell centres all over the full domain, if a dry/wet front crosses the boundary between two processors, correction fluxes affecting cells of both processors may appear. Virtual cell refreshing and flux communication are local processes, involving pairs of neighbour processors. To complete the variables updating it is necessary to broadcast the global minimum Δt. For optimization purposes in the global communication steps (minimum Δt and load distribution) a tree structure in two branches as sketched in Figure 2 has been implemented. First we start an up-tree communication process so that the useful information arrives at processor labelled as 0 and then starts a down-tree communication propagating the information to all processors.

In the broadcast of the global minimum Δt each processor sends also information about the consumed computational time in the last time step and about the number of wet cells and edges of the assigned portions. According to this information, processor labelled as 0 estimates at each iteration the time lost because of the bad load balancing. When this time loss exceeds the necessary time for a new load assignment, portions are reassigned in order to achieve the load balance. A collective communication process then starts where the processors exchange domain portions. The load assignment is performed according to the effective computational power of every processor. This strategy is suitable for the efficient use of heterogeneous computer networks or those with a fluctuating computational capacity.

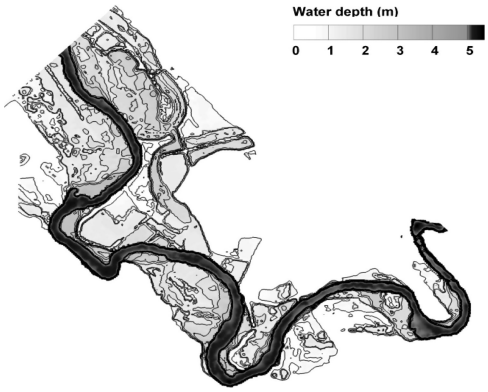


Figure 4. Flooded domain at $t = 3$ h.

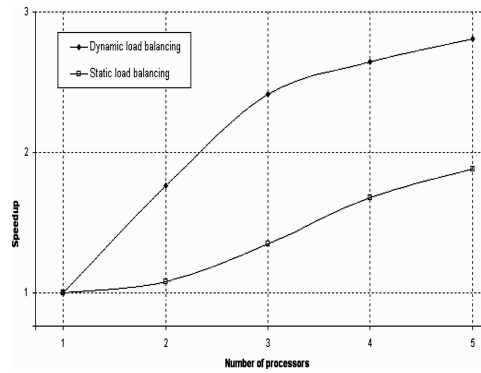


Figure 5. Speedup measures.

4 APPLICATIONS

The simulation of an inundation flow in a real scenario was considered in order to validate the performance of the parallel implementation. The area of interest has a size of 35 km^2. Simulations were performed using a triangular mesh of 130.000 cells. An input discharge of 3.000 m^3/s was considered to flow over the initially dry domain during 3 hours.

Two sets of simulations have been run. In the first, a constant number of portions were assigned to the processors. Since a uniform distribution of the portions among the processors would lead to an overly inefficient method, a non-uniform assignment was made in order to optimize the computational load during the unsteady flow. The size of the upstream processor domains is smaller and increases smoothly in the downstream direction. This is called static load balancing. In the second, a time dependent number of portions are assigned to the processors in order to optimize the computational load during the unsteady flow. This is called dynamic load balancing.

Simulation results are plotted in Figure 4, where the speedup is defined as the ratio between the parallel and the serial run time.

5 CONCLUSIONS

We have presented the basic features and performance of a parallel solver for the simulation of transient shallow water flows in the presence of wetting/drying fronts. Simulation results show that a dynamic work assignment increases the efficiency of the parallel implementation when considering simulations with transient wet domain.

REFERENCES

J. Murillo, P. García-Navarro, P. Brufau and J. Burguete (2006). A conservative 2d model of inundation flow with solute transport over dry bed (in press) Published online in Wiley InterScience. International Journal for Numerical Methods in Fluids.45: 1047–1082.

M. J. Castro, J. A. García, J. M. González and C. Parés (2005). A parallel 2D finite volume scheme for solving systems of balance laws with nonconservative products: application to shallow flows. Computer Methods in Applied Mechanics and Engineering: 195:288–2815.

J. Murillo, J. Burguete, P. Brufau and P. García-Navarro (2005). Coupling between shallow water and solute flow equations: analysis and management of source terms in 2D. International Journal for Numerical Methods in Fluids, 49:267–299.

P. Roe (1986). A basis for upwind differencing of the two dimensional unsteady Euler equations, Numerical Methods in Fluid Dynamics II. Oxford Univ. Press: Oxford, 1986.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

A 2D finite volume multiblock flow solver applied to flood extension forecasting

S. Erpicum, P. Archambeau, S. Detrembleur, B. Dewals* & M. Pirotton
Applied Hydrodynamics and Hydraulic Constructions Research Unit, MS²F – ArGEnCo Department, University of Liege, Liege, Belgium

ABSTRACT: This paper describes a quasi-3D free surface flow solver, developed for almost ten years at the University of Liege, and its application to flood extension forecasting. Based on the classical SWE, the finite volume numerical model deals with real topography and wetting-drying cells while exactly conserving the water volume. Multiblock options allow local refinement of the structured calculation grid and automatic mesh refinement enhances greatly computation time. An additional depth-averaged k-ε type turbulence model increases the potentialities of the solver to represent real complex flows. Following a Regional Government decision, this solver has been officially chosen in 2003 to be used to compute inundation maps on 800 km of the main rivers of the Walloon Region in Belgium with a spatial resolution of 5×5 to 2×2 meters and using high resolution Digital Elevation Models. Results are presented in this paper.

1 INTRODUCTION

In the global framework of recurrent large flood events all over the world and in Europe in particular, relevant inundation maps represent a tool of prime interest to help decision makers in defining land use and housing policies and to help insurance companies to asses the risks associated with flooding. In this scope, relevant and very precise flood extension and velocity fields maps are needed for several probabilistic extreme events.

The present paper covers a detailed description of a state-of-the-art complete 2D numerical model to be used as a strategic tool in the process of flood risk assessment and mitigation.

This model is integrated in the software package WOLF, which has been developed for more than ten years at the University of Liege. WOLF includes a complete set of numerical models for simulating free surface flows (process-oriented and spatially distributed hydrology, 1D and 2D hydrodynamics, sediment transport, air entrainment …) as well as an optimisation algorithm. This last tool, based on the robust Genetic Algorithms technique, allows an objective calibration of friction coefficients for example.

A user-friendly GIS interface, entirely designed and implemented by the authors, makes the pre- and post-processing operations very convenient. Import and export operations are easily feasible from and to various classical GIS tools. Different layers of maps can be handled to analyse information related to the topography, the ground characteristics, the vegetation density and the hydrodynamic fields.

*FNRS Post-doctoral researcher

2 WOLF2D NUMERICAL MODEL

2.1 *Mathematical model*

The model WOLF2D solves the classical SWE. In this approach, the only assumption states that velocities normal to a main flow direction are smaller than those in the main flow direction. As a consequence the pressure field is found to be almost hydrostatic everywhere. The large majority of flows occurring in rivers can reasonably be seen as shallow everywhere, except in the vicinity of some singularities (e.g. weirs). The divergence form of the shallow-water equations includes the mass balance:

$$\frac{\partial h}{\partial t} + \frac{\partial q_i}{\partial x_i} = 0 \tag{1}$$

and the momentum balance:

$$\left[\frac{\partial q_i}{\partial t} + \frac{\partial}{\partial x_i}\left(\frac{q_i q_j}{h}\right)\right] + gh\left(S_{fi} + \frac{\partial H}{\partial x_i}\right) + \frac{\partial \tau_{ji}}{\partial x_i} = 0; \quad j = 1, 2 \tag{2}$$

where Einstein's convention of summation over repeated subscripts has been used. H represents the free surface elevation, h is the water height, q_i designates the specific discharge in direction i, S_{fi} is the friction slope and τ_{ij} are the depth integrated turbulence stresses.

Several turbulence models are implemented in WOLF2D, starting from turbulent viscosity algebraic relations to an original depth-integrated two additional equations model (Erpicum 2006).

2.2 *Algorithmic implementation*

The space discretization of the 2D conservative shallow-water equations is performed by a finite volume method. This ensures a proper mass and momentum conservation, which is a prerequisite for handling reliably discontinuous solutions such as moving hydraulic jumps. As a consequence, no assumption is required as regards to the smoothness of the unknowns.

Designing a stable flux computation has always been a challenging and tough issue in computing fluid dynamics, especially if discontinuous solutions are expected. Flux treatment is here based on an original flux-vector splitting technique developed for WOLF (Mouzelard 2002). The hydrodynamic fluxes are split and evaluated according to the requirements of a Von Neumann stability analysis. Much care has been taken to handle properly the source terms representing topography gradients (Erpicum 2006).

Since the model is applied to transient flows and flood waves, the time integration is performed by means of a second order accurate and hardly dissipative explicit Runge-Kutta method.

2.3 *Friction modelling*

River and floodplain flows are mainly driven by topography gradients and by friction effects. The total friction includes three components: bottom friction (drag and roughness), wall friction and internal friction.

The bottom friction is classically modelled thanks to an empirical law, such as the Manning formula. The model enables the definition of a spatially distributed roughness coefficient. This parameter can thus easily be locally adjusted as a function of local soil properties, vegetation or sub-grid bed forms. An original evaluation of the real shear surfaces is realized and the friction along vertical boundaries, such as bank walls, is reproduced through a process-oriented model developed by the authors.

The internal friction is taken into account by the turbulence model (see § 2.1).

2.4 *Multiblock grid and automatic grid adaptation*

WOLF2D deals with multiblock structured grids (Erpicum et al. 2004). This feature enables a mesh refinement close to interesting areas without leading to prohibitive CPU times. A grid adaptation technique restricts the simulation domain to the wet cells, thus achieving potentially drastic reductions in CPU times.

Besides, the model incorporates an original method to handle covered and uncovered (wet and dry) cells. Thanks to an efficient iterative resolution of the continuity equation at each time step, based on a correction of the discharge fluxes prior to any evaluation of momentum balances, a correct mass conservation is ensured in the whole domain (Erpicum et al. 2004).

2.5 *Mesh refinement*

In case of transient computations, an automatic mesh refinement technique is used to enhance the convergence rate towards accurate steady-state solutions (Archambeau et al. 2004).

The computations are performed on several successive grids, starting from a very coarse one gradually refined up to the finest one. When the hydrodynamic fields are almost stabilized, the solver automatically jumps onto the next grid. The successive "initial solutions" are interpolated from the coarser towards the finer grid in terms of both water heights and discharges. This fully automatic method considerably reduces the number of cells in the first grids and then substantially decreases computation time, despite extra computation time for meshing and interpolation operations.

3 FLOOD EXTENSION COMPUTATION

3.1 *Introduction*

Following a Regional Government decision in 2003, the solver WOLF2D has been officially chosen to be used to compute inundation maps on 800 km of the main rivers of the Walloon Region in Belgium.

From constant discharge values with a 25-, 50- and 100-year return period, water elevations and velocity fields have to be compute on the main rivers of the South part of Belgium using 5×5 to 2×2 m grids and a high resolution Digital Elevation Models. For each river, the simulation of real large flood events has to be realized to calibrate the roughness coefficient of the model and to validate the approach.

3.2 *DEM*

Very recently, the Belgian Ministry of Facilities and Transport (MET), and in particular the Service of Hydrology Studies (SETHY), acquired an accurate DEM on the floodplains in the whole Walloon Region. An airborne laser has been used to characterise the inundation area of the main rivers network and an echo-sonar technique has been applied to measure the bathymetry of the main channel on navigable rivers. On smaller non navigable rivers, cross sections have been interpolated to generate the main bed topography information.

Consequently, the poor and inaccurate 3D information available for many years, 30 m of plan resolution with a precision of several meters in altitude, has been replaced by an exceptional DEM since the precision in altitude is 15 cm and the information density is one point per square meter (Fig. 1).

Generally, specific features of flows in urban area, such as blockage by buildings, are taken into account in hydraulic numerical models by a local modification of the roughness coefficient. With the new set of topographic data, irregularities of the topography influence directly the inundation flows. This allows focusing on proper physical values of the roughness coefficient.

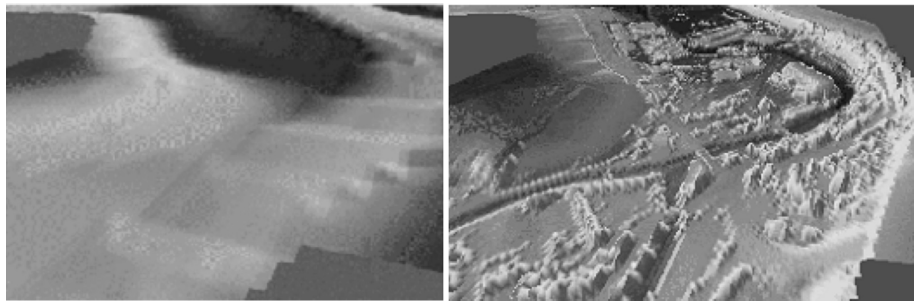


Figure 1. Town of Eupen: old IGN topography (left) and recent laser data (right).



Figure 2. Lesse river – Comparison between a real flood event and the corresponding simulation.

3.3 *Example of results*

The following pictures (Fig. 2) show a comparison between the flood extension of a real event on the river Lesse and the corresponding numerical results.

The accuracy of both the DEM and the numerical model allow very precise prediction of the flood extension, even in difficult modelling situations such as bridges (A), villages (B), or flat area (C).

4 CONCLUSIONS

The application to flood extension forecasting of a conscientiously developed 2D numerical free surface flow solver using high resolution topography data allows obtaining reliable results on 800 km of the main rivers of the South part of Belgium. Together with the multiplicity of the results available, the detailed management of the singularities in the rivers beds leads to a better understanding of the rivers functioning and of the inundation processes.

Beyond these useful results for risk analysis and land use policies, the fine grid of the simulation associated with the considerable elevation precision allow very precise impact and remediation studies, and the model plays thus the role of a genuine tool to help decision makers.

Resulting from several years of research and scientific development, WOLF2D is now fully operational and efficient for large scale real application thanks to the fast evolution of both calculation and data acquisition technologic resources.

The perspectives of improvements in the near future have to link these flow simulations with hydrologic computations of the whole river watershed to predict the flood extension from real time rain data.

REFERENCES

Archambeau, P., Dewals, B. Detrembleur, S., Erpicum, S., Pirotton, M. 2004. A set of efficient numerical tools for floodplain modeling. In Jirka G.H. & Uittewaal W.S.J. (ed.), *Shallow flows*. Leiden: Balkema

Erpicum, S., Archambeau, P., Dewals, B., Detrembleur, S., Pirotton, M. 2004. Computation of the Malpasset dam break with a 2D conservative flow solver on multiblock structured grid. In Liong S., Phoon K. & Babovic V. (ed), *Proc. of the 6th Hydro informatics Conf. Singapore, 21–24 June 2004*

Erpicum, S. 2006. *Optimisation objective de paramètres en écoulements turbulents à surface libre sur maillages multibloc*, PhD Thesis. Liege: HACH, University of Liege

Mouzelard, T. 2002. *Contribution à la modélisation des écoulements quasi tridimensionnels instationnaires à surface libre*, PhD Thesis. Liege: HACH, University of Liege

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

RKDG2 scheme for the 1D open channel flow

G. Kesserwani, J. Vazquez, R. Ghostine & R. Mosé
U.P.R. Système Hydrauliques Urbains, Ecole Nationale du Génie de l'Eau et de l'Environnement de Strasbourg, Strasbourg, France

A. Ghenaim
Institue Nationale des Sciences Appliqués, Strasbourg, France

ABSTRACT: Runge-Kutta discontinous Galerkin (RKDG) schemes are finite element methods using also finite volume (FV) techniques such as: Riemann solver, slope limiter and Total-Variation-Diminishing (TVD) Runge-Kutta time discretization. The intension of this survey is to establish a one-dimensional numerical computation of the shallow water equations (SWE) by means of a second order Runge-Kutta discontinuous Galerkin (RKDG2) scheme. Hence, this method is briefly described and incorporated with a robust handling of the boundary conditions (BC) which was performed by the use of the characteristics theory. By the support of: (a) results reported in the literature; (b) applications to the SWE, we illustrate the effectiveness of the proposed numerical scheme near the results of recent and traditional FV methods.

1 INTRODUCTION

In the last few years, several explicit numerical techniques for solving hyperbolic conservation laws have been proposed by many researchers. Particular applications to the SWE were examined by: García-Navarro & Savirón (1992) who extended a TVD version of the McCormack scheme; Delis & Skeels (1998) who presented and studied the performance of four widely used TVD schemes; Delis & Katsaounis (2003) who initiated a new approach for incorporating source terms in the relaxation model; and Burguete & García-Navarro (2001) who pioneered a new class of well-balanced high resolution TVD schemes. Recently, Črnjaríc-Žic & al. (2004) extended FV-WENO (Weighted Essentially Non-Oscillatory) schemes and central WENO schemes to the hyperbolic balance laws with the geometrical source term and the spatially variable flux. Most recently, Xing & Shu (2006b) generalized high order FV-WENO schemes and RKDG finite element methods to the same class of hyperbolic systems in order to maintain a well-balanced property.

The purpose of this study is to perform a one-dimensional numerical approximation of open channel flows where source terms are involved in the shallow water model. In practical applications, it could prove advantageous to decrease the number of the computational cells in order to save CPU-time cost without having a deterioration in the quality of the numerical solution. Therefore, supported by relevant references reported in the literature, the finite element RKDG2 scheme is selected and briefly overviewed. The algorithm is implemented in conjunction with Roe's scheme where a robust treatment of the BC by the characteristics theory (García-Navarro & Savirón 1992) is invoked. A particular numerical investigation for the SWE is presented illustrating the performance of the chosen scheme close to the performance of some conventional second order TVD schemes. Therefore, two steady tests having analytical solutions are selected: one characterized by a wholly supercritical flow; and the other is about a hydraulic jump occurrence.

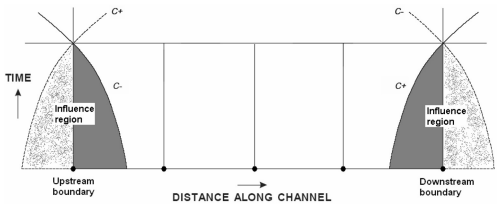


Figure 1a. Characteristic curves at the boundaries for subcritical flow in the $(x-t)$ plan.

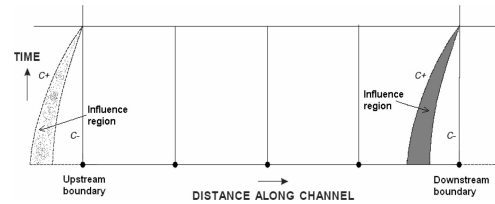


Figure 1b. Characteristic curves at the boundaries for supercritical flow in the $(x-t)$ plan.

Strong numerical evidence show that the RKDG2 scheme can improve the solution accuracy with less computational cells than the FV schemes and is of practical consideration for river flood applications.

2 MATHEMATICAL MODEL

It is generally accepted that the unsteady flow of water in a wide channel of slowly varying cross-section with a sufficiently gentle bottom slope can be described by the one-dimensional SWE. Recent schemes are based on the conservative formulation of the Saint Venant system, because it is related closely to the fluxes action on the flow. For flow in a prismatic channel of rectangular cross-section, the system has the following conservative form:

$$U_t + F(U)_x = G(U) \tag{1}$$

where, $U = [A; Q]^{\mathrm{T}}$; $F(U) = [Q;\ Q^2/A + 0.5\,g\,B\,h^2]^{\mathrm{T}}$; and $G(U) = [0; g\,A\,(S_o - S_f)]^{\mathrm{T}}$. $t =$ time in s; $x =$ longitudinal distance in m; $U =$ flow vector; $F =$ flux vector and $G =$ source term vector. Where: $A =$ cross-sectional area in m^2; $Q =$ flow discharge in m^3/s; $\mathrm{h} =$ water depth in m; $g =$ acceleration due to gravity in $\mathrm{m/s}^2$; $B =$ channel bottom width in m; $S_o =$ the bed slope; and $S_f =$ friction slope defined in terms of the Manning's roughness coefficient.

3 NUMERICAL SCHEME AND BOUNDARY CONDITIONS

The governing equations are discretized. The computational domain is divided into N uniform cells $I_i = [x_{i-1/2}; x_{i+1/2}]$ of length Δx. The RKDG2 method is considered according to the approach of Cockburn & Shu (2001). We seek a local approximation U_h to U such that for each time $0 \le t \le T$, U_h belongs to the finite dimensional space of polynomial in I_i of degree at most k. Therefore, the system 1 is multiplied by a continuous test function v_h and integrated over I_i. Afterward, the flux term is integrated by part to obtain the following weak formulation:

$$\int_{I_i} \partial_t U_h v_h dx - \int_{I_i} F(U_h)\partial_x v_h dx + [F(U_h)v_h]_{i+1/2} - [F(U_h)v_h]_{i-1/2} = \int_{I_i} G(U_h)v_h dx \tag{2}$$

In the aim of decoupling the system, we adopt the Legender polynomials as local basis functions. The degree k was chosen equal to unity giving a second order accuracy approximation in space. In the treatment of integral terms we used Gaussian quadrature rules, and to maintain the stability of the method in the vicinity of strong shocks the classical *minmod* slope limiter was applied. The flux at the cell's interface was handled by a numerical flux function based on the approximate Riemann solver of Roe (1981). The second space order semi-discrete scheme is discretized in time by a nonlinearly stable two step Runge-Kutta time discretization (Shu & Osher 1988) leading to a

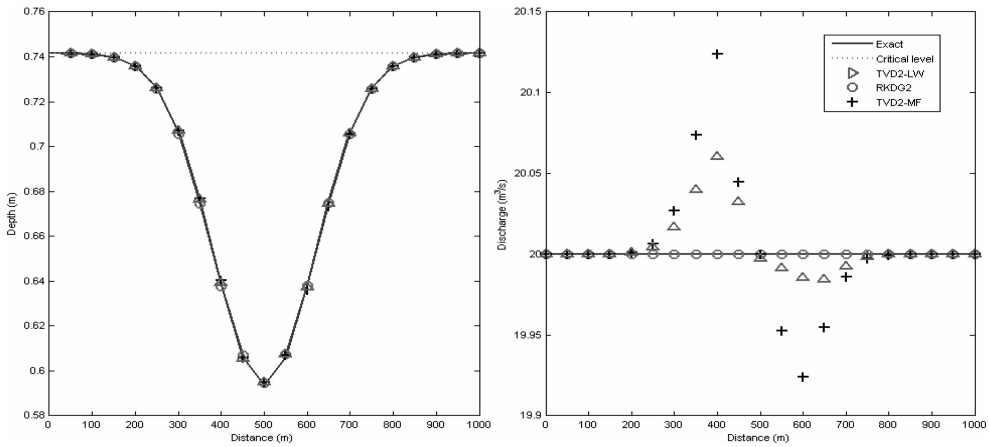


Figure 2. Depth and discharge profiles for the supercritical flow problem simulated with the RKDG2 scheme and compared with the two traditional second order TVD schemes.

second order accuracy scheme in time and space where the corresponding CFL number was chosen equal to 0.333.

