GeoPlanet: Earth and Planetary Sciences

Renata J. Romanowicz
Marzena Osuch *Editors*

Stochastic Flood Forecasting System

The Middle River Vistula Case Study

Springer

GeoPlanet: Earth and Planetary Sciences

More information about this series at http://www.springer.com/series/8821

Renata J. Romanowicz · Marzena Osuch
Editors

Stochastic Flood Forecasting System

The Middle River Vistula Case Study

Springer

Editors
Renata J. Romanowicz
Institute of Geophysics
Warsaw
Poland

Marzena Osuch
Institute of Geophysics
Warsaw
Poland

The GeoPlanet: Earth and Planetary Sciences Book Series is in part a continuation of Monographic Volumes of Publications of the Institute of Geophysics, Polish Academy of Sciences, the journal published since 1962 (http://pub.igf.edu.pl/index.php).

ISSN 2190-5193 ISSN 2190-5207 (electronic)
GeoPlanet: Earth and Planetary Sciences
ISBN 978-3-319-18853-9 ISBN 978-3-319-18854-6 (eBook)
DOI 10.1007/978-3-319-18854-6

Library of Congress Control Number: 2015941860

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Series Editors

Managing Editor

Anna Dziembowska
Institute of Geophysics, Polish Academy of Sciences

Advisory Board

Preface

Governments and communities around the world are increasingly adopting risk-based approaches to the planning, design and operation of systems to reduce flood damage. This approach underpinned the EU Directive (2007/60/EC) on the assessment and management of flood risk. Increasing practical experience of flood risk analysis is revealing the severe uncertainties associated with management decisions. Decisions that may have appeared sound when based on best estimates can be called into question when uncertainty is taken into account. The risk from floods is defined as the multiple of the probability of flooding and the estimated costs of flood damage. Both factors require an uncertainty analysis. Floods on the River Odra in 1997 and on the River Vistula in 2001 and 2010 serve as examples of the limitations of a purely deterministic, defensive approach to flow modelling. The immense flood in 1997 triggered a number of central government flood protection initiatives that included *inter alia* the modernisation of the monitoring system (Systems of Monitoring and Country Protection; in Polish: Systemy Monitoringu i Osłony Kraju, Lindel et al. 1997; Mierkiewicz et al. 1999; Kadłubowski 2005). This system automatically transfers measurements to the operational systems responsible for flood forecasting and protection. Flood forecasts prepared by national operational systems in the Vistula basin are based on HBV rainfall-run-off models and flow routing hydrodynamic models (Lindel et al. 1997; Kadłubowski et al. 2011). However, flood incidents in the River Vistula in Spring 2010 revealed shortcomings in the national system of flood forecasting and protection. The development of fully realistic probabilistic flood forecasting models has not yet been attempted and has been a severe drawback to management decisions and infrastructure planning. Nor has there been any attempt so far to develop probabilistic online flood forecasting systems for both long-term planning and short-term emergency warning.

The research project "Stochastic flood forecasting system (The River Vistula reach from Zawichost to Warsaw)" carried out by order of the National Science Centre (contract No. 2011/01/B/ST10/06866) in the Institute of Geophysics, Warsaw, is an example of the work on flood forecasting problems that takes into account stochastic aspects of risk assessment. This book gives a summary of the project's achievements.

The novelty of the approach consists in the integration of a number of different hydrological modelling tools, including distributed, lumped parameter, deterministic

and stochastic models that can be run separately or in parallel. The system has a modular structure, including models describing rainfall-run-off and snow-melt processes for tributary catchments and the transformation of a flood wave within the reach.

This is the first study to integrate different aspects of flow forecasting at different temporal and spatial scales. These include catchment scale rainfall-run-off modelling and flow routing processes. Catchment geography and land cover changes during the last 25 years are described by Gutry-Korycka et al. (2015; this book). Long-term changes of the river flow regime are analysed using mathematical modelling based on 50 years of daily flow and water-level records (Karamuz et al. 2015; this book). On the other hand, detailed flow and bed topography measurements performed locally at one of the gauging stations provide information on the short-term flow variability (Bialik et al. 2015; this book). Flood frequency analysis (FFA) plays an important part in flood risk assessment. However, that approach requires the assumption of a stationarity of the flow regime. Bogdanowicz et al. (2015; this book) apply a first-order Markov model to describe flood-poor and flood-rich period lengths in a Warsaw maximum flow series in order to assess the changes in flow regime. The rainfall-run-off models described by Osuch (2015a; this book) consist of the conceptual HBV models developed for the catchments located in the study area. The sensitivity and uncertainty analyses of model parameters and their predictions provide information on the dominant processes that control rainfall-run-off in the basin.