It is well know in one-dimensional numerical modelling that the numerical scheme is only defined for the interior cells, hence boundaries must be defined. If the flow is subcritical, the interval of dependence (Hirsch 1990) for the upstream most point on the channel is somewhat upstream from the boundary point (Fig. 1a). Thus, estimation of flow conditions at this boundary point requires information about the flow conditions upstream of the boundary. The $C+$ trajectory from the upstream points affects the flow at that point in the $(x - t)$ plan; therefore, a single condition must be specified at the boundary point. This kind of BC can be called physical BC. As the C- curve leaves the domain, the influence region belongs to the computational domain and the information linked to that characteristic curve depends exclusively on the interior points. Again, at the downstream boundary, part of the interval of dependence falls outside the channel length being analyzed, and a physical downstream BC must be supplied. A numerical boundary condition is to be integrated from the Saint Venant characteristic form along the $C+$ characteristic curve (García-Navarro & Savirón 1992). If the flow is supercritical at a boundary, the required number of conditions changes. At an upstream boundary, the flow and the elevation of water must both be supplied because, as shown in figure 1b, the interval of dependence of points on the upstream boundary in the $(x - t)$ plan are outside the length of channel analyzed. At the downstream end, no BC are required because the interval of dependence falls within the length of channel analyzed. In this situation, numerical BC are calculated by differentiating the Saint Venant characteristics form along both of the $C+$ and $C-$ curves.

4 DISCUSSIONS AND NUMERICAL INVESTIGATION

Many existing references point out the advantages of employing the RKDG schemes. Zhou & al. (2001) presented a quantitative comparison of a third order FV-WENO method and a third order discontinuous Galerkin scheme, together with a third order Runge-Kutta time discretization. The researchers' numerical results showed that to achieve the same error magnitude, the RKDG scheme usually require fewer spatial mesh points and less CPU-time cost than the FV-WENO method, although a CFL number of 0.6 was chosen for WENO and a CFL number of 0.2 was chosen for the RKDG scheme. Xing & Shu (2006a) considered high order WENO schemes and RKDG schemes. Authors claimed that the RKDG method is by far a straightforward approach

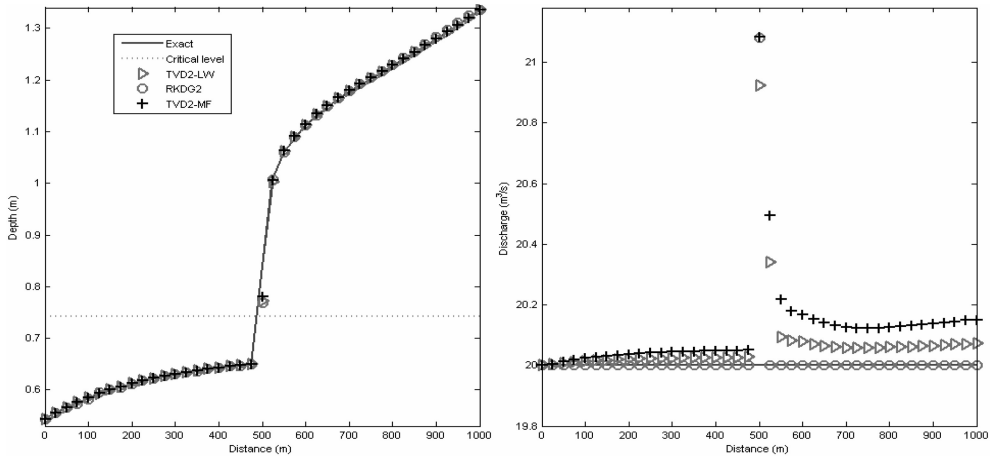


Figure 3. Flow variables for the steady hydraulic jump problem where the results of the RKDG2 scheme are compared with the results performed by the two traditional second order TVD schemes.

to obtain a well-balanced numerical scheme. Further, the researchers flaunted that the traditional RKDG method can achieve the well-balanced property with a small modification of either the initial value or the flux.

To illustrate the benefits of the RKDG2 scheme for the open channel flow simulation, two steady flow problems with analytical solutions were taken from MacDonald & al. (1997): a wholly supercritical flow; and a partially supercritical flow changing to subcritical flow via a formation of a hydraulic jump. The unsteady model, computed with the RKDG2 scheme, was given steady BC and the limiting steady solution was compared with the exact solution.

The steady discharge was equal to 20 m^3/s, bed and friction slopes were not equal to zero. Both of the water depth and the flow discharge were imposed at the inflow as physical BC. At the outflow, one physical boundary value (the flow discharge) was specified for the hydraulic jump problem. The other outflow boundary values were integrated numerically. In the computation, 21 and 41 cells were chosen, respectively, for the supercritical flow test and the hydraulic jump test.

The widely used second order Roe's extension TVD version of the Lax-Wendroff (TVD2-LW) scheme (Hirch 1990, p. 562) and the second order Roe's extension TVD upwind modified-flux approach (TVD2-MF) (Delis & Skeels 1998) – both implemented with a simple pointwise source vector discretization – were invoked to be compared with the RKDG2 method.

Figure 2 contains the simulation results of the first test problem. Analyzing the supercritical flow region, a poor maintenance of the discharge balance was carried out by each of the TVD2-MF and the TVD2-LW. On the contrary, a perfect agreement with the analytical solution is noticed for the discharge profile corresponding to the RKDG2 scheme. As well for the hydraulic jump problem, exposed in the figure 3, diffusive discharges profiles were marked by both of the traditional second order TVD schemes, seriously after the transition to subcritical flow regime.

5 CONCLUSION

In this work, we gave a brief description of the RKDG2 scheme. The BC treatment with the method of characteristics was also reported. Advantages of the RKDG2 method according to literature reports; and to a particular investigation for the SWE were discussed. The numerical results carried out by the RKDG2 scheme were compared with the results performed by two tradition second order TVD schemes jointly with a reference solution. With a small number of computational cells, the RKDG2 method led to a suitable resolution for the non-idealized simulated problems. Hence,

it can be concluded that the addressed scheme is accurate, well-balanced and can be of practical consideration for complex flow field such as combining channel branches at a junction where propagating accurate BC is required.

REFERENCES

Burguete, J. & García-Navarro, P. 2001. Efficient construction of high-resolution TVD conservative schemes for equations with source terms: application to shallow water flows. *Int. J. Numer. Meth. Fluids* 37(2): 209–248.

Cockburn, B. & Shu, C.W. 2001. Runge–Kutta discontinuous Galerkin method for convection-dominated problems. *J. Sci. Comput.* 16(3): 173–261.

Črnjarić-Žic, N., Vuković, S. & Sopta, L. 2004. Balanced finite volume WENO and central WENO schemes for the shallow water and the open-channel flow equations. *J. Comput. Phys.* 200(2): 512–548.

Delis, A.I. & Katsaounis, Th. 2003. Relaxation schemes for the shallow water equations. *Int. J. Numer. Meth. Fluids* 41(7): 695–719.

Delis, A.I. & Skeels, C.P. 1998. TVD schemes for open channel flow. *Int. J. Numer. Meth. Fluids* 26(7): 791–809.

García-Navarro, P. & Savirón, J.M. 1992. McCormack's method for the numerical simulation of one-dimensional discontinuous unsteady open channel flow. *J. Hydr. Res.*, 30(1): 95–105.

Hirsch, C. 1990. *Numerical computation of internal and external flows, in Vol. 2, Computational methods for inviscid and viscous flows.* Wiley: Chichester.

MacDonald, I., Baines, M.J., Nichols, N.K. & Samuels, P.G. 1997. Analytical benchmark solutions for open-channel flows. *J. Hydr. Engrg.* 123(11): 1041–1045.

Xing, Y. & Shu, C.W. 2006a. A new approach of high order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms. *Commun. Comput. Phys.* 1(1): 101–135.

Xing, Y. & Shu, C.W. 2006b. High order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms. *J. Comput. Phys.* 214(2): 567–598.

Zhou, T., Li, Y. & Shu, C.W. 2001. Numerical comparison of WENO finite volume and Runge-Kutta discontinuous Galerkin Methods. *J. Sci. Comput.* 16(2): 145–171.

Roe, P.L. 1981. Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comput. Phys.* 43(2): 357–372.

Shu, C.W. & Osher, S. 1988. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comput. Phys.* 77(2): 439–471.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Use of a nested modeling framework to study multiscale hydrological processes

J.E. Hunink
Técnica y Proyectos S.A., Barcelona, Spain

W. Bouten & E.E. van Loon
Universiteit van Amsterdam, Institute of Biodiversity and Ecosystem Dynamics, Amsterdam, The Netherlands

ABSTRACT: Within hydrology the trend is to integrate different mathematical models coming from related disciplines. These models operate most often on different scales. Therefore, there is a need for a modeling environment in which stand-alone models can operate simultaneously on the corresponding scales with possible cross-scale interference. We propose a framework that takes advantage of the inherent nested spatial character of hydrological processes and lets the process knowledge exactly operate on the scale it was studied. We show that with this multi-scale framework, nonlinear hydrological response can be simulated and argue that using a nested modeling strategy, our understanding of hydrological response can be improved and cross-scale interaction of multiscale controls can be investigated.

1 INTRODUCTION

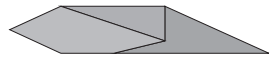
Especially hydrology is a multi-disciplinary field of study with observations and model descriptions on different scales. Scale-issues create a challenge for many hydrological scientists as it is inevitable to link observations and models on different scales (Philips, 2005). When process knowledge and observations are transferred to a mathematical model, scientists tend to focus on just one scale supposing that the 'first principles' or the main driving forces processes reside on one specific scale (Blöschl 2001). As a consequence, in hydrological modeling studies a-priori assumptions are made about the cross-scale linkages of processes. Most hydrological models are scale-dependent (Reed et al. 2004), hence no insight can be obtained through the modeling analysis about how processes influence each other across scales and how scale-specific nonlinearities and thresholds affect other scales.

Besides, the tendency in hydrological modeling is to couple model descriptions coming from different related disciplines (e.g. European OpenMI project). This implies inevitably having to deal with scale-issues. Scale problems are found for example in the coupling of overlandflow and routing models with sediment and chemical transport models (e.g. Simpson and Castelltort, 2006).

One of the biggest challenges in hydrological studies in is dealing with nonlinearities and thresholds. Nonlinear system response is thought to be related with patterns and multiple controls acting on multiple scales (Sivapalan, 2005). The influence of nonlinearities and thresholds on a certain scale on the system response has been studied in field- and modeling studies (e.g. Kokkonen et al. 2004). Especially in semi-arid areas, the nonlinear relationship between rainfall and runoff makes the modeling process more complex (Ye et al. 1997) and nonlinear (Goodrich et al. 1997).

In hydrology, nested approaches can be found in field and theoretical studies (e.g. Cammeraat, 2004); however, in modeling little has been done so far. The semi-distributed approaches (e.g. HEC-HMS) can be considered close to nested modeling but no real nesting on different hierarchical levels

Table 1. Illustration of framework concepts: the scale domain is discretized in scale-levels, each one having corresponding spatial and temporal boundary values and 1 or more models applied to it.

Scale-level	Spatial discretization	Timestep	Models
1		1 year	Model X & Model Y
2		1 day	Model Y
3		1 min	Model Z & Model W

with different model descriptions is possible. A promising hierarchical approach is the use of spatial object hierarchies in GIS (Band et al. 2000).

This study proposes a nested modeling framework through which different scale-dependent hydrological models are able to operate simultaneously on different scales and communicate with each other. The framework counts with a hierarchical tree structure for the state variables and model parameters of the different models. This data structure makes use of the nesting capabilities of cell and structure arrays in MATLAB. Every hierarchic level in the data structure corresponds to a scale-level and every node to one of the spatial model elements. To every scale-level one or more models can be applied.

2 FRAMEWORK CONSIDERATIONS AND DETAILS

The notion we have of the hydrological structure of a watershed is an interconnected network of streams with contributing areas on both sides, resulting in a mosaic being composed of many nested sub-catchments. This nested hierarchy and the relation between the spatial and temporal scale are the main piles of the modeling framework. Some framework concepts are schematized in Table 1.

Besides coordinating the output analysis and visualizing the data, the framework accomplishes three main tasks; shortly commented in the following sections.

2.1 *Multiscale time frames*

The framework functions as a manager handing out the jobs to his workers and tells the models between which two time instants the job has to be finished. Within these temporal boundary values, the models are able to work with their own timesteps. The framework uses a dynamic function with a recursive algorithm that handles the different scale-dependent time intervals.

2.2 *Data transfer to the models*

The framework provides the models with the correct data of the spatial model elements. This task requires a common data structure in which the state variables and model parameters of the different models can be stored. The data structure needs to preserve the multi-scale and multi-model character of the framework. Therefore, a so-called tree cell array was designed in MATLAB, forming the backbone of the framework. For every scale-level it passes the right model parameters and state variables to the models. This approach has the advantage that every spatial sub-element can be taken by the models as an autonomous functional sub-watershed.

2.3 *Inter-model communication*

One of the main objectives of the framework is to offer a clear communication structure to the models. The framework counts with the necessary message functions that exchange and update data with

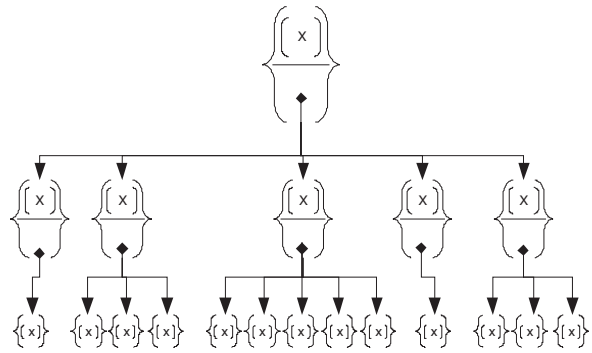


Figure 1. Schematic representation of a tree cell array with three nested hierarchic levels. The brackets {} represent nested cell arrays and [x] symbols the arrays with the data.

the tree cell array. It has to be noted that the modeling framework already passes directly the right parameters and variables to the models for every timestep. Only when inter-model communication is necessary, these function calls are required.

3 CASE STUDY: A NESTED OVERLANDFLOW MODEL

A basic multiscale framework application was carried out to show the capability of the framework to reproduce a nonlinear system response. Two models were implemented in the framework, one being a simple description of Hortonian overlandflow, and the other of detention storage on different scale-levels. No calibration or comparison with real data was done as the only objective was to test whether the behavior of the system.

Three scale-levels were distinguished. On the finest scale-level both models were applied (infiltration and detention storage) and on the other 2 coarser scale-levels only the detention storage model. The case study uses an arbitrary spatial discretization with the corresponding tree cell array having 1 element on the first scale-level, 10 elements on the second and 100 on the 3rd scale-level. These elements were randomly connected with each other (similar to Figure 1).

The infiltration model is based on the Green-Ampt function:

$$f = A + \frac{B}{S_i} \tag{1}$$

in which f is the infiltration capacity, A and B are constants and S_i the infiltrated water volume.

The detention storage model consists of a conceptual bucket for every spatial model element, of which the outflow can be described as:

$$Q = \frac{\max\left(S_d - S_{d,\max}, 0\right)}{\Delta t} \tag{2}$$

in which Q is the outflow, S_d the detention storage volume, $S_{d,max}$ the maximum storage volume and Δt the integration timestep.

The relation of the input magnitude of the system with the output response was examined multiplying a 60-timestep design rainfall record with different scaling factors. This was compared with the system output being the overflow of the detention storage element on the coarsest scale-level.

Figure 2 shows the input variable (rainfall intensity) and the output depending on the input magnitude. The system starts to react with a relative magnitude of 0.4. The output is clearly non-proportional to the input in terms of volume and amplitude of the output dynamics. The system

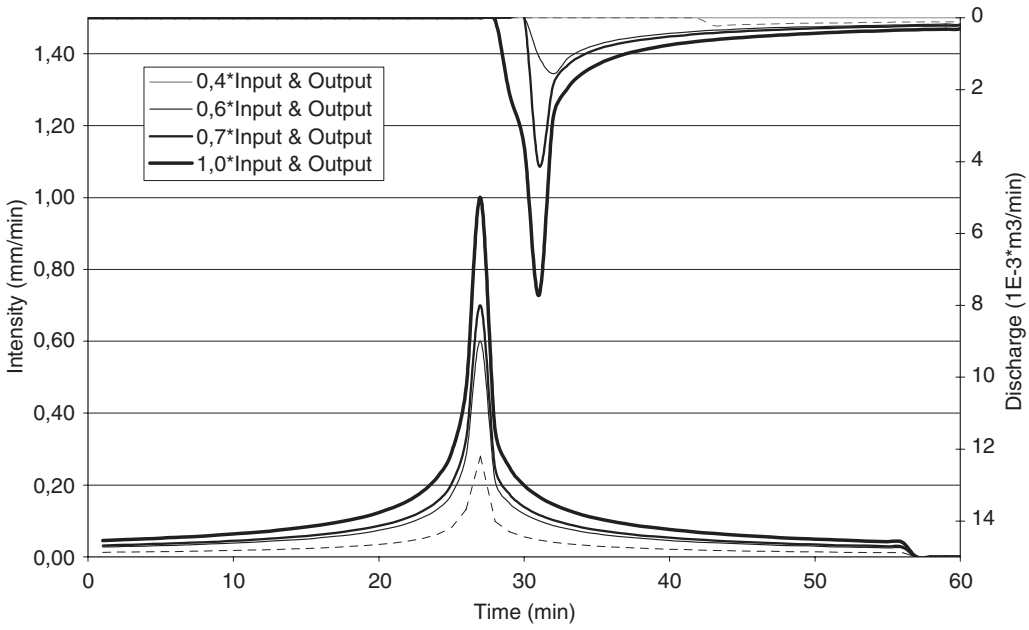


Figure 2. Comparison of the system output with varying input magnitudes.

shows a strong and fast transition being the result of the cumulative effects of the finer scale system controls.

4 CONCLUSIONS

The framework counts with a hierarchical tree data structure with the corresponding interface functions. The main advantage of this tree data structure is its flexibility for multiscale discretizations and the fact that it is capable to retain the nested hierarchy of hydrological processes itself.

The framework assigns to every scale-level a different temporal discretization. Within these time frames, the framework passes exactly the information of the right node in the data tree to the models. Inter-model communication is possible through a message-type function of the framework.

To test the framework, a simple application was done with two models related to overlandflow. Due to the multiscale character of the system with threshold controls acting on different scales, nonlinearity was demonstrated between input magnitude and response dynamics.

The next step in the investigation will be the development of an experimental design for the observation input parameters in order to carry out a real application of the framework. The questions that will have to be answered is how to make use of hierarchical measurements and how to perform a multiscale calibration. This study indicated that a nested hydrological modeling framework can serve as an intuitive tool for the coupling of multiscale models and studying nonlinear hydrologic responses.

REFERENCES

Band, L.E., Brun, S.E., Fernandes, R.A., LTague, C.L. and D.E. Tenenbaum (2000). “Modelling Watershed as Spatial Object Hierarchies: Structure and Dynamics” Transactions in GIS 4(3): 181–196.

Blöschl, G. (2001). “Scaling in hydrology.” Hydrological Processes 15(4): 709–711.
Cammeraat, E.L.H. (2004). “Scale dependent thresholds in hydrological and erosion response of a semi-arid catchment in southeast Spain.” Agriculture Ecosystems & Environment 104(2): 317–332.
Goodrich D.C., Lane L.J., Shillito R.M., Miller S.N., Syed K.H. and D.A. Woolhiser (1997) “Linearity of basin response as a function of scale in a semiarid watershed” Water Resources Research 33(12): 2951–2965.
Kokkonen, T., Koivusalo, H., Karvonen, T., Croke, B. and A. Jakeman (2004). “Exploring streamflow response to effective rainfall across event magnitude scale” Hydrological Processes 18(8): 1467–1486.
Philips, J.D. (2005). “Entropy Analysis of Multiple Scale Causality and Qualitative Causal Shifts in Spatial Systems.” The Professional Geographer 57(1): 83–93.
Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo D.J. and DMIP-Participants (2004). “Overall distributed model intercomparison project results.” Journal of Hydrology 298(1–4): 27–60.
Simpson, G. and S. Castelltort (2006) “Coupled model of surface water flow, sediment transport and morphological evolution” Computers & Geosciences 32(10): 1600–1614.
Sivapalan, M. (2005). “Pattern, Process and Function: Elements of a Unified Theroy of Hydrology at the Catchment Scale” Encyclopedia of Hydrological Sciences. Edited by M.G. Anderson. Johan Wiley & Sons, Ltd.
Ye, W., Bates, B.C., Viney, N.R., Sivapalan M. and A.J. Jakeman (1997). “Performance of conceptual rainfall-runoff models in low-yielding ephemeral catchments.” Water Resources Research 33(1): 153–166.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Furrow fertigation simulation model, calibration, validation and application

N. Zapata
Soils and Irrigation Department, Agro-Food Research Center, Zaragoza, Spain

J. Burguete
Soil and Water Department, EEAD-CSIC, Zaragoza, Spain

E.P. Santos
Fluid Mechanics Department, CPS-University of Zaragoza, Zaragoza, Spain

E. Playán
Soil and Water Department, EEAD-CSIC, Zaragoza, Spain

P. García-Navarro
Fluid Mechanics Department, CPS-University of Zaragoza, Zaragoza, Spain

ABSTRACT: Furrow fertigation can be an advantageous practice over traditional methods for fertilizer application. Fertigation efficiency and uniformity are tightly related to irrigation water efficiency and uniformity. This work presents a simulation model of furrow fertigation that was calibrated and validated with experimental data. The model predicts water flow and solute transport throughout Saint-Venant and advection-dispersion equations. Values for the three semi-empirical parameters of the model were obtained through a calibration process. The model can reproduce water advance times, flow depths and fertilizer hydrodynamics adequately, particularly for medium and large irrigation discharges.

1 INTRODUCTION

The numerical simulation of hydrodynamics is a very common technique for irrigation analysis, from conveyance networks to on-farm systems. In this paper, the application of these techniques to furrow irrigation fertigation is presented. The application of chemical fertilizers to surface irrigated crops is often hindered by the crop canopy, which can make the use of tractors and fertilizer machinery impossible. Under these circumstances, fertigation – the application of chemicals in the irrigation water – becomes a practical alternative. Several authors have addressed this problem using simulation approaches in borders and basins (Playán & Faci, 1997; García-Navarro *et al*., 2000) and in furrows (Zapata *et al*., 2005; Strelkoff *et al*., 2006). The objective of this paper is to simulate furrow fertigation using an infiltration equation accounting for the furrow wetted perimeter and incorporating the model proposed by Rutherford (1994) to describe the chemical dispersion coefficient. Experimental results are used to calibrate the semi-empirical model parameters (infiltration and roughness) and to validate the proposed model.

2 MATERIALS AND METHODS

For the development of the furrow fertigation simulation model it was supposed that the furrow irrigation flow is one-dimensional and unsteady, hence governed by the St. Venant (Cunge *et al.*,

1980) or one-dimensional shallow water equations that are frequently formulated as:

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = -Pi; \quad \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x}\left(\frac{Q^2}{A} + gI_1\right) = gA\left(S_0 - S_f\right); \quad \frac{\partial A_i}{\partial t} = Pi \tag{1}$$

where A is the transversal wetted area of the furrow, Q is the irrigation discharge, P is the wetted perimeter, i is the infiltration rate, A_i is the volume of infiltrated water per unit of furrow length, g is the acceleration due to gravity, S_0 is the bed slope and S_f is the friction slope (Cunge *et al.*, 1980). The last term was modeled by the Manning equation:

$$S_f = \frac{n^2|Q|QP^{4/3}}{A^{10/3}} \tag{2}$$

with n Manning roughness coefficient, and finally, I_1 represents the hydrostatic pressure integral over the wetted area:

$$I_1 = \int_0^h (h - z)\sigma \, dz \tag{3}$$

with h representing water depth and σ being the furrow width at a height z, referred to the bottom of the furrow section.

Using the Kostiakov approach, the cumulative infiltration Z was computed as a function of two experimental parameters, K and a, dependent on soil type characteristics. These parameters should be calibrated with field experiments.

$$Z = K\tau^a, \quad i = \frac{dZ}{dt} = aK\tau^{a-1} \tag{4}$$

In (1), the independent variable (τ) is the opportunity time that corresponds to the wetted time of each section. This approximation could introduce serious errors when estimating the infiltrated volume and the advance phase, since there is no reference to the furrow wetted perimeter. To account for this peculiarity of furrow infiltration, the model proposed by Maikaka (2004) was used. In this model the infiltration rate is dependent on the furrow wetted perimeter (P) and on the furrow wetted area (A_i), and the opportunity time is not explicitly present:

$$\frac{\partial A_i}{\partial t} = PaK\left(\frac{A_i}{KP'}\right)^{\frac{a-1}{a}} \tag{5}$$

where A_i represents the infiltrated area, P the wetted perimeter and P' the furrow spacing. As described in Maikaka (2004), this infiltration model is based on the experimental parameters K and a of the traditional Kostiakov equation.

The fertigation model is based on the following solute transport equation:

$$\frac{\partial(AC)}{\partial t} + \frac{\partial(QC)}{\partial x} = -PiC + \frac{\partial}{\partial x}\left(K_x A \frac{\partial C}{\partial x}\right) \quad \frac{\partial \phi}{\partial t} = PiC \tag{6}$$

where C is the average solute concentration at the section, ϕ is the solute infiltrated mass per unit of furrow length and K_x is the dispersion coefficient (strictly positive), accounting for molecular diffusion, turbulent diffusion and the discordance resulting from using the average section approximation. The model proposed by Rutherford (1994) was used to describe the dispersion coefficient:

$$K_x = 10\sqrt{gPAS_f} \tag{7}$$

The Saint Venant equations and the advection-dispersion equation result in a system of three first-order, non-linear, hyperbolic, partial derivative equations. Initial and boundary conditions (discharge, water depth and solute concentration) are needed to solve the governing equations. The initial conditions were uniform zero discharge and water depth. The transport equation was initialized with a solute surge distribution, rectangular in shape at the solute application point.

The model was used to simulate water advance and solute concentration. Three semi-empirical parameters were needed: the Manning roughness coefficient (n) and the infiltration parameters (K and a).

The experimental data were taken from Zapata *et al*. (2005). A summary of the experimental layout is presented. Four isolated furrows were built using the same field machinery. The furrow cross-sectional geometry was hand measured, the furrow length was 100 m and the furrow spacing was 1 m. Each furrow was fertigated with a different inflow discharge (1, 2, 3 and 4 L s^{-1}; Q1, Q2, Q3 and Q4, respectively), the same fertilizer concentration and under free draining conditions. The fertilizer was applied following the same procedure for all four experiments.

The furrows were monitored to follow the hydrodynamics of water and fertilizer during the furrow irrigation evaluations. Stations were marked along each furrow every 10 m. The advance time to each station and the evolution of the upstream flow depth were registered. In five stations – every 20 m (E20, E40, E60 and E80) – fertilizer concentration was monitored at different times during the irrigation event by taking water samples for laboratory analysis.

Two of the furrow experiments (Q1 and Q4) were used for model calibration purposes; the other two (Q2 and Q3) were used for model validation. The calibration process pursued the identification of three semi-empirical parameters. The calibration consisted on simulating the experiments with different values of K, a and n, and comparing simulated furrow fertigation with the experimental measurements. The terms of comparison were irrigation advance phase, flow depth and fertilizer concentration at the stations. The following statistics were used for calibration purposes: the root mean square error (RMSE), the slope of the simple regression forced trough the origin (A) and the coefficient of determination of this regression model (R^2).The values of the three parameters that better fit field observations were selected as representative of the experiments. These values were used to simulate the validation experiments.

3 RESULTS AND DISCUSSION

Comparison statistics between observed and simulated water advance fronts and water flow depths were presented in Table 1. The advance phase was better simulated (lower RMSE) for the large inflow discharges (3 and 4 L s^{-1}). The selected semi empirical parameters ($K = 0.0048\ m^2 s^{-a}$, $a = 0.2$ and $n = 0.04\ m^{-1/3}$ s) represent a compromise between the calibration experiments. The large error for the experiment of the lowest inflow discharge (Q1) can be attributed to the error induced by the use of the Manning equation. Manning n was only considered dependent on surface

Table 1. Comparison between observed and simulated water advance fronts and upstream water depths.

	Process	Inflow discharge (L s^{-1})	RMSE	R^2	A
Calibration	Advance (min)	1	212.12	0.983	1.12
		4	44.07	0.998	1.08
	Flow depth (mm)	1	0.006	0.864	0.99
		4	0.015	0.982	1.13
Validation	Advance (min)	2	85.91	0.994	0.93
		3	23.49	0.997	1.01
	Flow depth (mm)	2	0.039	0.593	1.66
		3	0.024	0.908	1.25

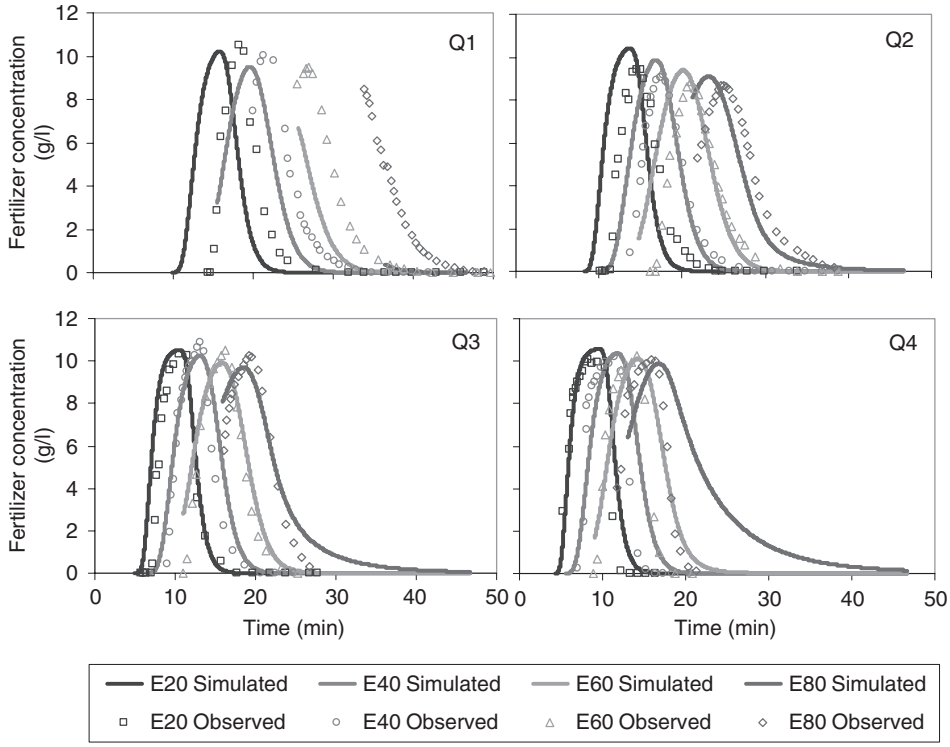


Figure 1. Observed (symbols) and simulated (lines) fertilizer concentration at each fertilizer station for the furrow experiments as a function of time.

roughness. However, other authors (Chow, 1959; Walker & Skogerboe, 1989) pointed out that this coefficient also depends on flow depth. This effect is particularly relevant for small discharges, resulting in low flow velocity and depth. In these conditions viscosity forces commence to be relevant and Manning law – based on turbulent flow – is far from being adequate.

Figure 1 presents a comparison between observed and simulated fertilizer concentrations for the experiments Q1, Q2, Q3 and Q4 at fertilizer stations E20, E40, E60 and E80. Fertilizer hydrodynamics were satisfactory reproduced for Q2, Q3 and Q4, with Q1 showing larger errors. The comparison statistics between observed and simulated data (not presented) confirmed the graphical analysis.

4 CONCLUSIONS

1. The field experiments were adequate to calibrate and validate the furrow fertigation model.
2. The infiltration process has been modeled following Maikaka (2004). This approach relates the infiltration parameters to the wet area.
3. The largest differences between observed and simulated process were obtained for the lower inflow discharge experiment. The Manning law approach can introduce serious error when the discharge is small since the viscosity forces may be relevant. This subject needs to be studied in further research.
4. Solute diffusion was been modeled following Rutherford (1994) obtaining very reasonable results.

REFERENCES

Chow, V. T. (1959). Open Channel Hydraulics, McGraw-Hill Book Co. Inc.

Cunge, J. A., Holly, F. M. & Verwey, A. (1980). Practical aspects of computational river hydraulics. Ptman Publising Ltd. London,.

García-Navarro, P., Playán, E. & Zapata, N. (2000). Solute transport simulation in overland flow: application to fertigation. J. Irrig. and Drain. Engrg., ASCE, 126(1):36–40.

Maikaka, M. (2004). Tesis Doctoral. Modelos Numéricos Unidimensionales en Riego por Superficie. Universidad de Zaragoza.

Playán, E. & Faci, J. M. (1997). Border fertigation: field experiments and a simple model. Irrig. Sci, Springer-Verlang. 17: 163–171.

Rutherford, J.C. (1994). River Mixing, Wiley.

Strelkoff, T. S., Clemmens, A. J. & Perea-Estrada H. (2006), Calculation of non-reactive chemical distribution in surface fertigation: Agricultural Water Management, 86: 93–101.

Walker, W. R. & Skogerboe, G. V. (1987). Surface irrigation. Theory and practice. Prentice-Hall, Inc., Englewood Cliffs, New Jersey. 386 pp.

Zapata, N., García-Navarro, P., Burguete, J., Maikaka, M. & González, S. (2005). Calibración y Validación de un Modelo de Fertirriego en surcos. Primeros resultados. Congreso Nacional de Riegos y Drenajes.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Numerical methods for dispersion in water distribution pipes and networks

V.G. Tzatchkov
Mexican Institute of Water Technology, Jiutepec, Morelos, Mexico

ABSTRACT: A review of recent developments in numerical modeling of unsteady dispersion in water distribution pipes is presented. The hydrodynamic process is described by the two-dimensional advection-dispersion-reaction equation. Although this equation can be numerically solved, such a solution is impractical for modeling the dispersion in networks, so that the equivalent one dimensional equation is used. A numerical solution for pipe networks is then proposed, based on numerically computed Green functions. Thus, the large system of equations that represents the discretized network is decomposed exactly in three easy-to-solve tridiagonal systems for each pipe, and one low order system for the concentration at the pipe junctions. In each pipe the solution is represented by the superposition of a homogeneous (zero boundary conditions) solution, and two numerically obtained Green function solutions. Henceforth, continuity balance relations are used to construct a system of linear equations for the unknown quantities at the network nodes. Comments on the applicability of the proposed model in laminar turbulent flow pipes and are presented.

1 INTRODUCTION

Distribution network water quality models are useful in network design and operation. Such models can be used to analyze water quality degradation problems, to assess alternative operational and control strategies for improving and maintaining water quality, to design water-quality-sampling programs, to optimize disinfection processes and to evaluate water quality aspects of distribution network improvement projects. Most distribution network water quality models, e.g., the widely used public domain EPANET model and known commercial models neglect dispersion. Those models works well for the main network arteries, where flow velocity is medium to high, and flow is turbulent. In pipes along the network periphery, however, sporadic water demands from sparse downstream users may lead to random intermittent laminar flows where stagnation is common (Buchberger *et al.* 2003). Recent research has shown that while unimportant in turbulent flow, the dispersion component of the transport process is important in laminar flow pipes (Tzatchkov *et al.* 2000, Tzatchkov *et al.* 2002, Li *et al.* 2005). Li *et al.* (2005) analyzed the relative importance of the three basic mechanisms on the overall transport process in laminar flow and found that dispersion accounts for more then 28 percent of that process. This paper resumes some developments of considering dispersion in water distribution pipe network numerical models.

2 BACKGROUND

Transport of a non-conservative chemical tracer moving in steady laminar flow through a pipe is described with the two-dimensional advection-diffusion equation:

$$\frac{\partial C}{\partial t} = D\left(\frac{\partial^2 C}{\partial r^2} + \frac{1}{r}\frac{\partial C}{\partial r} + \frac{\partial^2 C}{\partial x^2}\right) - 2U\left(1 - \frac{r^2}{a^2}\right)\frac{\partial C}{\partial x} \tag{1}$$

where $C = C(r, x, t)$ is the solute concentration at any point in the cross section, D is the molecular diffusion coefficient (diffusivity) of the solute in water, U is the mean velocity in the axial direction, K is the first-order reaction rate constant for bulk decay (no wall decay is considered), a is the pipe radius, r is radial position, x is axial position, and t is time.

According to the classical Taylor's theory for dispersion Equation 1 can be simplified to the one-dimensional advection-dispersion equation:

$$\frac{\partial C}{\partial t} + U\frac{\partial C}{\partial x} = E\frac{\partial^2 C}{\partial x^2} - KC \tag{2}$$

where E is the dispersion coefficient:

$$E = \frac{U^2 a^2}{48D}, \quad \text{for laminar flow} \tag{3a}$$

$$E = 10.06 a U\sqrt{\frac{f}{8}}, \quad \text{for turbulent flow} \tag{3b}$$

where f in Equation 3b is the Moody friction factor. Because of the small value of the diffusivity (D is about 10^{-9} m^2/s for salt or chlorine in water), Equation 3a gives values for E much larger than those computed by Equation 3b. For example, for a 10 cm diameter pipe with 0.01 m/s flow velocity (corresponding to laminar flow) Equation 3a gives E = 20.83 m^2/s, while for the same pipe and a higher 1 m/s flow velocity (corresponding to turbulent flow), and $f = 0.016$, Equation 3b gives E = 0.045 m^2/s. Being the longitudinal dispersion in pipe flow produced by transversal mixing due to non uniform cross sectional flow velocities, the physical meaning of this theoretical result is expressed by the fact that flow velocity distribution is more uniform in turbulent pipe flow, compared to laminar flow.

In steady flow, the 1-D model in Equation 2 with constant dispersion coefficient E (steady dispersion) works well provided a certain initialization period has elapsed. Expressed in dimensionless time T, the initialization condition is $0 < T < 0.5$, where $T = D \cdot t/a^2$. Because of the small value of the diffusivity and the constantly changing flow, in water distribution networks the dispersion process is always in the initialization period (unsteady dispersion).

In a network of pipes, Equation (2) applies for each pipe, and the following boundary conditions hold at the network nodes:

a) The concentration C is prescribed at some nodes, such as constituent sources.
b) Mixing at the network nodes. Two or more pipes, each of them with different flow and constituent concentration, may convey flow to a node. Water is mixed at the node and a new concentration is obtained, then water leaves the node with that concentration to the outflowing pipes and to the consumption abstracted. A complete mixing is usually assumed in the network models.
c) Mass conservation at the network nodes. Equation (2), is obtained by applying the mass conservation principle to an elementary pipe segment with a length dx in a time period dt, such that $dx = u.dt$. Tzatchkov *et al.* (2002) generalized this derivation to the case of a junction where several pipes meet, obtaining the following *nodal equivalent* of Equation (2):

$$\sum_{j=1}^{m}\left(\frac{dx_j}{2}A_j\right)\frac{\partial C}{\partial t} = \sum_{j=1}^{m}\left(A_j D_j\frac{\partial C}{\partial x} + Q_j C - K_j C A_j\frac{dx_j}{2}\right) - q_i\, C \tag{4}$$

where m = number of pipes connected to the node; A_j = cross sectional area of pipe j; Q_j = flow rate in pipe j; D_j = dispersion coefficient of pipe j; K_j = first order decay constant of pipe j; and q_i = flow rate abstracted at the node. For the case where two pipes of equal characteristics meet at the node and $q_i = 0$, Equation (4) reduces to Equation (2) if dx_j and dt tend to be infinitesimally small.

3 NUMERICAL METHODS

Numerical solutions for Equation 1 in steady and random intermittent unsteady laminar flow are obtained by Li *et al.* (2005). A Lagrangian-Eulerian numerical scheme was employed splitting Equation 1 is into two equations, one for molecular diffusion and another for advection. Each advected point retains its radial position but is translated downstream by an amount equal to the computed travel distance. Travel distance is computed as the product of the flow velocity and the time step. Radial and axial diffusion are computed with an Eulerian approach. In intermittent flow, during a busy period only radial diffusion is considered, using an unconditionally stable Crank-Nicholson scheme, because axial diffusion is negligible compared to advection. During idle times, both axial and radial diffusion are simulated, applying an unconditionally stable alternating direction implicit scheme using two half-steps to advance one full time step.

Large water distribution systems often contain hundreds of dead-end stems serving thousands of consumers. Considering the massive computational burden, it is impractical to apply a 2-D transport model to simulate water quality in all branching pipes of a municipal network. Most network water quality models simulate solute transport with pipe hydraulics based on 1-D plug flow. Hence, the approximation for unsteady laminar dispersion is merged with the 1-D model given by Equation (2), with an equivalent dispersion coefficient E for intermittent unsteady laminar flow (Li *et al.* 2005).

The numerical solution in networks poses three main problems:

a) Boundary conditions at the nodes common to several domains have to be formulated and considered.
b) The direct application of the numerical schemes produces large non-banded, unsymmetric and unstructured systems of equations to be solved, especially when the network is large.
c) The difficulty increases when advection dominates over dispersion. Sharp concentration gradients are expected in this case, and an extremely fine discretization would be needed if Eulerian methods are to be applied.

Tzatchkov *et al.* (2002) develop a numerical solution for the advection-dispersion-reaction equation in pipe networks, based on the domain decomposition strategy used to efficiently solve the resulting finite difference equations. A two-stage Eulerian-Lagrangian numerical scheme is applied. An implicit numerical scheme is used in the Eulerian stage leading to a system of linear equations. To achieve computational efficiency in the solution of this system, a special technique that employs numerically computed Green's functions within each pipe is proposed.

Equation (2) is split in two parts, an advective part and a dispersive part, and numerically solved in each time step in two stages. In the first (Lagrangian) stage the advective (or advective-reactive) part is solved for each pipe. The backward method of characteristics is used.

In the Eulerian stage, the direct application of a finite difference approximation to a network of pipes produces a non banded system of linear equations, not amenable for an efficient numerical solution. To overcome these difficulties, Aldama *et al.* (1998) proposed an efficient numerical solution of boundary value problems in networks, called *the numerical Green's function technique*. Using Green's functions, the solution inside each network pipe is expressed by the sum of a *homogenous solution* $H(\xi)$ and two Green's function $G(\xi, R)$ and $G(\xi, F)$ (one for each pipe end R and F), computed from tri-diagonal systems of linear equations. The first of them accounts for initial conditions, and the other two for boundary conditions. The same matrix is used for each of them. The two Green's functions are symmetric, so only one of them needs to be computed.

The expression for C is then substituted in the balance condition at the network nodes (Equation 4) obtaining a system of equations for the values of C at the network nodes. The matrix of the system of equations for the network nodes is symmetric and sparse, and reflects the structure of the network itself: for each network node there is a row in the matrix whose non zero elements correspond to the diagonal and to the nodes to which the current node is connected. Efficient sparse matrix

algorithms can be applied to store and to solve systems of equations with this type of matrix. Thus the large non banded systems of equations, that would otherwise produce the direct application of the finite difference scheme for the network, is decomposed in three easy-to-solve systems for each network reach and a much smaller sparse system for the network nodes, and the solution can be computed more efficiently, especially for large networks. More information can be found in Tzatchkov *et al.* (2003), Tzatchkov *et al.* (2000), Tzatchkov *et al.* (1998), Tzatchkov *et al.* (2002), and Tzatchkov *et al.* (2001).

4 MODEL APPLICABILITY

The proposed model is equally applicable to laminar, transitional and turbulent flow, which is important for its application in water distribution networks, where flow is laminar in some pipes and turbulent or transitional in others. The transport process normally is dispersion dominated in laminar flow pipes, and advection dominated in turbulent flow pipes. Because of its Eulerian-Lagrangian nature, the proposed numerical method works well in both cases. Eulerian-Lagrangian methods are proven to be stable and sufficiently accurate over a wide range of the Peclet number *Pe*. The model is applied to real water distribution networks, for which field observations and results from an advection-reaction model are available (Tzatchkov *et al.* 1998, Tzatchkov *et al.* 2000, Tzatchkov *et al.* 2001, Tzatchkov *et al.* 2002), Tzatchkov *et al.* 2003, and Li *et al.* 2006). In these networks the mesh Peclet number ($Pe = U.\Delta x/E$) ranges from 0 to 5250. In general, the proposed model and the simpler advection-reaction model provide similar results for turbulent flow pipes, but in laminar flow pipes the predictions of the proposed model are much better than those from the advection-reaction model.