A step-by-step procedure for the distributed modelling of river flow by contrasting two commonly used modelling tools, MIKE 11 and HEC-RAS, is presented by Kochanek et al. (2015; this book). This chapter is aimed at the end-user to facilitate the application of the scenario analysis research by water management bodies (e.g. the Warsaw Water Management Board). There is a clear priority of automated calibration over the hand-tuning of model parameters, as exemplified by the MIKE11 (automatically calibrated) and HEC-RAS (hand-tuned) modelling results. The sensitivity and uncertainty analyses of the MIKE11 model predictions are presented by Osuch (2015b; this book). Three methods of sensitivity analysis are compared, namely Morris, Sobol and a correlation-based approach allowing for the choice of an appropriate model calibration strategy. The results indicate that an appropriate strategy should not consist of a series of model optimisations for independent sub-reaches but should rather be integrated over all reaches and simultaneously take into account the model fit at each controlled cross section. The 1-D modelling of the river flow presented so far is extended to a 2-D model by Magnuszewski (2015; this book), by the application of a two-dimensional hydrodynamic model CCHE2D to model flow in the contemporary River Vistula channel in Warsaw. The application of the 2-D hydrodynamic model for different scenarios of floodplain maintenance has helped to predict flood conveyance in highly urbanised and vulnerable areas of Warsaw.

In the next part of the study, lumped parameter models were applied to describe flow routing in the studied river reach. A semi-distributed stochastic flood forecasting system for the middle River Vistula is presented by Romanowicz and Osuch (2015a; this book). Due to the stochastic nature of the methods applied, the confidence limits of the predictions are also estimated in a straightforward way. The ability to tackle the nonlinearity of flood wave transformation by means of the Hammerstein–Wiener modelling approach is assessed. Due to the lack of the historical series of short-time

period observations necessary to obtain short-term online flow forecasts, a stochastic transfer function-based emulator is introduced by Romanowicz and Osuch (2015b; this book). Application of a distributed flow routing model makes it possible to interpolate forecasts both in time and in space. The application of the ECMWF ensemble forecasts to prolong the forecast lead time for the Biala Tarnowska catchment, southern Poland, is presented by Kiczko et al. (2015; this book). This study presents the first application in Poland of online assimilation of ECMWF forecasts. Medium-range probabilistic weather forecasts (ECMWF) and online observations of temperature, precipitation and water levels are used to prolong the forecast lead time.

The interests of end-users are discussed by Rucinska (2015; this book) in an analysis of the social aspects of flood risk assessment. Potential end-users will benefit from a description of the social vulnerability to natural hazards in the study area.

In summary, the book covers a broad range of problems related to stochastic flow forecasting. The fourth and first chapters describe the spatial and temporal variability of flow regimes, flood frequency analysis under non-stationary conditions and river morphology based on analytical studies and measurements. These are followed by physically based modelling, including rainfall-flow and flow routing models. Particular attention is paid to the uncertainty aspects of flood predictions. The semi-distributed approach, following the idea of the data-based modelling, is presented in the next three chapters, which also provide a bridge between distributed and lumped flow routing models in the form of emulators. The ensemble forecasts presented in the last section of that group illustrate an ability to extend the flood forecast lead beyond the flood wave travel times. The social aspects arising in flood risk assessment are discussed in the final chapter.

This book aims to stimulate similar approaches in Poland and elsewhere. An important aim is to raise awareness of the need for a critical, rigorous stochastic approach to flood forecasting. It also provides the tools to decision-makers (Warsaw Regional Water Board) in the form of an integrated system for scenario analysis. Our main objective has been to raise interest and to motivate further research in this difficult area.

Acknowledgments

The authors would like to thank Ray Macdonald and Anna Dziembowska for their help in editing this book. Finally, we would like to thank all reviewers for their constructive and critical reviews which helped to improve the quality of the papers. This work was partly supported by the project “Stochastic flood forecasting system (The River Vistula reach from Zawichost to Warsaw)” carried by the Institute of Geophysics, Polish Academy of Sciences on the order of the National Science Centre (contract No. 2011/01/B/ST10/06866).

Renata J. Romanowicz
Marzena Osuch

References

Bialik RJ, Sziło J, Karpiński M, Rajwa-Kuligiewicz A, Głowacki O (2015) Bed topography and discharge measurements in the Świderskie Islands nature reserve, River Vistula, Poland. This book, chapter 4

Bogdanowicz E, Strupczewski WG, Kochanek K (2015) On the run length in annual maximum flow series in the Middle Vistula Basin in the context of climate change impact (an attempt to determine the indicator for the assessment of changes in the regime of high waters). This book, chapter 3

Gutry-Korycka M, Mirończuk A, Hościło A (2015) Land cover change in the Middle River Vistula catchment. This book, chapter 1

Karamuz E, Romanowicz RJ, Booij MJ (2015) Influence of land-use and water management practices on water levels in the Middle River Vistula. This book, chapter 2

Kadłubowski A (2005) Szkoła Tematyczna MANHAZ, Zarządzanie Zagrożeniami dla Zdrowia i Środowiska, 26–30 Sept 2005

Kadłubowski A, Mierkiewicz M, Budzyńska H (2011) Operational rainfall/snowmelt-runoff model for upper narew river. In: Modelling of hydrological processes in the Narew Catchment. Geoplanet: Earth and Planetary Sciences. Springer, Berlin, Heidelberg

Kiczko A, Romanowicz RJ, Osuch M, Pappenberger F (2015) Adaptation of the integrated catchment system to on-line assimilation of ECMWF forecasts. This book, chapter 11

Kochanek K, Karamuz E, Osuch M (2015) Distributed modelling of flow in the middle reach of the River Vistula. This book, chapter 6

Lindel S, Ericsson L, Mierkiewicz M, Kadłubowski A (1997) Integrated hydrological monitoring and forecasting system for Vistula River Basin. Final report, IMGW (Poland) and SMHI (Sweden)

Magnuszewski A (2015) Influence on flood safety of channel processes and vegetation in the River Vistula valley in Warsaw. This book, chapter 8

Mierkiewicz M, Kadłubowski A, Sasim M (1999) System of hydrological protecting in the middle and lower Vistula River Basin. Wiadomości IMGW, t. XXXII(XLII), z. 4 (in Polish)

Osuch M (2015a) Rainfall-runoff modelling in the Middle River Vistula tributaries. This book, chapter 5

Osuch M (2015b) Sensitivity analysis of the flow routing model for the Middle River Vistula-multi-method approach. This book, chapter 7

Romanowicz RJ, Osuch M (2015a) Stochastic semi-distributed flood forecasting system for the Middle Vistula reach. This book, chapter 9

Romanowicz RJ, Osuch M (2015b) Stochastic transfer function based emulator for the on-line flow forecasting. This book, chapter 10

Rucinska D (2015) Social aspects in flood risk assessment. This book, chapter 12

Contents

Contributors

Robert J. Bialik Department of Hydrology and Hydrodynamics, Institute of Geophysics, PAS, Warsaw, Poland

Ewa Bogdanowicz Project CHIHE, Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

M.J. Booij Department of Water Engineering and Management, University of Twente, Enschede, The Netherlands

Oskar Głowacki Department of Polar and Marine Research, Institute of Geophysics, PAS, Warsaw, Poland

M. Gutry-Korycka Faculty of Geography and Regional Studies, University of Warsaw, Warsaw, Poland

A. Hościło Institute of Geodesy and Cartography, Warsaw, Poland

Emilia Karamuz Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

Mikołaj Karpiński Department of Hydrology and Hydrodynamics, Institute of Geophysics, PAS, Warsaw, Poland

Adam Kiczko Warsaw University of Life Sciences—SGGW, Warsaw, Poland

Krzysztof Kochanek Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

Artur Magnuszewski Hydrology Department, Faculty of Geography and Regional Studies, University of Warsaw, Warsaw, Poland

A. Mirończuk Institute of Geodesy and Cartography, Warsaw, Poland

Marzena Osuch Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

Florian Pappenberger European Centre for Medium Range Weather Forecasts, Reading, UK