Thanks to the decomposition of the system of equations for the network in tri-diagonal systems for each pipe, the proposed solution is efficient in terms of computer time and storage and can be applied to networks with large number of pipes. While for the solution of the original full system the required storage is proportional to n^2 (where n is the total number of points in all pipes), and the computer time is proportional to n^3; in the proposed solution both of them are proportional (roughly) to $2n$.

By now, simple linear interpolation at the characteristics feet is used at the Lagrangian stage. Since this may lead to some undesirable numerical dispersion, future work is needed to determinate the most appropriate time step size and the influence of different interpolation schemes, given that these are well understood in one domain but are not studied in networks.

REFERENCES

Aldama, A.A., Tzatchkov, V.G. & Arreguin, F.I. 1998. The numerical Green's function technique for boundary value problems in networks. In W.R. Blain (ed), *Hydraulic Engineering Software VII:* 121–130. Southampton: Computational Mechanics Publications/WIT Press.

Buchberger, S.G. Li, Z. & Tzatchkov, V.G. 2003. Hydraulic characterization of pipe network subject to stochastic water demands, In C.A. Brebbia (ed), *Water Resources Management II:* 161–170, Southampton: WIT Press.

Li, Z., Buchberger, S.G. & Tzatchkov, V.G. 2005. Importance of dispersion in network water quality modeling. In *Proceedings of the ASCE World Water & Environmental Resources Congress, Anchorage, AL, May 15–19*.

Li, Z., Buchberger, S.G. & Tzatchkov, V.G. 2006. Integrating Distribution Network Models with Stochastic Water Demands and Mass Dispersion. In *Proc. 8th Annual Water Distribution System Analysis Symposium, University of Cincinnati, 27–30 August 2006*.

Tzatchkov, V.G., Aldama, A.A. & Arreguin, F.I. 2000. Modelación numérica de la advección y dispersión de solutos en redes de distribución de agua potable. *Ingeniería Hidráulica en México*, 15(3): 101–115 (in Spanish).

Tzatchkov, V.G., Aldama, A.A. & Arreguin, F.I. 2002. Advection-dispersion-reaction modeling in water distribution networks. *J. of Water Resources Planning and Management.* 128(5): 334–342.

Tzatchkov, V.G., Aldama, A.A. & Arreguín, F.I. 2003. A domain decomposition strategy for the numerical simulation of contaminant transport in pipe networks. In I. Herrera, D.E. Keys, O.B. Widlund, R. Yates (eds), *Domain Decomposition Methods in Science and Engineering:* 483–490, Mexico: UNAM.

Tzatchkov, V. G., Aldama, A.A., Arreguín, F.I., Buchberger, S.G & Lee Y. 2001. Advection-dispersion transport in water distribution networks with intermittent flow. In C.A. Brebbia, P. Anagnostopoulos, K. Katsifarakis & A.H-D. Cheng (eds) *Water Resources Management:* 217–226, Southampton: WIT Press.

Tzatchkov, V.G., Arreguin, F.I. & Aldama, A.A. 1998. An application of the numerical Green's function technique to advection-diffusion contaminant transport in water supply networks. In W.R. Blain (ed), *Hydraulic Engineering Software VII:* 587–596, Southampton: WIT Press.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

3D Modeling of hydrodynamics processes of lagoon-open sea interaction

D. Bellafiore & G. Umgiesser
Institute of Marine Sciences, ISMAR-CNR, Venice, Italy

ABSTRACT: In this paper some evidences on the 3D finite element hydrodynamic model SHYFEM development are presented to simulate the interaction lagoon-open sea processes. A study has been performed to identify the rate of impact of each forcing (tides, wind, freshwater, heat fluxes) in simulating hydrodynamic processes in coastal areas. A particular interest is for phenomena in proximity of rivers and inlets, such as the three inlets connecting the Venice Lagoon with the Adriatic Sea (Lido, Malamocco and Chioggia Inlets). The model has been checked by applying a test case from Chao and Boicourt (1986), reproducing the freshwater input of a river in an idealized basin. The test case permitted to assess the 3D behavior of water masses, in terms of advective, barotropic and baroclinic pressure gradients and diffusive contributions on circulation. In the test case baroclinic pressure gradient terms in the momentum equations are computed explicitly, while temperature and salinity gradients are applied as forcings. The real case run, applied to the Venetian Lagoon connected to the Adriatic Sea, permitted to conclude that two additional forcings are important in proximity of inlets: wind stress and tidal cicles. As a future step, the thermohaline coastal circulation will be investigated, even introducing to the real case run the temperature and salinity initial conditions to compute the baroclinic terms.

1 INTRODUCTION

Coastal processes modeling represents one of the most interesting topics for modelers, because of the variety of phenomena, on different spatial and temporal scales. Littoral areas have multiple aspects to consider since small scale phenomena appear near the connection channels-basin, primarily due to the tidal cicle, the wind and the turbulent mixing processes. These hydrodynamics interact with the general circulation of the main basin, and mesoscale and sub-mesoscale phenomena can be appreciated in the coastal area (permanent eddies). Because of the multiplicity of processes to simulate, a lower degree of approximation is called for the model, both in terms of a new numerical development and of an increase of spatial resolution in the study area. There is a challenge in reproducing the study area with resolute bathymetrical data and in covering it with the right forcings. A high spatial resolution is required near the coastline and in the proximity of complex morphological areas, such as inlets or rivers. 3D finite element modeling, solving the shallow water momentum equations, is suited for this purpose, since it permits to increase spatial resolution near the coastline, decreasing it offshore (Fig. 1).

The development of these kind of models results important to investigate the particular area of the Venice Lagoon, its inlets and the North Adriatic littoral, which are characterized by a shallow bathymetry, increasing in depth from 0.5 m, inside the lagoon, to 20 m, 15 km offshore. This is the shallowest zone in the Adriatic Sea, and the impact of forcings as wind and tides produces visible effects here, particularly near the connection inlets (Gačić et al. 2002).

The study of the hydrodynamics of the interaction channels between a lagoon and the open sea requires to consider even that different termohaline environments are connected.

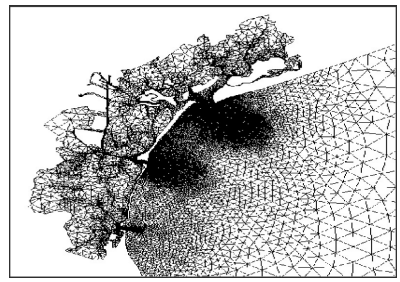


Figure 1. High resolution finite element grid study area – north adriatic, venice lagoon.

This aspect and the fact that near the studied lagoon there are several rivers flowing into the open sea, suggest the need to do an additional effort to model even the baroclinic pressure gradients.

In section 2 the 3D SHYFEM (Shallow Water Hydrodynamic Finite Element Model) is described. The test case used to check the new version of the model is presented in section 3. Some evidence of the hydrodynamic processes in the study area and preliminary considerations on a 3D real case run are located in section 4. Finally, in section 5 some conclusions are drawn.

2 CONCEPTUAL MODEL

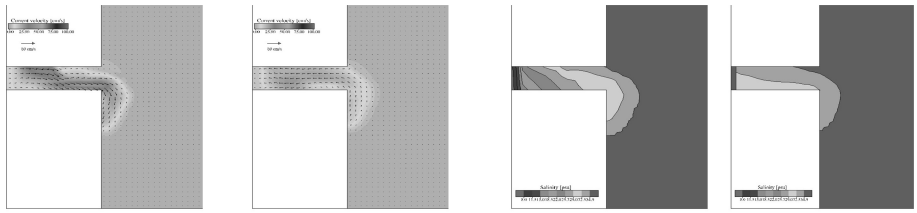
The 3D SHYFEM model has been implemented to simulate the lagoon-open sea interaction phenomena. The model, in 2D versions, has been applied in many studies to the Venice Lagoon Area (Umgiesser et al. 2004; Cucco and Umgiesser 2006; Umgiesser 2000). It is a primitive equation model, based on the solution of momentum and continuity shallow water equations. The complete transport equations, including also the baroclinic pressure gradient terms, are:

$$\frac{\partial U_i}{\partial t} - f V_i = A dv_i^x - g\frac{\partial \zeta}{\partial x}h_i - \frac{gh_i}{\rho_0}\frac{\partial}{\partial x}\int_{-H_i}^{\zeta}\rho' dz - \frac{1}{\rho_0}\frac{\partial p_a}{\partial x} + \frac{1}{\rho_0}(\tau_x^{i-1} - \tau_x^i)$$
$$+ A_H\left(\frac{\partial^2 U_i}{\partial x^2} + \frac{\partial^2 U_i}{\partial y^2}\right) + A_v\left(\left.\frac{\partial u}{\partial z}\right|_{\text{top}(l)} - \left.\frac{\partial u}{\partial z}\right|_{\text{bottom}(l)}\right) \quad (1)$$

$$\frac{\partial V_i}{\partial t} + f U_i = A dv_i^y - g\frac{\partial \zeta}{\partial x}h_i - \frac{gh_i}{\rho_0}\frac{\partial}{\partial y}\int_{-H_i}^{\zeta}\rho' dz - \frac{1}{\rho_0}\frac{\partial p_a}{\partial y} + \frac{1}{\rho_0}(\tau_y^{i-1} - \tau_y^i)$$
$$+ A_H\left(\frac{\partial^2 V_i}{\partial x^2} + \frac{\partial^2 V_i}{\partial y^2}\right) + A_V\left(\left.\frac{\partial v}{\partial z}\right|_{\text{top}(l)} - \left.\frac{\partial v}{\partial z}\right|_{\text{bottom}(l)}\right) \quad (2)$$

with i vertical layer, (U, V) horizontal transport, p_a atmospheric pressure, g gravitational constant, ζ sea level, $\rho = \rho_0 + \rho'$ water density, h_i layer thickness, H_i depth, A_H horizontal eddy viscosity and A_v vertical eddy viscosity. The stress terms can be written for each vertical interface as

$$\tau_x^i = -k\left(\frac{\partial u}{\partial z}\right)_i \tau_y^i = -k\left(\frac{\partial v}{\partial z}\right)_i \quad (3)$$



(a) (b) (c) (d)

Figure 2. Velocity plots, first layer (a), fourth layer (b). Salinity plots, first layer (c), fourth layer (d).

and the boundary conditions for stress terms are

$$\tau_x^{\text{sup}} = -c_D \rho_a w_x \sqrt{w_x^2 + w_y^2} \quad \tau_y^{\text{sup}} = c_D \rho_a w_y \sqrt{w_x^2 + w_y^2} \tag{4}$$

$$\tau_x^{\text{bottom}} = c_B \rho_0 u_L \sqrt{u_L^2 + v_L^2} \quad \tau_y^{\text{bottom}} = c_B \rho_0 v_L \sqrt{u_L^2 + v_L^2} \tag{5}$$

C_B is the bottom friction coefficient, c_D the wind drag coefficient, ρ_a the air density, ρ_0 the constant water density, (w_x, w_y) the wind velocity and (u_L, v_L) the bottom velocity.

$$w_0 = \frac{\partial \zeta}{\partial t} w_L = -\left(u_L \frac{\partial H_L}{\partial x} + v_L \frac{\partial H_L}{\partial y} \right) \tag{6}$$

w_0 and w_L are the vertical velocities at the surface and at the bottom layers. The SHYFEM model applies an Arakawa B grid for the horizontal discretization and z layers. The barotropic and Coriolis terms are semi-implicitly discretized, while the baroclinical, the advective and the horizontal diffusion terms are explicitly discretized. Some developments and changes have been done in the computation of the Coriolis term from the first 3D model version presented in (Umgiesser and Bergamasco 1995).

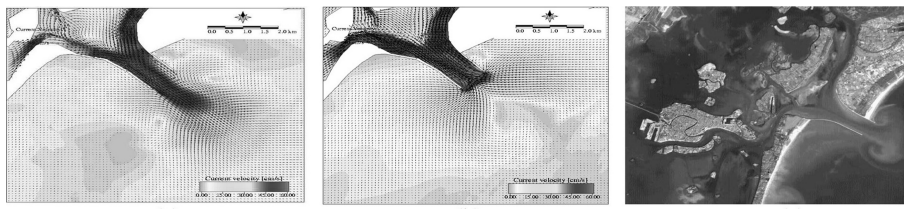
3 TEST CASE

In order to test the model with the introduction of baroclinic terms, a computational test case, proposed by Chao and Boicourt (1986), was run. A 0.01 ms^{-1}freshwater inflow velocity is set at the open boundary upper layers of a channel connected to an ideal basin. Outflow is set at the lower layers to obtain no barotropic flux at the open boundary. The run setup considers bottom friction 0.01, vertical eddy viscosity 10E-4 m^2s^{-1}, horizontal diffusion 10m^2s^{-1} and beta plane Coriolis approximation for latitude 30 degrees North. After 5 days the velocity and the salinity fields are as presented in Figure 2. These results are in agreement with the results presented by Chao and Boicourt (1986).

This work comes after and is meant to be a development of previous tests. Applying the first version of 3D SHYFEM Model a real case has been run, the Po river outflow (Umgiesser and Bergamasco 1998). Starting from that version, the baroclinic pressure gradient terms computation has been developed and changed and now new application to real basins tests are needed.

4 INLET HYDRODYNAMICS – REAL CASE

To simulate interaction processes at the three Venice Lagoon inlets, a coupled finite element grid Venice lagoon-open sea has been used with an open boundary in the Southern part of the Adriatic Sea. As a first step, the impact of tidal forcings (imposed at Otranto) and realistic wind forcings (taken from ECMWF model) has been considered. The termohaline forcings were not introduced



(a) (b) (c)

Figure 3. Tidal cicle eddies at the Lido Inlet (a). Dipole formation as seen in Wells theoretical model. (c) Tidal cicle eddies formation as seen from satellite images. Lido Inlet,Venice.

in this preliminary run. The results of a two months run (January and February 2004) evidenced the formation of symmetrical eddies near the inlets, due to the tidal cycle (Fig. 3ab). After some tidal cycles the two eddies merge in a dipole propagating offshore. This kind of phenomenon has already been characterized in satellite images (Fig. 3c) and even in laboratory experiments proving the theoretical model proposed by Wells and Van Heijst (2003).

5 CONCLUSIONS

In this paper a test of the state of the art 3D SHYFEM model in a coastal area has been presented. A step forward has been concluded in the modeling of baroclinic pressure gradient terms, defining the termohaline contribution to circulation. The first implementation of this last version of the model to a real case, the coupled basin Venetian Lagoon-Adriatic Sea, on a coastal high resolution finite element grid, permitted to draw some conclusions and to establish future steps. Tidal forcing reveals to be important in proximity of the lagoon inlets, as reported by Gačić et al. (2004). The next steps will be in the direction of a run with termohaline forcings introduced both in a diagnostic and prognostic way and with freshwater input by rivers. We work on the hypothesis that this additional information could explain the sub-mesoscale eddies forming along the coast in the area between the Venice Lagoon inlets, as registered by HF Radars Kovačević et al. (2004).

REFERENCES

Chao, S. and W.C. Boicourt (1986). Onset of estuarine plumes. *Journal of Physical Ocenography 16*, 2137–2149.

Cucco, A. and G. Umgiesser (2006). Modeling the Venice Lagoon Residence Time. *Ecological Modelling, 193*(1–2),34–51.

Gačić, M., V. Kovačević, A. Mazzoldi, J. D. Paduan, I.M. Mosquera, G. Gelsi, and G. Arcari (2002). Measuring water exchange between the Venetian Lagoon and the open sea. Eos, *Transaction, American Physical Union 83*(20).

Gačić, M., I.M. Mosquera, V. Kovačević, A. Mazzoldi, V. Cardin, and F. Gelsi (2004). Temporal variation od water flow between the Venetian lagoon and the open sea. *Journal of Marine Systems 51*.

Kovačević, V., M. Gačić, I.M. Mosquera, A. Mazzoldi, and S. Marinetti (2004). HF radar observation in the northern Adriatic: surface current field in front of the Venetian Lagoon. *Journal of Marine Sistems 51(1–4)*,95–122.

Umgiesser, G. (2000). Modeling residual currents in the Venice Lagoon, pp. 107–124. Tokio: Terra Scientific Publishing Company (TERRAPUB).

Umgiesser, G. and A. Bergamasco (1995). Outline of a primitive equation finite element model. In *Rapporto e Studi*, Volume XII, pp. 291–320. Venice, Italy: Istituto Veneto di Scienze, Lettere ed Arti.

Umgiesser, G. and A. Bergamasco (1998). The spreading of the Po plume and the Italian coastal current. In *Physics of Estuaries and Coastal Seas*, pp. 267–275. Rotterdam: Dronkers and Scheffers (eds).

Umgiesser, G., D. Melaku-Canu, A. Cucco, and C. Solidoro (2004) A finite element model for the Venice Lagoon. Development, set-up, calibration and validation. *Journal of Marine Systems 51*, 123–145.

Wells, M.G. and G.-J.F.V. Heijst (2003). A model of tidal flushing on an estuary by dipole formation. *Dynamics of Atmospheres and Oceans 37*, 223–244.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

On the wave-front shape and the advancing of the wetting front of a dam-break flood over an inclined plane of arbitrary bottom slope

P. Bohórquez*
Universidad de Málaga, E.T.S. Ingenieros Industriales, Málaga, Spain

ABSTRACT: In this note we consider the analysis of the wave-front shape resulting from the sudden release of a finite volume of water over an inclined plane bed of arbitrary bottom slope. To that end we use the one-dimensional turbulent shallow water equations with a constant friction factor. We propose an asymptotic analytical solution for the height and velocity in the wave tip region based on the velocity of the wetting front and its temporal derivatives, in a similar fashion as Whitham (1955). For large time, an analytical solution is obtained for the advance of the wetting front. This proposal is compared and validated with a numerical simulation computed with a second-order TVD-MinMod numerical scheme. The divergence of the Hunt's solution (Hunt 1982; Hunt 1984) for the advancing of the wetting front as time increases is also established. Our solution for the tip region is tested against the numerical one, and is also compared with that deduced by Hunt (1984) for small slopes of the bottom – a really good agreement is found. An additional and novel result is also presented, the formation of roll waves in the dam-break problem for Froude numbers larger enough when a quasi-steady and quasi-uniform regimen is reached.

1 INTRODUCTION

As far as the author knows the first theoretical studies on the effect of hydraulic resistance on dam-break waves were performed by Dressler (1952) and Whitham (1955) – these investigations started with the analysis of horizontal beds. Later on more realistic problems were analysed, for instance dam-break floods of a finite volume of water over plane inclined beds [i.e., Hunt (1982, 1984)]. Recently this topic have achieved new interest and insight in its theoretical aspect [see, for instance, Hogg & Pritchard (2004)] as well as in its experimental version (Lauber 1997).

However, the above theoretical works used the general expressions of Saint Venant, often called the nonlinear shallow-water equations, valid only for small slopes of the bottom. In this line we have considered the analysis of the release of a finite volume of water over an inclined bed of arbitrary bottom slope, using the expressions provided by Fernandez-Feria (2006) (but adding basal drag effects) to appropriately describe the fluid movement over arbitrary but constant slopes.

The paper is organized as follows: we firstly formulate the problem (§2). We then analyse the flow movement under the kinematic wave approximation and obtain the velocity and height field as well as the wave-front location (§3.1). In §3.2 we investigate the wave shape in the tip region and provide the solution as an asymptotic expansion, indeed the explicit expressions for the whole coefficients are deduced. This solution is also shown to be in close agreement with Hunt's solution

* This research has been supported by a fellowship from the Ministerio de Educación y Ciencia of Spain (AP-2004-3719).

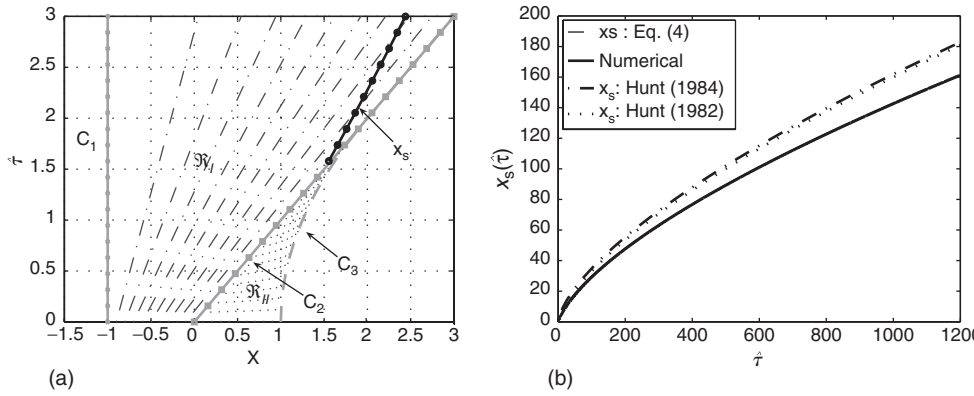


Figure 1. (a) Characteristic curves in a space-time diagram under the kinematic wave approximation for $e = 1$. (b) Temporal evolution of the wave-front obtained numerically compared with the asymptotic result by Hunt (1982, 1984) and with Equation (4); we have selected $\eta_0 = 1$ m, $e = 1$ and $Fr_{eq} \equiv \sqrt{8e/f} = 2$.

(Hunt 1984) for small slopes of the bottom. We compare our theoretical predictions with numerical experiments in §3.1 and §3.2. A summary and some concluding comments are given in §4.

2 FORMULATION OF THE PROBLEM

We consider here the one-dimensional flow over a constant slope bed. In the shallow-water approximation the dimensionless equations for the mass conservation and momentum in the direction of the flow can be written as

$$\partial_t \eta + \partial_X(\eta U) = 0, \quad \partial_t U + U\partial_X U + \cos\theta\, \partial_X \eta = \sin\theta - f U^2/(8\eta). \tag{1}$$

The notation and initial conditions of the problem under consideration are exactly the same as used in Fernandez-Feria (2006), therefore we do not repeat this information. We have only added the effects of basal drag using the Darcy-Weisbach friction factor f (considered constant in the present work) in the momentum balance equation. These equations, are valid for *any* slope θ of the *constant-slope* bed, not just for small channel slope, provided that the ratio of the depth η to the streamwise extent is sufficiently low.

3 RESULTS

3.1 *Kinematic wave solution*

In this section we follow the same theoretical approach as Hunt (1982) but solving Equations in (1) valid for arbitrary bottom slope to obtain the forward wave (wetting front) advancing over a dry bed.