Agnieszka Rajwa-Kuligiewicz Department of Hydrology and Hydrodynamics, Institute of Geophysics, PAS, Warsaw, Poland

Renata J. Romanowicz Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

Dorota Rucinska Faculty of Geography and Regional Studies, University of Warsaw, Warsaw, Poland

Witold G. Strupczewski Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

Joanna Sziło Centre for Polar Studies, Institute of Geophysics, PAS, Warsaw, Poland

Part I
Spatial and Temporal Variability of Flow Regime and River Morphology Based on Analytical Studies and Measurements

Land Cover Change in the Middle River Vistula Catchment

M. Gutry-Korycka, A. Mirończuk and A. Hościło

Abstract This chapter presents an analysis of land cover and its changes in the Middle River Vistula catchment over the period 1990–2012. The river reach is situated between hydrological gauging stations at Zawichost and at Port Praski (Warsaw). Comprehensive information on land cover and land cover change is a crucial step in understanding and modelling hydrological processes. The land cover and land cover change were derived from the CORINE Land Cover (CLC) inventories. The land cover classes (level 3) were aggregated from the point of view of hydrological processes and anthropogenic impacts into five classes: urban, agricultural, meadow, forest, open spaces and surface water. The analysis of land cover changes was conducted for the entire Middle River Vistula catchment and for the 10 tributaries and 11 adjacent areas. The results indicate that most of the changes (≈ -25 %) displayed belong to agricultural areas. Human impacts are mostly noted in the forested areas, bushes, meadows and urban areas.

1 Importance of the Land Cover in the Modelling of Hydrological Processes in the Catchment

The aim of the present chapter is to assess the changes of land cover in the modelled catchment, within the limits of practical viable description, from the point of view of its possible influence on the processes included in water circulation in

M. Gutry-Korycka (✉)
Faculty of Geography & Regional Studies, University of Warsaw, Warsaw, Poland
e-mail: msgutryk@uw.edu.pl

A. Mirończuk · A. Hościło
Institute of Geodesy and Cartography, Warsaw, Poland
e-mail: anna.mironczuk@igik.edu.pl

A. Hościło
e-mail: agata.hoscilo@igik.edu.pl

R.J. Romanowicz and M. Osuch (eds.), *Stochastic Flood Forecasting System*,
GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-319-18854-6_1

natural environment. Nowadays, external anthropogenic pressure imposed on the hydrological cycle increases and it has become clear that mathematical modelling of the processes taking into account feedback phenomena from land usage improves forecasts and interpretation of the behaviour of the natural systems caused by anthropogenic activity (Jarsjö et al. 2012). Many authors, e.g. Schulze (2000) and Montanari et al. (2013) stress the arising need for a more detailed examination of the complex relationships between water, economic changes and terrain cover and usage. In order to achieve this goal, the cooperation of interdisciplinary research teams and use of a variety of approaches are needed. These approaches describe the operation of the hydrological systems where the stationary character of the processes is changed, which was justified by, *inter alia*, Sivapalan et al. (2012), as shown in Fig. 1.

Many authors have stated that the most important geomorphological factors determining the variation of humidity of the ground and terrain surface include: the depth of the groundwater; aggregation composition; terrain shape; climate conditions, such as sum of rainfall, air temperature and humidity; and type and status of the plants in the basin. The dominant role in the hierarchy of listed factors is played by the terrain shape—an element which is beneficial for the dynamics of the processes (Beven and Kirkby 1979; Crave and Gascuel-Odoux 1997). However, Vachaud et al. (1985) have proven earlier that the granulometry of the soil cover is the key factor which determines the spatial layout of soil moisture in the catchment. To date, there has been a great interest in applying remote sensed data and methods to derive spatial information on soil moisture, hydrological droughts and vegetation condition (Dąbrowska-Zielińska et al. 2011, 2012). Satellite images have been widely used for national and global land cover and land cover change mapping as well as for estimating the distributed rainfall-runoff model parameters.