In the kinematic wave approximation, one neglects the left hand side of the momentum equation and obtains the relation that follows between η and U, $U^2 = (8\eta/f)\sin\theta$. Taking into account this relationship, mass conservation yields

$$\frac{\partial \eta}{\partial \tau} + \frac{3}{2}\sqrt{\eta}\frac{\partial \eta}{\partial X} = 0\,, \quad \tau \equiv \left(\frac{8\sin\theta}{f}\right)^{1/2} t. \tag{2}$$

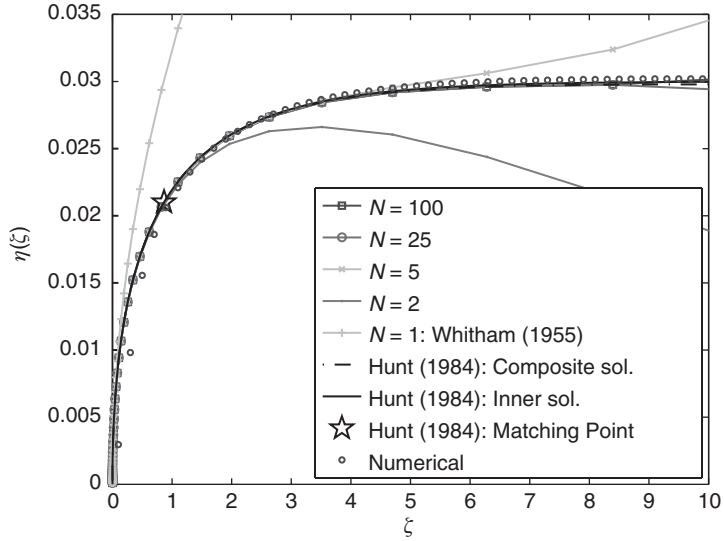


Figure 2. Asymptotic solution (7) for the height $\eta(\zeta)$ for a given instant of time for several values of N. The case $N=1$ corresponds to the analogous solution given by Whitham (1955). Also included is the asymptotic solution by Hunt (1984) and the numerical one for the same conditions. We have selected $t=4411$, $\eta_0=1$ m, $e=0.0175$ and $Fr=2.5$.

The equation above can be readily integrated together with the initial condition for η [see Fernandez-Feria-(2006) for its definition]. The wave profile should be defined in two different regions: the first area (denoted by $\Re_I$) is enclosed by the two real characteristics $C_1 \equiv -1/e$ and $C_2(\tau) \equiv 3\tau/2$, and this region exists for any τ; the second region ($\Re_{II}$) ranges between the characteristic $C_2(\tau)$ and $C_3(\tau) \equiv (16e^2+9\tau^2)/(8e)$, and it exists just for $\tau \leq 4e/3$. After defining these two different areas the solution reads

$$\eta(X,\tau) = \begin{cases} 1+(e/2)\left(e\hat{\tau}^2 + 2X - \hat{\tau}\sqrt{4+e^2\hat{\tau}^2+4eX}\right) & \text{if } X \in \Re_I \\ 1-(2eX-\hat{\tau}^2-\hat{\tau}\sqrt{4e^2+\hat{\tau}^2+4eX})/(2e^2) & \text{if } X \in \Re_{II} \end{cases}, \tag{3}$$

being $\hat{\tau} \equiv 3\tau/2$. As in Hunt's (1982) solution, a shock must be inserted in the solution to satisfy the mass conservation requirement. For simplicity we provide the location of the shock $x_s(\tau)$ just for $\tau > 3x_s/2$. In closed-form it is given by

$$\hat{\tau}\left(4+e^2\hat{\tau}^2+4ex_s\right)^{3/2} = e^3\hat{\tau}^4 + 12x_s + 6e^2\hat{\tau}^2x_s + 6e\left(\hat{\tau}^2+x_s^2-1\right). \tag{4}$$

Figure 1a depicts both the region $\Re_I$ (dashed-dotted line) and the region $\Re_{II}$ (dotted line) in the characteristic plane $\{X,\hat{\tau}\}$ for $e=1$. The border between these two regions is established by the characteristic curve C_2 (solid, squared line). One should note the presence of the shock x_s (solid, circled line), which establishes the validity of the solution (3): on the right hand side of the shock the solution is no longer valid, and both the height and the velocity field vanish. The other two characteristics C_1 (solid, crossed line) and C_3 (dashed line) are also shown in the same figure. To validate the front location (4), Equations in (1) (with the corresponding initial conditions) have been solved numerically on an uniform grid using an upwind TVD (total variation diminishing) method, second-order accurate in both space and time, with semi-implicit and upwind treatment of the source terms, as described by Burguete & Garcia-Navarro (2001). The boundary condition used for the wetting front is a cut-off height ($\eta/\eta_{\max} < 10^{-4}$), the Courant-Friedrich-Lewy number is

$CFL = 0.45$, and the number of nodes is $nx = 10^4$. Figure 1b shows a close agreement between the numerical (solid line) and the analytical result (4) (dotted line). We have also plotted the asymptotic result provided by Hunt (1982, 1984) considering the shallow-water equations due to Saint Venant. The solutions by this author becomes worse as time goes on because the slope e is not small.

3.2 *Wave-front shape*

If we define the variable $\zeta = x_f(t) - X$, where x_f is the position of the wetting front, Equation (1) read (note that $U > 0$)

$$\partial_t \eta + (x_f' - U)\partial_\zeta \eta - \eta \partial_\zeta U = 0\,, \tag{5}$$

$$\partial_t U + (x_f' - U)\partial_\zeta U - \cos\theta\, \partial_\zeta \eta = \sin\theta - fU^2/(8\eta)\,, \tag{6}$$

where primes denote differentiation with respect to t. One can obtain a solution of the resulting equations in powers of ζ using the same expansion as Whitham (1955), which can be written as

$$\eta(\zeta, t) = \sum_{j=0}^{N} \eta_j(t)\zeta^{\frac{j+1}{2}}\,, \quad U(\zeta, t) = \sum_{j=-1}^{N} U_j(t)\zeta^{\frac{j+1}{2}}\,. \tag{7}$$

If the wave-front is taken to be $X = x_f(t)$, then $U_{-1} = x_f'$ at the wave-front. On the other hand, the particle acceleration, $\partial_t U + (x_f' - U)\partial_\zeta U$, is expected to be finite while the effect of resistance is to pile up the fluid near the wave-front, so that $\partial_\zeta \eta$ becomes large. Therefore, to first approximation, $-\eta \cos\theta\, \partial_\zeta \eta = \eta \sin\theta - fU^2/8$. Taking into account (7) and grouping terms of order ζ one obtains $\eta_0 = x_f'/2(f/\cos\theta)^{1/2}$. Now, substituting the expansion (7) in the continuity equation (5) and grouping again terms of order ζ one obtains $U_0 = 0$. Repeating this process for the next exponents in ζ, firstly with the momentum equation (6), and then with the mass balance (5), the coefficients η_i and U_i are obtained explicitly. For instance, grouping terms $O(\zeta^{1/2})$ in the mass balance provides $\eta_1 = 2(x_f'' - \sin\theta)/(3\cos\theta)$. All the coefficients can be obtained analytically:

$$U_N = \frac{1}{\eta_0}\left(\frac{2}{N+2}\eta_{N-1}' - \sum_{j=1}^{N-1} U_j \eta_{N-j}\right) \quad \text{for } N \geq 1, \tag{8}$$

$$\eta_N = \frac{2}{N+2}\frac{1}{\eta_0 \cos\theta}\left[-\cos\theta \sum_{j=1}^{N-1} \frac{N-j+1}{2}\eta_j \eta_{N-j} - \sin\theta\, \eta_{N-1} + \frac{f}{8}\sum_{j=-1}^{N-1} U_j U_{N-j-2} + \sum_{j=0}^{N-1} \eta_j U_{N-2-j}' + \sum_{j=0}^{N-1}\left(-\eta_j \sum_{k=1}^{N+1-j} \frac{N-k-j+1}{2} U_{N-k-j}U_{k-1}\right)\right] \quad \text{for } N \geq 2. \tag{9}$$

This asymptotic solution, which actually converges even for $\zeta > 1$ (see Fig. 2), was originally found by Whitham (1955), but this author only provided explicit expressions for the first few coefficients, just up to $N = 1$. The first result we must point out is that the velocity field depends on ζ from the third term in the expansion, and therefore the assumption performed by Whitham "that U is nearly uniform in the tip region" (Whitham 1955) is valid only until $O(\zeta)$. Second, the solution converges slowly as N is increased if ζ is not too small, as it is shown in Figure 2, where the height η has been plotted as a function of ζ for a given t. To compute x_f' and next derivatives we have used the previous result (4) valid for *any* slope of the bottom.

In that figure we also include the form of the wetting front given by the asymptotic solution by Hunt (1984) showing a very good agreement with the asymptotic solution (7)–(9) if N is sufficiently large. This is so because we have used a very small bottom slope (note that $e = 0.0175$ implies

$\theta = 1°$), and thus the Saint Venant equations for a plane bed approximates with precision (1), in order to check (7) with a known solution. But this solution is actually valid for any constant slope. Our expansion is in agreement with the inner solution (plotted in continuous line) for $N \geq 2$. From the matching point (big star) to the right hand side, the inner and the composed solution (dashed-dotted line) by Hunt are almost identical, and we need quite a lot terms in the expansion (e.g., $N > 25$) to obtain a similar result. Furthermore, we have also computed the numerical solution with the same technique described in §3.1($CFL = 0.45$ and $nx = 10^4$) and a very good agreement has been found with the theoretical ones (Fig. 2). Finally, although the slope of the bed is small, the kinematic wave approximation is valid [one can verify $x_s \gg 4(e + 1/e)$], and thus the source terms in Equation (1) dominate the fluid movement. Even in this case the cut-off height method used to compute the advance of the wetting-front is good enough for the problem under consideration.

4 CONCLUSIONS

In this study we have considered the effects of hydraulic resistance on the dam-break flood over an inclined plane of arbitrary but constant slope.

For the kinematic wave regime the theoretical location of the wetting front has been supplied together with the velocity and height field (§3.1). We have also solved the shallow-water equations (with the appropriate initial conditions) using an upwind TVD method, second-order accurate in both space and time, with semi-implicit and upwind treatment of the source terms. We find that both solutions are in agreement. One of the most noticeable effects of the results for large slopes of the bottom is the divergence of the previous solution by Hunt (1982, 1984), who provided a similar asymptotic solution using the traditional Saint Venant equations. This divergence arises mainly because the whole gravity g moves the fluid when the Saint Venant equations are used, instead of its projection along the perpendicular to the plane.

We have also considered the analysis of the shape of the flood in the tip region (§3.2). The results are not only valid for a dam-break flood but also for *any mass of fluid* that spreads on a plane bed. The solution has been obtained as a full expansion in terms of the streamwise coordinate located at the wetting front. Our results are directly comparable to those obtained by Hunt (1984) for small slopes of the bottom. We find that the earlier result exhibits a good agreement with our solution when we use enough terms in the expansion, but Hunt's solution is valid only for small slopes of the bottom. The accuracy of the numerical solution has been also checked, and very good results are obtained by just using a cut-off height as boundary condition for the wetting front.

Finally, we note that the solutions developed could be used to check, for instance, the precision of other numerical schemes or boundary conditions.

REFERENCES

Burguete, J. & Garcia-Navarro, P. (2001). Efficient construction of high-resolution TVD conservative schemes for equations with source terms: application to shallow water flows. *Int. J. Num. Meth. Fluids* 37(1), 209–248.

Dressler, R. (1952). Hydraulic resistance effect upon the dam-break functions. *J. Res. Nat. Bur. Stand.* 49, 217–225.

Fernandez-Feria, R. (2006). Dam-break flow for arbitrary slopes of the bottom. *Journal of Engineering Mathematics* 54(7), 319–331.

Hogg, A. J. & D. Pritchard (2004). The effects of hydraulic resistance on dam-break and other shallow inertial flows. *Journal of Fluid Mechanics* 501, 179–212.

Hunt, B. (1982). Asymptotic solution for dam-break problem. *J. Hydraul. Div.* 108(1), 115–126.

Hunt, B. (1984). Perturbation solution for dam-break floods. *J. Hydraul. Eng.* 110(8), 1058–1071.

Lauber, G. (1997). *Experimente zur Talsperrenbruchwelle im glatten geneigten Rechteckkanal*. Ph. D. thesis, VAW-ETH, Zrich, Switzerland.

Whitham, G. (1955). The effects of hydraulic resistance in the dam-break problem. *Proc. Roy. Soc. of London 227*, 399–407.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Roll waves in floods on inclines

P. Bohórquez*
Universidad de Málaga, E.T.S. Ingenieros Industriales, Málaga, Spain

ABSTRACT: The unstable nature of kinematic waves, resulting from outburst floods on steep inclines, is investigated theoretically using the shallow water equations. A numerical experiment is designed in order to efficiently perform a stability analysis. By analysing the linear and non-linear numerical evolution of small perturbations, it is possible to reveal how the nature of the ensuing flow depends not only on the Froude number but also on non-parallel and time-varying characteristics of the background flow. It is also demonstrated that these effects stabilize turbulent roll-waves and raise the critical Froude number required for instability.

1 INTRODUCTION

Our goal in the current article is to establish the existence of physical instabilities in floods on steep inclines. To gain some inroads into this problem, we shall consider turbulent flows governed by the one-dimensional shallow-water equations. We shall limit our study to large times, in which the suddenly released, fixed mass of fluid lengthens and reaches a kinematic state.

The spreading of a suddenly released mass of fluid over an inclined plane bed has been widely studied in the above conditions [see, for instance, Weir (1983), Hunt (1984), Pritchard (2005)], and it is of hydraulic interest not only for the hydrodynamic but also for the associated transport of sediments. However, none of the previous studies have taken into account the strongly unstable character of this flow (Bohorquez 2006).

Another point of interest to be noted is that nowadays studies on the onset of instability, performed with the one-dimensional shallow-water equations, have been limited to steady flows over plane beds [e.g., Prokopiou *et al.* (1991)] or with spatially periodic bottom topography (Balmforth & Mandre 2004). Many complications are introduced by non-parallel and time-dependent effects of the background state, which enrich the dynamics of real flows, when analysing its linear stability (Schmid 2007). Hence, the author considers that the study about this time-varying space-dependent system is an outstanding issue.

The reader can find in this volume the definition of the dam-break problem for a *constant but arbitrary slope* of the bottom as well as the asymptotic solution under the kinematic wave approximation (Bohorquez 2007). However, when friction is taken into account in such problem, it is well known that the flow loses memory of its initial distribution. Hence, considering a point source of mass, located at $X = 0$ with initial conditions $U(X, 0) = 0$ and $\eta(X, 0) = A\delta(X)$, being A the released volume, allows to obtain a more simple solution. In fact, defining a scaling factor ϵ, and rescaling non-dimensional variables in Bohorquez (2007), § 2, by the following relations,

$$x \equiv X\epsilon \tan\theta, \quad \tau \equiv t\epsilon \tan\theta\sqrt{(8/f)\sin\theta}, \quad u \equiv U\sqrt{f/(8\sin\theta)}, \tag{1}$$

* This research has been supported by a fellowship from the Ministerio de Educación y Ciencia of Spain (AP-2004-3719).

conservation of mass and momentum read

$$\partial_\tau \eta + \partial_x(\eta u) = 0, \quad \epsilon \left[Fr_{eq}^2 \left(\partial_\tau u + u \partial_x u \right) + \partial_x \eta \right] = 1 - u^2/\eta, \tag{2}$$

respectively, where Fr_{eq} is defined as $Fr_{eq}^2 \equiv (8/f)\tan\theta$. The *Froude* number, which states the local ratio of the flow velocity U to the wave celerity $c \equiv (\eta\cos\theta)^{1/2}$, is given in the new variables by $Fr \equiv U/c = Fr_{eq} u/\sqrt{\eta}$, and reduces to the constant parameter Fr_{eq} when a balance is struck between the friction of the bottom and the streamwise component of gravity, in other words, when ϵ is neglected in the momentum balance.

The set of non-dimensional equations (2) is exactly the same as those used by Hunt, but they come from different dimensional equations – since Hunt applied the St. Venant equations valid just for small slopes of the bottom. In view of this, we find that *dimensionless Weir's (1983) and Hunt's (1984) solutions for floods on inclines are also valid for steep slopes when non-dimensional variables are defined correctly*. This result will be used for the formulation of the stability problem (see bellow).

We start defining the linear stability equations and the base state in § 2, as well as the numerical schemes used in following sections. Next, in § 3, we study numerically the linear and non-linear stability of our background flow. We devote § 4 to summarize our results. Overall, the study is focused on non-uniform and unsteady base flows, but some of the results carry over parallel (uniform and steady) flows, in order to check with previous analytical results the ability of the numerical schemes to reproduce the behaviour of the perturbations.

2 FORMULATION OF THE PROBLEM AND NUMERICAL SCHEMES

Here, we are interested in kinematic waves as base state, obtained when the parameter ϵ is small in the momentum balance (2), provided u and η as well as their gradients are order unity. In this case, the velocity and height can be expanded in powers of ϵ to find an analytical solution:

$$u_b(x,\tau) = V(x,\tau) + \epsilon V_1(x,\tau) + O(\epsilon^2), \quad \eta_b(x,\tau) = H(x,\tau) + \epsilon H_1(x,\tau) + O(\epsilon^2). \tag{3}$$

The first term in the expansion, called first-order *outer* solution, was found by Hunt (1984), and is given by

$$V(x,\tau) = 2x/(3\tau), \quad H(x,\tau) = V^2. \tag{4}$$

We shall carry the stability analysis on this basic flow.

To analyse the stability of such time-dependent system, we shall superimpose an infinitesimal disturbance, $v(x,\tau)$ and $h(x,\tau)$, on the background flow,

$$u = u_b(x,\tau) + v(x,\tau), \quad \eta = \eta_b(x,\tau) + h(x,\tau), \tag{5}$$

where $|h| \ll \eta_b$ and $|v| \ll |u_b|$. Substituting (5) in (2), and neglecting second order terms in the small perturbations, one is left with the following set of linear equations for the perturbations:

$$\frac{\partial h}{\partial \tau} + \frac{\partial \eta_b v}{\partial x} + \frac{\partial h u_b}{\partial x} = 0, \quad Fr_{eq}^2 \left[\frac{\partial v}{\partial \tau} + \frac{\partial v u_b}{\partial x} \right] + \frac{\partial h}{\partial x} = \frac{u_b^2}{\epsilon \eta_b} \left[\frac{h}{\eta_b} - \frac{2v}{u_b} \right], \tag{6}$$

The above partial differential equations (PDEs) require one initial condition for both the velocity v and height h fields. We shall consider the initial configuration (at $\tau = \tau_0$) of the perturbation v on the basic flow $u_b \approx V$ (henceforth, we will use just the first term in the expansion (3) to compute the base state) given by

$$v(x,\tau_0) = \begin{cases} B(l_1)\sin\left(2\pi n \frac{x-l_1}{l_2-l_1}\right) & \text{if } l_1 \le x \le l_2 \\ 0 & \text{otherwise} \end{cases}, \tag{7}$$

where B denotes the initial amplitude of the perturbation ($B \equiv \delta V(l_1, \tau_0)$, being δ a very small parameter), l_1 and l_2 establish the boundary enclosing the initial waves, and n denotes the number of waves. For the height, the relation that follows is used: $h(x, \tau_0) = (V + v)^2 - H \approx 2Vv$, which is $O(\delta H)$.

The convective nature of the perturbations is easily established in terms of the Froude number. This is understandable in view of the fact that both the linear (6) and nonlinear (2) system of PDEs have two real and distinct characteristic curves, and these are the same for small waves. Hence, perturbations travel along the characteristics $C_{\pm}$: $dx/d\tau = u_b \pm \eta_b/Fr_{eq}$, and the convective nature comes for supercritical kinematic waves: $dx/d\tau \approx V(1 \pm 1/Fr_{eq}) > 0$ with $Fr_{eq} > 1$. From now on, we are interested just in supercritical flows.

To compute the linear evolution of the perturbations, we use a finite-difference scheme on an uniform grid with mesh size Δx. The time integration is performed with a Crank-Nicolson method, and a fourth-order central-differences scheme is applied in space – the truncation error is $O(\Delta\tau^2, \Delta x^4)$, where $\Delta\tau$ is the time step. In order to use the minimal number of nodes, we initially solve the resulting linear system of equations just in a small region I_1 (which bounds the waves) with homogeneous boundary conditions. Before the wave train travelling downwards reaches the end of I_1, a new domain I_2 is defined containing the subregion in which waves live, and it is enlarged downwards adding new nodes. This process is repeated several times during the numerical simulation. Thus, we have to solve in each time step a linear system of equations with just eleven diagonals. Band storage mode is used to minimize memory requirements. To that end, subroutines DGBMV and DLSLRB into the IMSL Numerical Libraries are used.

On the other hand, the full non-linear set of equations (2) is solved with the same numerical technique as in Bohorquez (2007). In this case, taking into account that the flow is supercritical at both inlet and outlet, two boundaries conditions are imposed upstream (given by Eq. 4), and the characteristic variable extrapolation (CVE) method is used downstream. Next, we present some results.

3 RESULTS

First, we will show results corresponding with the parallel case. To that end, we have considered an uniform and steady basic flow with $U = 1$ and $\theta = 1^o$. This flow has been initially disturbed with n small waves ($n = 10$), and its nonlinear evolution has been computed following the approach introduced in § 2. Figure 1a depicts the maximum amplitude of the wave train as a function of time for several Froude numbers. Initially, the nonlinear growth rate is in agreement with theoretical

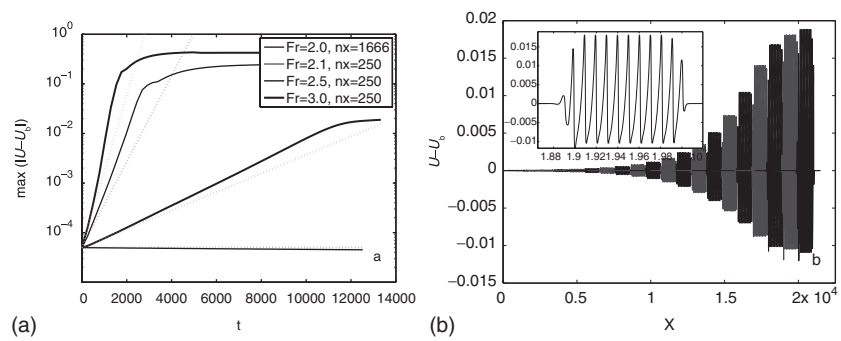


Figure 1. (a) Amplitude of the water wave velocity, $U - U_b$, against non-dimensional time, t, for different steady flow Froude numbers $Fr_{eq} \geq 2.0$ and an initial amplitude of the velocity perturbation $\delta = 5 \times 10^{-5}$. The solid curves represents the solution obtained from the non-linear numerical code, and the dashed lines indicate the corresponding growth rates obtained from the linear stability analysis (see, e.g., Brook *et al.* 1999). *nx* represents the number of nodes used by wave. (b) Plots of perturbation velocity, $U - U_b$, against distance along the channel, X, at several instants of time ($t < 1.4 \times 10^4$) corresponding with the case $Fr_{eq} = 2.1$.

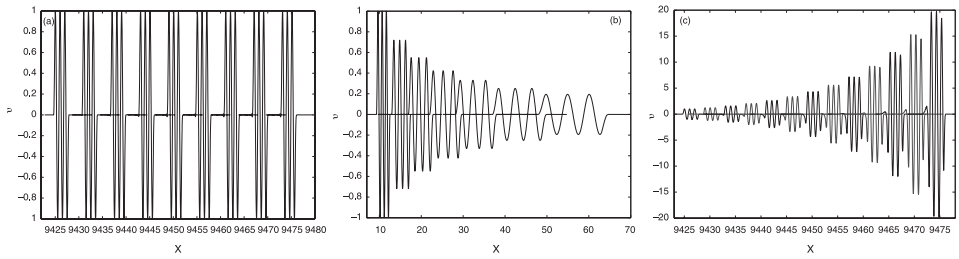


Figure 2. (a) Plots of perturbation velocity, v, against distance along the channel, X, at several instants of time with $Fr_{eq} = 2$, $\lambda = 1$, $U_b(l_1, \tau_0) = .1282$, and $\partial_X U_b(l_1, \tau_0) = 1.36 \times 10^{-5}$. The case (b) corresponds with the case (a) but with $U_b(l_1, \tau_0) = 3.9$ and $\partial_X U_b(l_1, \tau_0) = 4.09 \times 10^{-4}$. (c) As in (a) but with $Fr_{eq} = 2.2$. The solution has been obtained with the linear numerical code described in § 2.

results obtained using the linear theory, but as time goes on, and the wave amplitude is not so small, the wave changes its shape towards hydraulic jumps (as shown by the inset in Fig. 1b) and saturation appears due to nonlinear effects.

One should note that the main difference of the numerical technique we have used, with respect that applied by Brook *et al.* (1999), are the boundary conditions. Traditionally, roll waves are numerically studied by using their periodicity properties, and periodic boundary conditions are usually stated. However, in order to analyse the stability of more general basic flows (see bellow), which are not periodic, we have followed a different approach (see § 2 for details).

Next, we analyse the stability of the unsteady and non-uniform base state given by (4). One should focus the study on the question that follows: how temporal and spatial gradients of the basic flow affect the stability criteria. For simplicity, we study only the linear evolution of the perturbation, and due to the fact that the linear case conclusions do not depend on the wave amplitude B, we have normalized the perturbation velocity v ($B = 1$). In order to present the results, we have selected a wave length equal to unity, $\lambda \equiv (l_2 - l_1)/n = 1$, a wave train generated by three waves, $n = 3$, and the initial conditions of the base state will change from one simulation to another.

The first novel result we have found, when analysing the particular case $Fr_{eq} = 2$, is that non-parallel and unsteady effects stabilize the flow. Inspection of Figures 2a, b corroborates this statement. Furthermore, while the wave length remains constant in the parallel case as time goes on (Prokopiou *et al.* 1991; Brook 1999), disturbances lengthen in the present case, as it is shown in Figure 2b.

Second, when increasing the Froude number, the flow becomes unstable – as it is shown in Figure 2c, which corresponds with the numerical simulation of the same perturbation and base flow as in Figure 2a but rising the Froude number up to $Fr_{eq} = 2.2$. However, even background flows with Froude numbers larger than two are stable for some wave numbers. In fact, the numerical simulation of the case plotted in Figure 2b with Froude number $Fr_{eq} = 2.2$, instead of $Fr_{eq} = 2$, reports similar effects, i.e., the amplitude decreases while the wave lengthens (we do not show the graph due to space requirements).

4 CONCLUSIONS

In this study we have considered the stability analysis of a suddenly released, fixed mass of fluid that reaches a kinematic regime. In particular, we have designed a numerical experiment, and a clever numerical method, that allows to compute the linear and nonlinear evolution of small perturbations.

We have found that non-uniform and unsteady effects of the background flow stabilize turbulent roll-waves and raise the critical Froude number required for instability. The well known stability criteria for parallel flows at high Reynolds number (the basic flow is unstable for any wave length and Froude numbers larger than 2) differs abruptly of that resulting from kinematic waves. One of

the most noticeable effects is stabilization of disturbances on basic flows with $Fr_{eq} = 2$. In addition to that, for larger Froude numbers, $Fr_{eq} > 2$, the wave amplitude decreases or increases depending on the velocity, and both spatial and temporal gradients of the base flow. Furthermore, waves lengthen as time goes on – this behaviour is also different with respect to the parallel one, in which waves remain with constant wave length even when hydraulic jumps are developed.

P. B. gratefully acknowledges many interesting and fruitful discussions on this work with Prof. Ramón Fernández-Feria.

REFERENCES

Balmforth, N.J. & Mandre, S. (2004). Dynamics of roll waves. *J. Fluid Mech. 514*, 1–33.

Bohorquez, P. (2006). *Nonparallel spatial stability of shallow water flow down an inclined plane of arbitrary slope.* Series Monográficas: Trabajos de Investigación Tutelada, UMA, Málaga, Spain. (ISBN 84-690-3171-6).

Bohorquez, P. (2007). On the wave-front shape and the advancing of the wetting front of a dam-break flood over an inclined plane of arbitrary bottom slope. *Numerical Modelling of Hydrodynamics for Water Resources*. Taylor and Francis, /Balkema, Leiden, The Netherlands. (in press)

Brook B.S., Falle, S. & Pedley, T.J. (1999). Numerical solution for unsteady gravity-driven flows in collapsible tubes: evolution and roll-wave instability of a steady state. *J. Fluid Mech. 396*, 223–256.

Hunt, B. (1984). Perturbation solution for dam-break floods. *J.Hydraul. Eng. 110* (8), 1058–1071.

Pritchard, D. (2005). On fine sediment transport by a flood surge. *J. Fluid Mech. 543*, 239–248.

Prokopiou, Th., Cheng, M. & Chang, H.C. (1991). Long waves on inclined films at high Reynolds number. *J. Fluid Mech. 222*, 665–691.

Schmid, P.J. (2007). Nonmodal stability theory. *Annu. Rev. Fluid Mech. 39*, 129–162.

Weir, G.J. (1983). The asymptotic behaviour of simple kinematic waves of finite volume. *Proc. Roy. Soc. of London 387*, 459–467.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Effects of the irregularity in the cross sectional distribution of velocity on the 1D shallow-water model

J. Burguete
Suelo y Agua, Estación Experimental Aula Dei, CSIC, Zaragoza, Spain

P. García-Navarro & J. Murillo
Fluid Mechanics, CPS, Universidad de Zaragoza, Zaragoza, Spain

ABSTRACT: In this work, a decomposition of the channel cross section in subsections is performed to improve the 1D shallow-water simulation of unsteady open channel flows in cases of distorted cross sectional shapes. Some test cases with measured data are used to validate the quality of the approaches and a natural river is used to illustrate the effects due to the cross sections irregularity.

1 INTRODUCTION

To model friction in 1D shallow-water, empirical or semi-empirical methods like Gauckler-Manning have traditionally been applied. In cases of channel or river flow with overbank sections as well as those with non-uniform cross section bed roughness, the traditional hypothesis of uniform velocity is unrealistic. It is not obvious what is the correct form to incorporate a cross sectional variability of the frictional coefficient in 1D models. It becomes necessary to perform a decomposition of the channel cross section in subsections as, for instance, suggested in (Chow 1959; Knight 2006; Cao et al. 2006; Burguete et al. 2007). This procedure is based on the assumption of a single cross sectional energy slope. We seek a model less refined than (Shiono et al. 1999), who model lateral exchange of momentum requiring the solution of a differential equation in the transversal direction and therefore becomes excessively complex for our purpose. The lateral exchange of momentum is neglected in our approach. The main hypothesis will be that the wetted cross section is much wider than deep and our simplification will be supported by experimental evidence (MacGahey et al. 2006) that the lateral exchange of momentum effects and secondary flows have a relative magnitude smaller than boundary friction effects in most rivers. A laboratory channel and a gauge section in a river, both cases with measured data, are used to validate the quality of the approaches. Finally, the model is applied in a natural river to illustrate the effects due to the cross sections irregularity.

2 FRICTION MODEL

The 1D momentum conservation equation in open channel flow models can be expressed as:

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x}\left(\frac{\beta Q^2}{A} + gI_1\right) = g[I_2 + A(S_0 - S_f)] \tag{1}$$

where A is the wetted cross section, Q the discharge, g the gravitational acceleration, S_0 the main bed slope, S_f the friction stress over the solid surface in the channel reach, β is a coefficient that

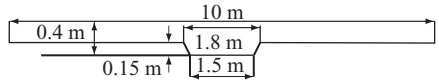


Figure 1. Cross section geometry of the serie I of the FCF channels.

appears as a result of the assumption of variable velocity in the cross section:

$$\beta = \frac{A}{Q^2} \int_A v_x^2 \, dA \tag{2}$$

with $v_x = v_x(x,\ y,\ z,\ t)$ the x component of the local point flow velocity, and I_1, I_2 pressure forces integrals (Chow 1959).

In the empirical Gauckler-Manning model, the friction slope S_f is:

$$S_f = \frac{n^2 |Q| Q}{R^{4/3}} \tag{3}$$

with $R = A/P$ the hydraulic radius, P the wetted perimeter and n the Gauckler-Manning coefficient. This model is only valid in cases of uniform flow velocity distribution in a cross section. In irregular sections or compound channels, the velocity cannot be considered uniform in a cross section. In these cases a constant slope model can be used (Burguete et al. 2007). In this case:

$$U = \frac{\pm\sqrt{|S_f|}h^{2/3}}{n} \Rightarrow Q = \int_A Uh \, dy = \pm\sqrt{|S_f|} \int_A \frac{h^{5/3}}{n} dy \Rightarrow S_f = \frac{|Q|Q}{\left(\int_A \frac{h^{5/3}}{n} dy \right)^2} \tag{4}$$

with h the local water depth, U the depth-averaged water velocity and y the transversal coordinate. This model also enables an estimation of the cross sectional momentum distribution β parameter:

$$\beta = \frac{A \int_A U^2 h \, dy}{Q^2} = \frac{AS_f}{|Q|Q} \int_A \frac{h^{7/3}}{n^2} dy \tag{5}$$

3 EXAMPLES

3.1 *Experimental FCF series*

The experimental data from 50 cases measured in a smooth compound channel laboratory setup from the FCF series of the University of Birmingham (Knight 2004) have been used as additional validation. Figure 1 is a sketch of the channel cross section geometry. We shall to define U_H as the depth averaged velocity and H the maximum depth, both at the canal axis. Figure 2(a) is a plot of U versus y for this case. The velocity values predicted by the model agree with the observed values within a 20% of error. Finally, Figure 2(b) displays the variation of β with H both as predicted by the model and as estimated from the measurements. Again, the models provide a reasonable prediction.

3.2 *Experimental river flow velocity data*

A set of more than 200 water velocity data, measured at a gauging station by the Statistics and Gauging Service of the Ebro River Water Authority (*Confederación Hidrográfica del Ebro*, Spain), have been used in the second case presented. The data were collected at an irregular

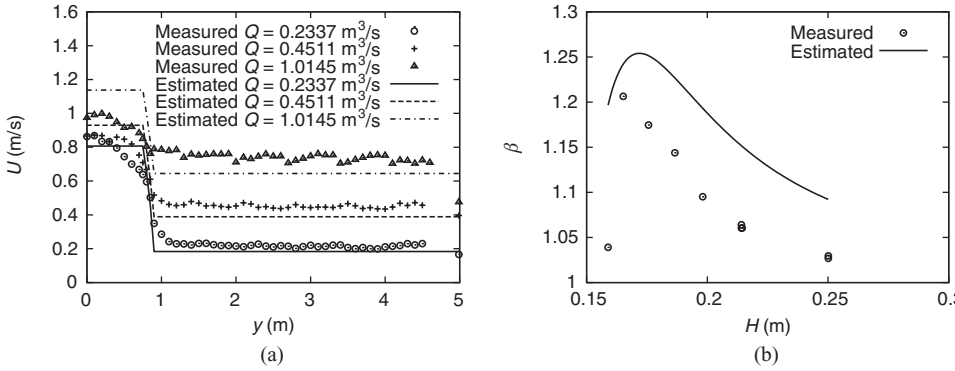


Figure 2. (a) Depth-averaged water velocity, (b) comparison of the β parameter versus depth both measured and predicted by model in the serie I of the FCF channel.

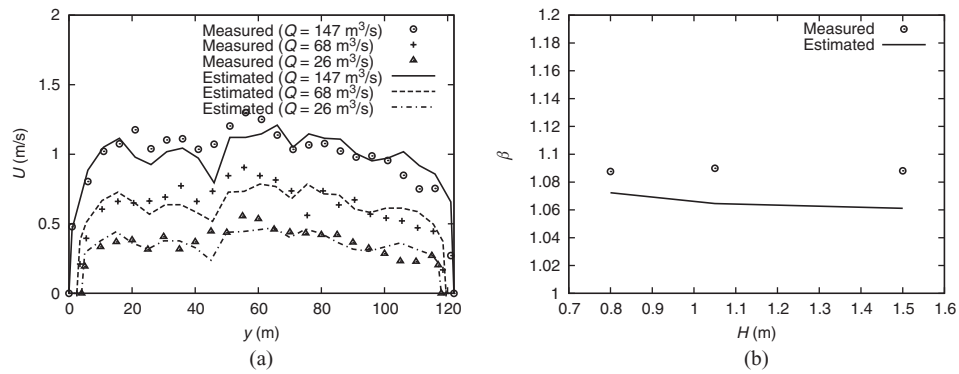


Figure 3. (a) Depth-averaged water velocity, (b) comparison of β coefficient versus depth both measured and predicted by model in the Ebro River cross section.

section of the Ebro River by means of a probe located at different depths and transverse position within the section, under three different water discharge conditions (26, 68 and 147 m^3/s). None of the discharges was extraordinary so that water was always in the main channel. The hypothesis made was that the bed river roughness is approximately the same all over the cross section.

Figure 3(a) represents the comparison of the measured and predicted U velocity values along the cross section for the three discharges. The model prediction is close to the experimental data. Finally, Figure 3(b) displays the estimate of the β value from the measured data as function of H for the three discharges. It is compared with the predictions of this parameters provided by the models. Despite the good agreement of the velocity profiles supplied by the model they slightly underestimate the value of β. This can be explained by the simplification made assuming that there is a uniform value of roughness coefficient in the cross section.

3.3 *Friction models in typical river cross sections*

The model predictions of the β values at four typical Ebro River cross sections will be next compared. The cross sections are represented in Figure 4. The value of the roughness coefficient used is $n = 0.03\ sm^{-1/3}$, taken from the literature (Chow 1959). Figure 5(a) shows the plot of the β values predictions at the four river sections. The value reached by this coefficient, as obtained with the

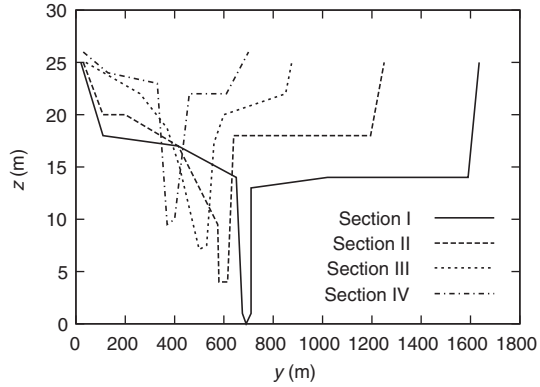


Figure 4. Cross sectional form of four typical Ebro River sections.

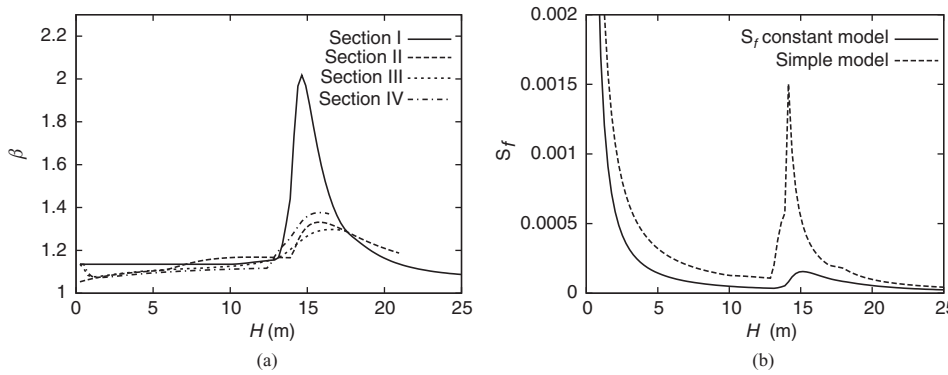


Figure 5. (a) Estimated values of the β coefficient in the four sections, (b) comparison of the S_f in the section I by simple (3) and constant friction slope (4) models, both as a function of the water depth.

friction model, is close to 2 in some cases, indicating that the basic approximation of this coefficient by 1 is not realistic and can lead to inaccurate models. Having stated the acceptable validity of the model predictions under laboratory conditions, this is a good example that provides an evidence of the values of β in realistic conditions. In Figure 5(b) S_f is plotted versus H, assuming a constant velocity $U = 1$ m/s according to the simple classical form, equation (3), and the S_f constant model, equation (4). Apart from the great discrepancy observed when the main channel is overflown, it can also be observed that the constant friction slope model predicts rather lower friction values even at low discharges.

4 CONCLUSIONS

In irregular or compound cross sections, a hypothesis has been analysed for the cross sectional velocity distribution based on the constant friction slope across the wetted cross section. The experimental validation has been helpful to show the agreement with the predictions of the model. The cross sectional velocity distributions predicted can also be used to estimate the factor β. Our calculations show that this factor must be evaluated since it can reach values much higher than the usually assumed value of $\beta = 1$ in cases of flood over a floodplain. Furthermore, it has been proved that the friction values are smaller when considering an irregular cross sectional velocity distribution than those predicted by the simpler uniform velocity assumption. Therefore, and taking

into account that most of the available friction coefficient values in the literature derive from the simple model, it can be stated that when using both the constant slope model in 1D problems or in 2D models the value of the roughness coefficient must be increased if an agreement with the simple 1D model is sought.

REFERENCES

Burguete, J., P. García-Navarro, and J. Murillo (2007). Friction term discretization and limitation to preserve stability and conservation in the 1d shallow-water model: application to unsteady irrigation and river flow. *ASCE Journal of Hydraulic Engineering in press*.

Cao, Z., J. Meng, G. Pender, and S. Wallis (2006). Flow resistance and momentum flux in compound open channels. *ASCE Journal of Hydraulic Engineering 132*(12), 1272–1282.

Chow, V. T. (1959). *Open channel hydraulics*. New York: McGraw Hill.

Knight, D. W. (2004). FCF Series, flow database at the University of Birmingham, <www.flowdata.bham.ac.uk> (Downloaded Jan. 22, 2007).

Knight, D. W. (2006). *River flood hydraulics: validation issues in one-dimensional flood routing models*. London: Taylor & Francis.

MacGahey, C., P. G. Samuels, and D. W. Knight (2006). A practical approach to estimating the flow capacity of rivers - application and analysis. Lisboa. River Flow: Taylor & Francis.

Shiono, K., Y. Muto, D. W. Knight, and A. F. L. Hyde (1999). Energy losses due to secondary flow and turbulence in meandering channels with overbank flow. *Journal of Hydraulic Research 37*(5), 641–664.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Conservative numerical boundary conditions

J. Burguete
Suelo y Agua, Estación Experimental Aula Dei, CSIC, Zaragoza, Spain

P. García-Navarro & J. Murillo
Mecánica de Fluidos, CPS, Universidad de Zaragoza, Zaragoza, Spain

ABSTRACT: A recent technique for the numerical boundary conditions in 1D shallow water flow models is able to preserve the good properties of a conservative scheme used for the interior points. The implementation in a conservative scheme is shown in a test cases with exact solution and applied to the simulation of a real river flood wave leading to very satisfactory results.

1 INTRODUCTION

In order to reach the correct numerical solution in a general unsteady case, it is clear that the numerical scheme chosen for the interior points is as important as the method applied to discretize the boundary conditions. As for the boundary conditions, the starting point is the theory of characteristics. It provides a clear idea of the space-time directions followed by the information, of the influence regions and of the required number of boundary conditions to have a well posed problem (Kreiss 1970). When, in the $x-t$ plane, a characteristic curve enters the computational domain, the region of influence is exterior to the domain and an additional condition is required. This is called a physical boundary condition. On the other hand, if the curve leaves the domain, the region of influence falls within the computational domain and the boundary condition depends exclusively in the interior variables. These are called numerical boundary conditions and their correct discrete representation is the main objective of this work.

The usual methods to discretize the numerical boundary conditions generate error in the global mass conservation. In this work, a recent method based on the integral form of the mass conservation equation extended to the full domain (global mass conservation) is also considered. It was first introduced in (Burguete et al. 2002; Burguete et al. 2006) and preliminarily used. The technique can be adapted to any conservative numerical scheme used for the interior points leading to a machine accuracy global mass conservation in all steady or unsteady situations.

2 METHODOLOGY

2.1 *Basic equations*

1D shallow water flows can be modelled by means of the Saint-Venant equations. The conservative form of such equations admits a vectorial expression as follows:

$$\frac{\partial(\mathbf{U})}{\partial t}+\frac{\partial(\mathbf{F})}{\partial x}=\mathbf{H} \tag{1}$$

with $\mathbf{U}$ the vector of conserved variables, $\mathbf{F}$ the fluxes and $\mathbf{H}$ the source terms:

$$\mathbf{U}=\begin{pmatrix} A \\ Q \end{pmatrix}, \quad \mathbf{F}=\begin{pmatrix} Q \\ \dfrac{Q^2}{A}+gI_1 \end{pmatrix}, \quad \mathbf{H}=\begin{pmatrix} 0 \\ g[I_2+A(S_0-S_f)] \end{pmatrix} \tag{2}$$

where A is the wetted cross section, Q the discharge, g the acceleration constant of gravity, I_1 and I_2 represent the pressure force integrals, S_0 the bed slope and S_f the friction slope that is commonly modelled with the Gauckler-Manning law (Chow 1959). Defining the Jacobian of the conservative flux:

$$\mathbf{J} = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \begin{pmatrix} 0 & 1 \\ c^2 - u^2 & 2u \end{pmatrix} \tag{3}$$

Here $u = \frac{Q}{A}$ is the flow velocity, $c = \sqrt{g\frac{A}{B}}$ the speed of the small surface waves and B the wetted cross section top width. The eigenvectors of $\mathbf{J}$ are:

$$\lambda_1 = u + c, \quad \lambda_2 = u - c \tag{4}$$

and the matrices which diagonalize the jacobian:

$$\mathbf{P} = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad \mathbf{J} = \mathbf{PLP}^{-1} \tag{5}$$

2.2 *Numerical scheme*

Having stated that our main goal is the derivation of a globally exact conservative scheme for the shallow water equations, conservative numerical methods are required both for the interior and the boundary points. The conservative form of the system of differential equations (1) can be solved by means of a conservative scheme. The following vector is defined for convenience (Burguete and García-Navarro 2001):

$$\mathbf{G}_{i+(1/2)} = \left(\mathbf{H} - \frac{\delta \mathbf{F}}{\delta x} \right)_{i+(1/2)} \tag{6}$$

The unified discretization of the friction term consists of discretizing flux, friction and other source terms in the equation in a similar form. A general conservative scheme with unified discretization can be expressed as (Burguete and García-Navarro 2001):

$$\Delta \mathbf{U}_i^p = \Delta t (\mathbf{G}_{i-(1/2)}^L + \mathbf{G}_{i+(1/2)}^R) \tag{7}$$

where $\mathbf{G}^{L,R}$ represents the wave discretization particular to each numerical scheme. We shall definine the upwind matrices as:

$$\mathbf{O}^{\pm} = \frac{1}{2}\mathbf{P} \begin{pmatrix} 1 \pm \text{sign}(\lambda_1) & 0 \\ 0 & 1 \pm \text{sign}(\lambda_2) \end{pmatrix} \mathbf{P}^{-1} \tag{8}$$

Then, the wave decomposition of the explicit upwind scheme is (Burguete and García-Navarro 2001):

$$\mathbf{G}_{i+(1/2)}^L = \left(\mathbf{O}^+ \mathbf{G} - \nu \frac{\delta \mathbf{U}}{\delta x} \right)_{i+(1/2)}, \quad \mathbf{G}_{i+(1/2)}^R = \left(\mathbf{O}^- \mathbf{G} - \nu \frac{\delta \mathbf{U}}{\delta x} \right)_{i+(1/2)} \tag{9}$$

with ν the entropy artificial viscosity as in (Burguete et al. 2006).

2.3 *Numerical boundary conditions*

Numerical boundary conditions are those additional equations required to enable the correct numerical resolution at the boundary points. They carry some kind of information from the calculation domain. The method of global mass conservation (Burguete et al. 2002; Burguete et al. 2006) for 1D schemes is based on enforcing the integral form of the mass conservation extended to all the computational domain in combination with a conservative scheme for the interior points to generate the numerical boundary condition. This method is sensitive to the form the mass of the system is evaluated and to the physical boundary conditions.

In a domain discretized using N cells, the mass increment ΔM^n in one time step is defined as:

$$\Delta M^n = \sum_{i=1}^{N} \Delta A_i^n \delta x \tag{10}$$

In a first step, a conservative scheme is used all over the domain neglecting contributions from outside cells. The cross section increments predicted in one time step are:

$$\Delta A_i^p = -\frac{\Delta t}{\delta x}(\delta Q^R_{i+(1/2)} + \delta Q^L_{i-(1/2)}) \tag{11}$$

The total numerical mass variation ΔM^p produced by the scheme considered is (Burguete, García-Navarro, and Aliod 2002):

$$\Delta M^p = \sum_{i=1}^{N} \Delta A_i^p \delta x = \Delta t(Q_1^n - Q_N^n) \tag{12}$$

Since the schemes used are conservative, this variation is only due to the boundaries and can be split into numerical contribution at the inlet ΔM^p_{in} and at the outlet ΔM^p_{out} in the following form:

$$\Delta M^p = \Delta M^p_{in} + \Delta M^p_{out}, \quad \Delta M^p_{in} = \Delta t Q_1^n, \quad \Delta M^p_{out} = -\Delta t Q_N^n \tag{13}$$

If the physical boundary condition is, for instance, a certain mass input at the inlet $\Delta \mathrm{M}^n_{in}$ or at the outlet in ΔM^n_{out}, in order to ensure the global mass conservation of the scheme the numerical mass increment must be corrected. This is achieved by (Burguete et al. 2006):

$$A_1^{n+1} = A_1^p + \frac{\Delta M^n_{in} - \Delta t Q_1^n}{\delta x}, \quad A_N^{n+1} = A_N^p + \frac{\Delta M^n_{out} + \Delta t Q_N^n}{\delta x} \tag{14}$$

Furthermore, for the inner points:

$$A_i^{n+1} = A_i^p, \quad Q_i^{n+1} = Q_i^p, \quad \forall\, i \neq 1, N \tag{15}$$

3 APPLICATIONS

3.1 *Test with analytical solution: discontinuous boundary conditions*

One of the best forms to study the performance of the numerical schemes is to compare their numerical solution with the analytical solution in problems in which this exists. If a discontinuity is enforced to enter the upstream boundary of an ideal flat and smooth channel of uniform depth and discharge so that the flow properties on both sides obey the Rankine-Hugoniot equations, the discontinuity will propagate unchanged along the channel at a constant speed. From the initial conditions, the water depth h_i and velocity u_i are known in the channel. Imposing h_* or u_* at the boundary, the other variable and the front speed can be worked out solving these equations for

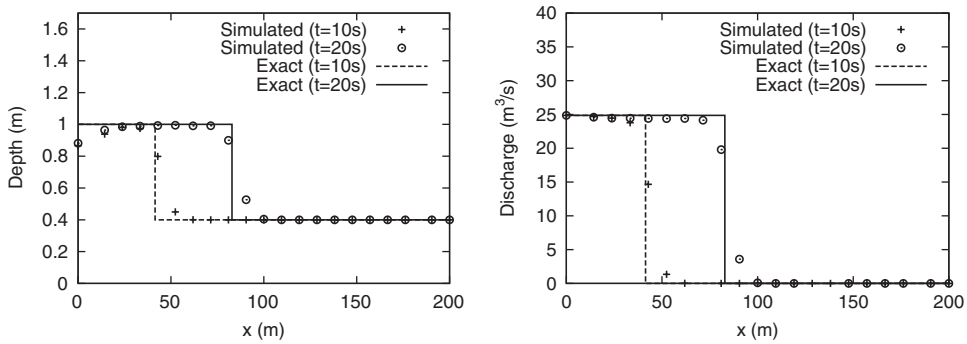


Figure 1. (a) Water depth, (b) discharge using the upwind scheme, $\delta x = 10$ m, CFL $= 0.9$ and the global mass conservation method for the upstream numerical boundary condition.

the discontinuity. In the case of a prismatic rectangular channel and still water initial conditions ($u_i = 0$) (Burguete et al. 2006):

$$U = \sqrt{g\frac{h_*}{h_i}\frac{h_* + h_i}{2}}, \quad u_* = \left(1 - \frac{h_i}{h_*}\right)U, \quad Q_* = \sigma_0\,(h_* - h_i)U \tag{16}$$

The propagation of a non-transcritical discontinuity in a rectangular channel 10 m width, with $h_i = 0.4$ m, $h_* = 1$ m, $Q_* = 24.86021$ m^3/s and $U = 4.143385$ m/s is used to study the accuracy of the resulting combinations of proposed numerical schemes for the interior points and methods for the numerical boundary conditions.

Figure 1 displays the numerical results obtained in this test using the explicit upwind scheme and the global mass conservation approach for the discretization of the upstream numerical boundary condition. This method achieves a reduction in the mass error to machine precision.

3.2 *River flow application: Ésera River*

Ésera River is a Spanish river on the left bank of the Ebro River basin. It flows through a touristic mountain area of the Pirynees. A study has been carried on led by the interest to evaluate the risk of inundation in a few nearby camping sites. It is a typical, irregular and sloping mountain river (average slope around 4%). 32 measured cross sections were used to define the 1 km long river reach bed form, a base discharge of 10 m^3/s to state the initial flow conditions, and a high roughness Manning coefficient $n = 0.03$ m$^{-1/3}$s to model the stony bed according to Chow (1959).

Four flooding discharges of 236, 344, 414 and 563 m^3/s, corresponding to return periods 10, 50, 100 and 500 year respectively were assumed in the study. There was no field information either at the reach outlet so that, due the sloping character of the river bed, a critical outflow condition was assumed in case of downstream subcritical flow. At the upstream boundary, the global mass conservation method was used for the numerical condition discretization.

The calculation was performed in two steps. In a first step, the steady base flow was determined by sudden introduction of the base discharge over dry bed initial conditions. After convergence, this led to the initial conditions for the second run. Figure 2 is a plot of the water depth profiles corresponding to the base flow discharge and the four inundation discharges all flowing in steady state as computed with the explicit upwind scheme.

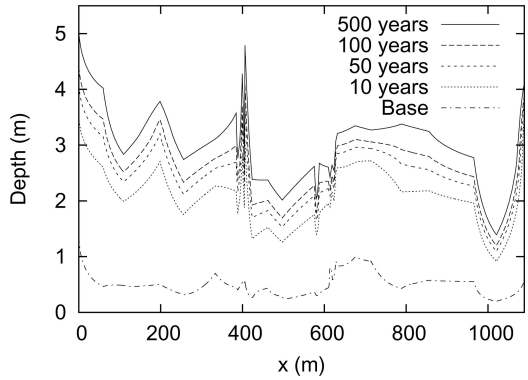


Figure 2. Several steady flow water depth profiles in the Ésera River reach as obtained from the upwind scheme using $\delta x = 1$ m and CFL $= 0.9$.

4 CONCLUSIONS

The global mass conservation method for the numerical boundary condition discretization is simple, is able to solve accurately unsteady flow with discontinuities at the boundary and leads to an exactly conservative solution when combined with any conservative explicit scheme for the interior points.

REFERENCES

Burguete, J. and P. García-Navarro (2001). Efficient construction of high-resolution tvd conservative schemes for equations with source terms: application to shallow water flows. *International Journal for Numerical Methods in Fluids 37*(2), 209–248.

Burguete, J., P. García-Navarro, and R. Aliod (2002). Numerical simulation of runoff from extreme rainfall events in a mountain water cachtment. *Natural Hazards in Earth System Sciences 2*, 1–9.

Burguete, J., P. García-Navarro, and J. Murillo (2006). Numerical boundary conditions for globally mass conservative methods to solve the shallow-water equations and applied to river flow. *International Journal for Numerical Methods in Fluids 51*(6), 585–615.

Chow, V. T. (1959). *Open channel hydraulics*. New York: McGraw Hill.

Kreiss, H. O. (1970). Initial boundary value problem for hyperbolic systems. *Communication of Pure and Applied Mathematics 23*, 277–298.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

A coupled 1D-2D modelling study of the 1907 flood in the Ebro delta

J. Raso & J. Hunink
Técnica y Proyectos, S.A. (TYPSA), Barcelona, Spain

ABSTRACT: In October 1907 occurred one of the biggest floods in the Ebro delta of the last 500 years. Until today, flood reconstruction and discharge calculations have been difficult as there was no numerical model capable of managing the amount of information of the entire Ebro floodplain. The advances in commercial software and integrated 1D-2D modelling techniques of the last 5 years have allowed reproducing the 1907 historical flood in the Ebro delta, achieving a good agreement with measured water level marks. This article reviews some key aspects of the historical flood modelling in the Ebro delta.

1 INTRODUCTION

The Ebro is the largest river of Spain, and has a long record of flooding events. The catchment went through various deforestation stages since the Roman period. This increased the sediment loads of the Ebro tributaries extremely during the last centuries. The sub-humid Mediterranean climate involves many local extreme rainfall events during the year. The resulting floods formed a huge delta, mainly during the last 500 years.

During the last century, the catchment has been regulated through the construction of numerous dams, converting some of its tributaries in one of the most regulated European rivers. Especially those affluents coming from the Pyrenees with considerable snowmelt contributions have been exploited by private companies for hydropower production.

Besides, the catchment flow regime and runoff coefficients have been altered by the increasing surface occupied by agricultural activities and the extensive irrigation network of which the implementation already started in the 19th century.

2 FLOOD HISTORY

2.1 *The dams effect*

Until the last century, the population near the river was used to suffer the impacts of several flood events during their lives. For this reason, all the detached houses had their 2nd floor prepared to "survive" a flood event during several days with the basic necessities (drinking water, toilette, food and other first-necessity material were available on that floor).

During the last 50 years, due to the presence of the big dams in the basin, floods with regular frequencies have been moderated and lowered, giving a false safety sensation to the floodplains' population and at the same time allowing expansion of the urban areas.

2.2 *Historical data*

In the past, water marks have been placed in many places along the river, some of them very close to the cities (Xerta, Tortosa, Amposta, etc).

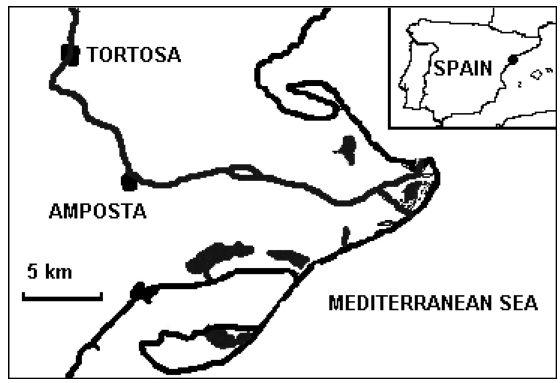


Figure 1. Situation of the Ebro delta.

Besides, the October 1907 flood event was studied in detail by Garcia Faria (1908), giving a considerable amount of data from this event, including flux velocities in the main channel, wetted perimeter and computed discharges each day during the month in the city of Tortosa.

3 MODELLING PROCESS

3.1 *1D modelling techniques and calibration*

A one-dimensional modelling approach was used in 2002 to delineate the 50, 100 and 500 year return period flood lines. However, this approach was not realistic taking into account the hydro-morphology of the Ebro delta: the fluxes can be considered highly two-dimensional. Indeed, computed water-levels were two meters higher than historical water marks using the discharges measured in 1907 with the current bathymetry, without even accounting for the lowering of the bathymetry since then, due to the sediment retention in upstream dams.

Consequently, the calibration process of the model parameters (e.g. manning roughness) and boundary conditions did not reach a good agreement between observed and calculated water levels.

3.2 *A 2D model for the Ebro delta*

A two-dimensional modelling analysis of the Ebro delta has been carried out based on a 30×30 m DEM. This model has been further scaled up in order to obtain a good agreement between computing velocity and result precision, resulting in a 300×300 m grid that covers an area of 291 km^2.

Obviously, this model is not capable of reproducing realistic channel flow response. Therefore, a finer scale is necessary in the river channel in order to compare results with measured water levels.

However, besides the main channel, secondary channels can also play an important role during big flood events. Therefore, a model is necessary that is able to represent both scales, simulating flow in 2 dimension on the floodplain. These features are included in the overland flow model developed by Delft Hydraulics in the SOBEK-Rural software package which is the selected model for the study.

3.3 *The 1D-2D coupling approach*

The resolution needed for river channel modelling requires the use of a coupled 1D-2D modelling approach. Mainly because of scale-issues, this kind of numerical algorithms tends to give stability problems in the 1D-2D models connection. For this reason, the choice of an efficient solution scheme is essential in order to avoid instabilities and on the same time have reasonable computing times.

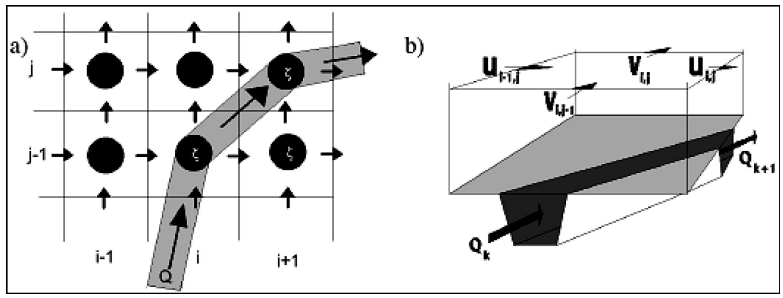


Figure 2. Schematisation of the hydraulic model: (a) combined 1D/2D staggered grid; (b) combined finite mass volume for 1D/2D computations.

Implicit coupling of 1D and 2D schematisations are based on the concept that both have to be defined on separate computational layers, similar to the concept of the storage of spatial data on different layers in a GIS.

The 2D layer describes, on the basis of a rectangular computational grid, the flow on a DTM-defined topography accounting for possible objects that block flow on the floodplains, such as dikes and natural levees. All sub-grid conveyance objects, such as channels, including remnants of dead river branches and similar local depressions, and all kinds of hydraulic structures are described on the 1D schematisation layer. Subsequently, 1D and 2D schematisations will be linked to each other via water level compatibility at selected computational nodes (Fig. 2).

It was found that this schematisation, implemented in SOBEK Rural by Delft Hydraulics, has a good balance between computing time and accuracy for the Ebro delta.

Both computational layers have finite difference formulations for volume and momentum balance equations, based upon a staggered grid approach. The water level compatibility between both layers is obtained by adding fluxes from the 1D and 2D schematisations through substitution into a shared continuity equation at the common grid points where water levels are defined. More details can be found in the references.

4 CALIBRATION PROCESS AND RESULTS

4.1 *Model calibration*

As mentioned above, the October 1907 flood was extensively documented by P. Garcia Faria in 1908. Water levels, wetted perimeters and channel velocities were recorded and discharges were computed for each day of October. For this reason, the discharge hydrograph could be introduced in the 1D-2D model in the city of Tortosa.

The first step was to recompute discharges introducing more realistic flow velocities on the floodplains, as in 1908 the discharges were computed directly multiplying wetted perimeters by channel velocities in the water surface. The model showed to be very useful for deducing floodplain velocities based on the digital topography.

The most important contribution of this step is that peak discharges of 23.000 m^3/s computed in 1908 have been drastically reduced to about 9.000 m^3/s taking into account more realistic average velocities on floodplains.

4.2 *Results*

Downstream of the city of Tortosa, the delta starts at the city of Amposta. The well-known suspension bridge is placed in a good control section, and it was constructed from 1915 to 1921. In spite of that, the cross-section has remained very similar to the one in 1907.

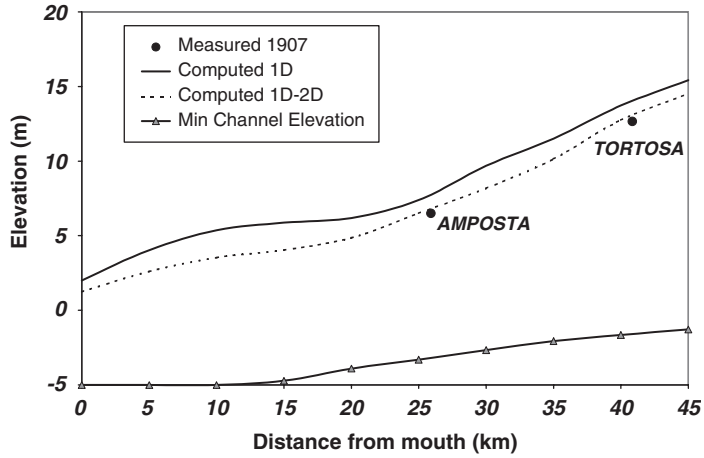


Figure 3. 1907 water level profiles compared with water marks.

Another thing to take into account is the altered river channel bathymetry because of the upstream dam constructions that lowered sediment loads. Because of this, the channel is more than 5 m deeper than in 1907. However, this effect can be partially neglected due to the erosion in the first 20 days of 1907 flood, before the peak discharge took place. This process transformed the cross-sections consistently with capacities similar to the current ones.

Taking this into consideration, the computed water levels at the Amposta cross-section bridge are less than 0.3 m higher than water marks, up to 6 meters over minimum river levels. This result is a balance between historical data interpretation, evaluation and implementation of new river modelling techniques and extensive interpretation of results in GIS tools.

5 CONCLUSIONS

The advances of the last 5 years in commercial software and integrated 1D-2D modelling techniques has allowed to reproduce the 1907 historical flood in the delta of the Ebro river, achieving a good agreement with measured water level marks in the cities of Tortosa and Amposta.

Results have confirmed the new reconstructed 1907 flood hydrograph based on the measured channel velocities and water levels in the city of Tortosa, lowering computed peak discharges from 23.000 m^3/s to 9.000 m^3/s.

REFERENCES

Dhondia J.F. & Stelling G.S. (2002). Application of One-dimensional two-dimensional integrated hydraulic model for flood simulation and damage assessment, Proceedings of the Fifth International Conference in Hydroinformatics, Cardiff , UK, July 2002, vol.1 pp. 265–276.

Frank, E., et al. (2001) Use of an integrated one-dimensional two-dimensional hydraulic Modelling approach for flood hazard and risk mapping. River Basin Management, eds R.A. Falconer & W.R. Blain, WIT Press, Southampton, UK, pp. 99–108.

Garcia-Faria, P. (1908). “Inundaciones de Octubre de 1907 en Catalunya. Previsiones Futuras”. Madrid.

Verwey, A, (2001), Latest Developments in Floodplain Modelling – 1D/2D Integration, 6th Conference on Hydraulics in Civil Engineering, Hobart, Australia.

WL|DELFT HYDRAULICS (2004), SOBEK Rural Reference Manual for Overland Flow. WL|DELFT HYDRAULICS, Delft, The Netherlands.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Numerical solution of transient flow problems in soils: The nonlinear conductivity model

A. Arrarás, L. Portero & J.C. Jorge
Departamento de Ingeniería Matemática e Informática, Universidad Pública de Navarra, Pamplona, Spain

ABSTRACT: This work deals with the efficient numerical solution of nonlinear transient flow problems posed on a two-dimensional soil profile of general geometry. The spatial semidiscretization of such problems is achieved by using a cell-centered finite difference scheme on a logically rectangular grid. The resulting nonlinear stiff initial-value problem is then integrated in time by means of a fractional step method, combined with a decomposition of the flow domain into a set of overlapping subdomains. The totally discrete scheme involves a set of nonlinear systems which can be linearized by means of suitable Taylor expansions. Thus, the proposed algorithm reduces the original problem to the solution of several linear systems per time step. Moreover, each one of such systems is a set of uncoupled linear subsystems which can be solved in parallel. A numerical example illustrates the behaviour of the method in the last section of the paper.

1 INTRODUCTION

Darcian water flow through non-swelling isothermal soils has been shown to obey Richards' equation (Celia et al. 1990). A simplified version of such equation, together with suitable initial and boundary conditions, gives rise to nonlinear parabolic initial-boundary value problems of the following form:

$$\begin{cases} \dfrac{\partial \psi(\mathbf{x},t)}{\partial t} = \mathbf{div}(K(\psi)\,\mathbf{grad}\psi(\mathbf{x},t)) + g(\psi) + f(\mathbf{x},t), & (\mathbf{x},t) \in \Omega \times (0,T], \\ \psi(\mathbf{x},0) = \psi_0(\mathbf{x}), & x \in \Omega, \\ \psi(\mathbf{x},t) = \psi_D(\mathbf{x},t), & (\mathbf{x},t) \in \partial\Omega \times (0,T], \end{cases} \tag{1}$$

where $\psi(\mathbf{x}, t)$ denotes the pressure head, $K(\psi)$ is a 2×2 symmetric positive-definite tensor which models nonlinear hydraulic conductivity, $g(\psi)$ is a smooth nonlinear function which may describe, for instance, root water uptake in soil profiles (Šimůnek et al. 2001) and $f(\mathbf{x}, t)$ is a source/sink term. The flow domain is assumed to be a bounded open set $\Omega \subseteq \mathbb{R}^2$ with boundary $\partial\Omega$, where we have considered Dirichlet boundary conditions.

This paper focuses on the design of an efficient method to numerically solve (1) by means of two discretization stages (in space and time) which are briefly explained in sections 2 and 3, respectively. The last section includes a numerical experiment which illustrates the main advantages of the numerical algorithm.

2 SPATIAL SEMIDISCRETIZATION

The semidiscrete scheme related to (1) will be obtained by using a finite difference spatial discretization based on the support-operator method. Such method, initially developed by Samarskiĭ

et al. (1981) and subsequently discussed by Shashkov (1996), provides a methodology for constructing discrete analogs of invariant first-order differential operators which appear in (1) (i.e., divergence and gradient).

Let us firstly discretize Ω by means of a logically rectangular grid Ω_h, where h denotes the spatial mesh size. The structure of such grid is indexed as follows: if N_x and N_y are positive integers, then the (i, j)-node is given by its coordinates $(x_{i,j}, y_{i,j})$, for $i = 1, 2, \ldots, N_x$ and $j = 1, 2, \ldots, N_y$. The quadrangle defined by the nodes (i, j), $(i+1, j)$, $(i, j+1)$ and $(i+1, j+1)$ is called the (i,j)-cell. The spatial semidiscretization used here considers cell-centered approximations for $\psi(\mathbf{x}, t)$, $g(\psi)$ and $f(\mathbf{x}, t)$ given by $\psi_h(t)$, $g_h(\psi_h)$ and $f_h(t)$, respectively. On the other hand, we use nodal discretizations for vector functions $\mathbf{w}(\mathbf{x}, t) \equiv (w^x(\mathbf{x}, t), w^y(\mathbf{x}, t))$ by means of $\mathbf{w}_h(t) \equiv (w_h^x(t), w_h^y(t))$.

For this approximation scheme, it is natural to use the divergence as the first-order *prime operator* (Shashkov & Steinberg 1997). Based on the invariant definition of the divergence, we can derive a discrete analog $\mathbf{div}_h$ of **div** in such a way that, for $i = 1, 2, \ldots, N_x - 1$ and $j = 1, 2, \ldots, N_y - 1$, we obtain the following expression for a semidiscrete vector

$$\begin{aligned}(\mathbf{div}_h \mathbf{w}_h(t))_{i,j} = \frac{1}{2\sigma_{i,j}}&(((w_h^x)_{i+1,j+1} - (w_h^x)_{i,j})(y_{i,j+1} - y_{i+1,j}) - ((w_h^x)_{i,j+1} - (w_h^x)_{i+1,j})(y_{i+1,j+1} - y_{i,j}) \\ &- ((w_h^y)_{i+1,j+1} - (w_h^y)_{i,j})(x_{i,j+1} - x_{i+1,j}) - ((w_h^y)_{i,j+1} - (w_h^y)_{i+1,j})(x_{i+1,j+1} - x_{i,j})),\end{aligned} \tag{2}$$

where $\sigma_{i,j}$ denotes the area of the (i,j)-cell and $(w_h^z)_{i,j}$ is the $((i-1)(Ny-1)+j)$-th component of $w_h^z(t)$ which approximates $w^z(x_{i,j}, y_{i,j}, t)$ for $z = x, y$.

Considering a discrete version of Gauss' theorem, together with the expression of $\mathbf{div}_h$, we shall construct the *derived operator* $\mathbf{grad}_h$ as the discrete analog of **grad** as follows:

$$\begin{aligned}(\mathbf{grad}_h^x \psi_h(t))_{i,j} = \frac{1}{2\eta_{i,j}}&((y_{i,j+1} - y_{i+1,j})(\psi_h)_{i,j} + (y_{i-1,j} - y_{i,j+1})(\psi_h)_{i-1,j} \\ &+ (y_{i+1,j} - y_{i,j-1})(\psi_h)_{i,j-1} + (y_{i,j-1} - y_{i-1,j})(\psi_h)_{i-1,j-1}),\end{aligned} \tag{3}$$

for $i = 2, \ldots, N_x - 1$ and $j = 2, \ldots, N_y - 1$, where $\eta_{i,j} = 0.25(\sigma_{i,j} + \sigma_{i-1,j} + \sigma_{i,j-1} + \sigma_{i-1,j-1})$ and $(\psi_h)_{i,j}$ is the $((i-1)(Ny-1)+j)$-th component of $\psi_h(t)$ which approximates $\psi(x_{i,j}, y_{i,j}, t)$. A similar expression holds for $(\mathbf{grad}_h^y \psi_h(t))_{i,j}$.

Finally, using a suitable discretization of $K(\psi)$ referred to as $K_h(\psi_h)$, it is immediate to obtain a finite difference approximation of the nonlinear elliptic operator $\mathbf{A}(\cdot) \equiv \mathbf{div}(K(\cdot)\mathbf{grad}\cdot)$. We will denote $\mathbf{A}_h(\cdot) \equiv \mathbf{div}_h(K_h(\cdot)\mathbf{grad}_h\cdot)$. In the linear case studied in Shashkov & Steinberg (1997) (i.e. K does not depend on ψ), the local stencil for $(\mathbf{A}_h(\psi_h))_{i,j}$ involves the values of $\psi_h(t)$ at the centers of the cells $(i-1, j-1)$, $(i, j-1)$, $(i+1, j-1)$, $(i-1, j)$, (i, j), $(i+1, j)$, $(i-1, j+1)$, $(i, j+1)$ and $(i+1, j+1)$, as well as the values of the components of K_h at the nodes (i,j), $(i+1,j)$, $(i,j+1)$ and $(i+1, j+1)$. In our case, taking into account the dependence of the conductiviy tensor on $\psi_h(t)$, suitable approximations of $\psi(t)$ at each one of these four nodes are also needed. Such values will be obtained by means of a bivariate interpolation method as linear combinations of the nine values of $\psi_h(t)$ at the previous cell centers.

In this framework, the semidiscrete scheme can be set as a nonlinear stiff initial-value problem of the form:

$$\begin{cases} \dfrac{d\psi_h(t)}{dt} &= \mathbf{div}_h(K_h(\psi_h)\mathbf{grad}_h\psi_h(t)) + g_h(\psi_h) + f_h(t), \quad t \in (0, T], \\ \psi_h(0) &= \psi_{0h}. \end{cases} \tag{4}$$

3 TIME INTEGRATION

Let us consider Ω decomposed as the union of m overlapping subdomains, where each one of them consists of a certain number of disjoint connected components, i.e. $\Omega = \cup_{i=1}^m \Omega_i$, where

$\Omega_i = \cup_{j=1}^{m_i} \Omega_{ij}$ such that $\Omega_{ij} \cap \Omega_{ik} = \emptyset$ if $j \neq k$. Related to this decomposition, we define a smooth partition of unity consisting of m functions $\{\rho_i(\mathbf{x})\}_{i=1}^m$ which satisfy $\sum_{i=1}^m \rho_i(\mathbf{x}) = 1 \; \forall \, \mathbf{x} \in \Omega$. Each function ρ_i equals 0 outside subdomain Ω_i, it varies smoothly in those regions where Ω_i overlaps with the other subdomains and, finally, it equals 1 at those points lying on just Ω_i.

By using this partition of unity, we shall define the following splittings for the nonlinear discrete operator $\mathbf{A}_h(\cdot)$ and the semidiscrete function $f_h(t)$ (Portero et al. 2004):

$$\mathbf{A}_h(\cdot) = \sum_{i=1}^{m} \mathbf{A}_{i,h}(\cdot), \quad f_h(t) = \sum_{i=1}^{m} f_{i,h}(t), \tag{5}$$

where we denote $\mathbf{A}_{i,h}(\cdot) \equiv \mathbf{div}_h(K_{i,h}(\cdot)\mathbf{grad}_h \cdot)$. Here, $K_{i,h}(\psi_h)$ is obtained from $(\rho_i(\mathbf{x})K(\psi))$ in the same way as $K_h(\psi_h)$ was obtained from $K(\psi)$, whereas $f_{i,h}(t)$ are cell-centered evaluations of $(\rho_i(\mathbf{x})f(\mathbf{x},t))$.

Considering the splittings given in (5), a linearly implicit variant of the fractional implicit Euler method with m internal stages (Bujanda & Jorge 2006) reduces the nonlinear stiff problem (4) to the following set of nonlinear systems (one per internal stage):

$$\begin{cases} \psi_{h,0} = \psi_{0h}, & \\ \psi_{h,n}^k = \psi_{h,n} + \tau \displaystyle\sum_{\ell=1}^{k} (\mathbf{A}_{\ell,h}(\psi_{h,n}^{\ell}) + f_{\ell,h}(t_{n+1})) + \tau g_h(\psi_{h,n}), & \text{for } k = 1,2,\ldots,m, \\ \psi_{h,n+1} = \psi_{h,n}^m, & \text{for } n = 0,1,\ldots,N_T, \end{cases} \tag{6}$$

where $N_T \equiv [T/\tau] - 1$. The discrete solution $\psi_{h,n+1}$ approximates $\psi_h(t_{n+1})$, where $t_{n+1} = (n+1)\tau$ and τ denotes the constant time step. Note that the choice of a linearly implicit method entails an explicit treatment of the nonlinear function $g(\psi)$.

In order to linearize the previous systems, we approximate $\mathbf{A}_{\ell,h}(\psi_{h,n}^{\ell})$ by the two first terms of its Taylor expansion around $\psi_{h,n}$, i.e.

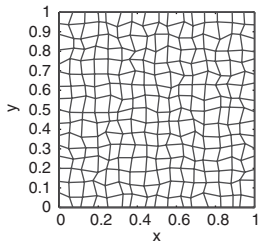
$$\mathbf{A}_{\ell,h}(\psi_{h,n}^{\ell}) \approx \mathbf{A}_{\ell,h}(\psi_{h,n}) + \left.\frac{d\mathbf{A}_{\ell,h}(\psi_h)}{d\psi_h}\right|_{\psi_h = \psi_{h,n}} (\psi_{h,n}^{\ell} - \psi_{h,n}). \tag{7}$$

If we denote by $(\psi_{h,n})_{i,j}$ the $((i-1)(N_y-1)+j)$-th component of $\psi_{h,n}$, for $i = 1,2,\ldots,N_x-1$ and $j = 1,2,\ldots,N_y-1$, then $(\mathbf{A}_{\ell,h}(\psi_{h,n}))_{i,j}$ depends nonlinearly on nine unknowns: $(\psi_{h,n})_{i-1,j-1}, (\psi_{h,n})_{i,j-1}, (\psi_{h,n})_{i+1,j-1}, (\psi_{h,n})_{i-1,j}, \quad (\psi_{h,n})_{i,j}, (\psi_{h,n})_{i+1,j}, (\psi_{h,n})_{i-1,j+1}, (\psi_{h,n})_{i,j+1}$ and $(\psi_{h,n})_{i+1,j+1}$. Therefore, the $((i-1)(N_y-1)+j)$-th row of the Jacobian matrix $\mathbf{J}_{\ell,h}(\psi_{h,n}) \equiv d\mathbf{A}_{\ell,h}(\psi_h)/d\psi_h|_{\psi_h=\psi_{h,n}}$ will contain nine non-zero elements representing the derivatives of $(\mathbf{A}_{\ell,h}(\psi_{h,n}))_{i,j}$ with respect to each one of the previous unknowns.

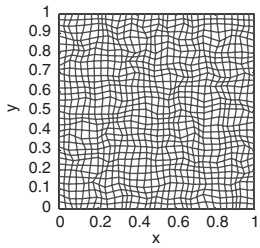
Inserting (7) into (6), we obtain the totally discrete scheme for problem (1):

$$\begin{cases} \psi_{h,0} = \psi_{0h}, & \\ (\mathbf{I} - \tau\mathbf{J}_{k,h}(\psi_{h,n}))\psi_{h,n}^k = \psi_{h,n} + \tau \displaystyle\sum_{\ell=1}^{k-1} (\mathbf{A}_{\ell,h}(\psi_{h,n}) + \mathbf{J}_{\ell,h}(\psi_{h,n})(\psi_{h,n}^{\ell} - \psi_{h,n}) & \\ \qquad + f_{\ell,h}(t_{n+1})) + \tau(\mathbf{A}_{k,h}(\psi_{h,n}) - \mathbf{J}_{k,h}(\psi_{h,n})\psi_{h,n} + f_{k,h}(t_{n+1})) + \tau g h(\psi_{h,n}), & \text{for } k = 1,2,\ldots,m, \\ \psi_{h,n+1} = \psi_{h,n}^m, & \text{for } n = 0,1,\ldots,N_T, \end{cases} \tag{8}$$

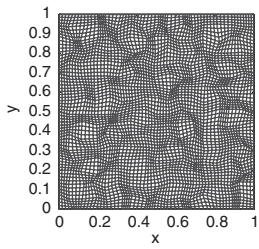
where each internal stage consists of a linear system which involves the unknowns lying on just one of the subdomains $\{\Omega_i\}_{i=1}^m$. Moreover, since each subdomain Ω_i comprises m_i disjoint connected components, this system can be immediately decomposed into m_i uncoupled subsystems which can be solved in parallel. A remarkable advantage of this technique, compared to classical domain decomposition methods, is that no Schwarz iterations are involved in the solution of (8).



(a) N =17



(b) N = 33



(c) N = 65

Figure 1. Logically rectangular grids used in the numerical experiment.

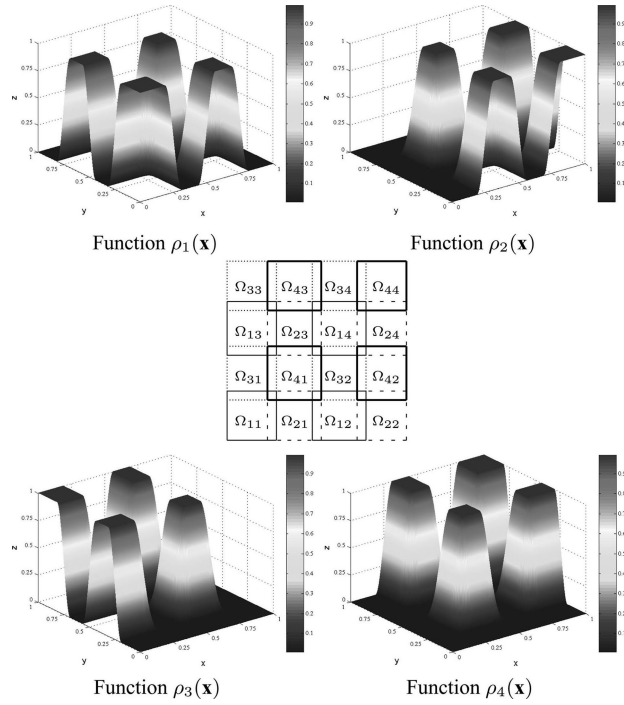


Figure 2. Decomposition of Ω into four overlapping subdomains (central plot) and corresponding functions $\rho_i(\mathbf{x})$ associated to each subdomain Ω_i, for $i \in \{1, 2, 3, 4\}$ (corner plots).

4 NUMERICAL EXPERIMENT

In this section we test the behaviour of the numerical algorithm on a set of pseudo-random logically rectangular grids. Let us consider an equation of type (1) posed on $\Omega \times (0, T] \equiv \{\mathbf{x} = (x, y) \in \mathbb{R}^2: 0 < x < 1,\ 0 < y < 1\} \times (0, 0.01]$. The hydraulic conductivity $K(\psi)$ is defined as a 2×2 diagonal tensor whose diagonal entries are set equal to $(1 + \psi^2)$, thus representing an isotropic soil. The nonlinear function is chosen to be $g(\psi) = 1/(1 + \psi^3)$, whereas the source/sink term $f(x, y, t)$ and both initial and Dirichlet boundary conditions are defined in such a way that $\psi(x, y, t) = e^{-2\pi^2 t}$ $\sin(\pi x)\ \sin(\pi y)$ is the exact solution of the problem.

The spatial semidiscretization is based on the finite difference method described in section 2. The flow domain Ω is firstly discretized by means of a pseudo-random logically rectangular grid $\Omega_h \equiv \{(xi,j, yi,j)\}_{i,j=1}^{N}$ with coordinates $x_{i,j} = (i - 1)h - 0.25h + 0.5\, h\, R_x$ and $y_{i,j} = (j - 1)h - 0.25\, h + 0.5\, h\, R_y$, where $h = 1/(N - 1)$and R_x, R_y are pseudo-random numbers generated on the interval (0, 1). Figure 1a shows an example of such type of grids for $N = 17$. In order to study the asymptotic behaviour of the error, we have successively refined the original pseudo-random grid by using the following procedure: starting from a given grid, we add the lines which connect, on each cell, the centers of its opposite sides. Figures 1b and 1c show the first two refinements for the grid of Figure 1a.

Table 1. Global errors and numerical orders of convergence in time for $N = 129$ and $\tau_0 = 10^{-3}$.

Time step, τ	τ_0	$\tau_0/2$	$\tau_0/4$	$\tau_0/8$	$\tau_0/16$	$\tau_0/32$
$E_{N,\tau}$	2.554 E-2	1.672 E-2	1.002 E-2	5.662 E-3	3.077 E-3	1.628 E-3
$p_{N,\tau}$	0.6112	0.7387	0.8235	0.8798	0.9184	–

Table 2. Global errors and numerical orders of convergence in space for $\tau = 10^{-7}$.

Number of nodes, N	17	33	65	129	257
$E_{N,\tau}$	4.506 E-3	1.886 E-3	4.595 E-4	1.144 E-4	2.857 E-5
$p_{N,\tau}$	1.2565	2.0372	2.0060	2.0015	–

Next, we consider a decomposition of Ω into $m = 4$ overlapping subdomains $\{\Omega_i\}_{i=1}^m$, each of which consists of $m_i = 4$ disjoint connected components, for $i \in \{1, 2, 3, 4\}$. On the other hand, piecewise functions $\{\rho_i(\mathbf{x})\}_{i=1}^m$ are defined at the overlapping regions by means of suitable exponential functions which give rise to a $\mathcal{C}^\infty$ partition of unity. Figure 2 shows the decomposition of Ω as well as the corresponding functions $\rho_i(\mathbf{x})$ associated to Ω_i. Subordinate to such decomposition, we use the fractional step method (6) with four internal stages ($m = 4$), as well as the linearization process described in (7), in order to obtain the totally discrete scheme (8).

We have included two tables which contain the global errors, $E_{N,\tau}$ (upper row), and numerical orders of convergence, $p_{N,\tau}$ (lower row), obtained for different values of N and τ when using the maximum norm in time and the L^2-norm in space, i.e. $||\cdot||_{L^\infty(0,T;L^2(\Omega))}$. The method shows unconditional convergence of first order in time (see Table 1) and second order in space (see Table 2).

REFERENCES

Bujanda, B. & Jorge, J.C. 2006. Stability results for linearly implicit fractional step discretizations of non-linear time dependent parabolic problems. *Appl. Numer. Math.* 56(8): 1061–1076.

Celia, M.A., Bouloutas, E.T. & Zarba, R.L. 1990. A general mass-conservative numerical solution for the unsaturated flow equation. *Water Resour. Res.* 26(7): 1483–1496.

Portero, L., Bujanda, B. & Jorge, J.C. 2004. A combined fractional step domain decomposition method for the numerical integration of parabolic problems. *Lecture Notes in Comput. Sci.* 3019: 1034–1041.

Samarskiĭ, A., Tishkin, V., Favorskiĭ, A. & Shashkov, M. 1981. Operational finite-difference schemes. *Differ. Equ.* 17(7): 854–862.

Shashkov, M. 1996. *Conservative Finite-Difference Methods on General Grids*. Boca Raton, FL: CRC Press.

Shashkov, M. & Steinberg, S. 1997. The numerical solution of diffusion problems in strongly heterogeneous non-isotropic materials. *J. Comput. Phys.* 132(1): 130–148.

Šimůnek, J., Hopmans, J.W., Vrugt, J.A. & van Wijt, M.T. 2001. One-, two- and three-dimensional root water uptake functions for transient modeling. *Water Resour. Res.* 37(10): 2457–2470.

Numerical Modelling of Hydrodynamics for Water Resources – García-Navarro & Playán (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-44056-1

Author index

Aliod, R. 293
Archambeau, P. 321
Arrarás, A. 383

Babovic, V. 67, 77
Beguería, S. 175
Bellafiore, D. 351
Bohórquez, P. 355, 361
Bouten, W. 333
Brufau, P. 47
Burguete, J. 47, 339, 367, 373

Castro, M.J. 211
Cea, L. 127

Detrembleur, S. 321
Dewals, B. 321

Erpicum, S. 321

Farmani, R. 25
Fernández-Nieto, E.D. 211
Ferreira, R.M.L. 189
Ferreiro, A.M. 211
Feyen, J. 153

García-Navarro, P. 47, 317, 339, 367, 373
García-Ruiz, J.M. 175
Ghenaim, A. 327
Ghostine, R. 327
Giustolisi, O. 25
González, C. 293

Hunink, J. 379
Hunink, J.E. 333

Kapelan, Z. 25
Kesserwani, G. 327
Knight, D.W. 3

Lana-Renault, N. 175
Latorre, B. 317

Malaterre, P.O. 271
Mosé, R. 327
Murillo, J. 47, 317, 367, 373

Nadal-Romero, E. 175

Parés, C. 211
Pirotton, M. 321
Playán, E. 339
Portero, L. 383
Puertas, J. 251

Raso, J. 379
Regüés, D. 175

Santos, E.P. 339
Savic, D. 25
Šejna, M. 95
Serrano-Muela, P. 175
Šimůnek, J. 95
Stansby, P.K. 115
Suárez, J. 251

Tzatchkov, V.G. 345

Umgiesser, G. 351

van Genuchten, M.Th. 95
van Loon, E.E. 333
Vazquez, J. 327
Vázquez-Cendón, M.E. 127
Villanueva, I. 233
Vázquez, R.F. 153

Zapata, N. 339