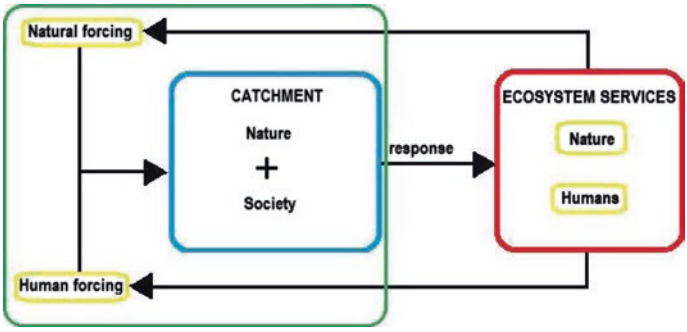

Fig. 1 Feedback between catchment, ecosystem services and natural and human forcing (modified after Sivapalan et al. 2012 and Montanari et al. 2013)

2 Structure and Location of the Study Area

Geomorphological profile is created in order to assess a variety of features that characterize the basin, which stress both differences as well as similarities, and can be subsequently used in a mathematical description of the basin using distributed or semi-distributed models. This refers to the type of characteristics of the catchments, as well as the biota covering the terrain, land cover or ground characteristics (soil, lithology). It should be noted that in case of large differentiation and heterogeneity of the landscape features (forests, agriculture areas or urbanized areas), identification of parameters is complex, and often no unique results can be achieved. The modelled system—middle section of the River Vistula—starts at the mouth of the River San (right tributary) and extends down to the biggest, also right, tributary—mouth of the River Narew; which in total forms a 256 km section of the river (Fig. 2). This includes the Vistula drainage divide with a total topographic surface of 34,981.7 km^2. It is characterized by a variety of right tributaries

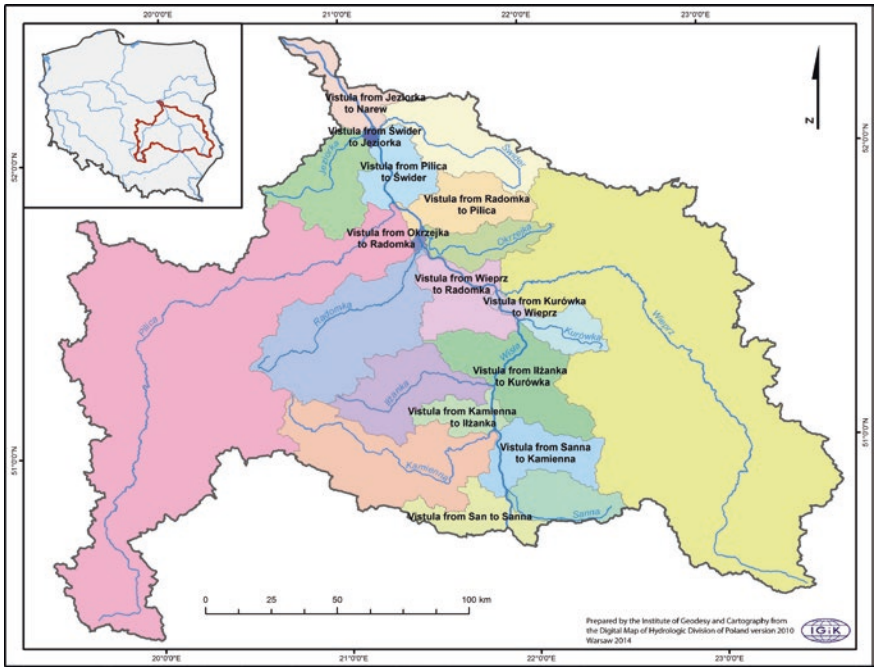

Fig. 2 A drainage divide of the Middle River Vistula catchment (acc. MPHP [MPHP—**M**apa **P**odziału **H**ydrograficznego **P**olski (Map of Hydrographical Divide of Poland 1:50,000 funded by the National Fund for Environmental Protection and Water Management)], version 2010)

(rivers: Sanna, Kurówka, Wieprz, Okrzejka and Świder), as well as left tributaries (rivers: Kamienna, Iłżanka, Radomka, Pilica and Jeziorka), which are shown in Fig. 2.

Catchments of the right tributaries take 38 % of the surface, while the catchments of the left tributaries are not significantly larger-they constitute 44.2 % of the surface, which in total results in 82.6 % of the catchment surface. The remaining part of the drainage area is taken by 11 adjacent areas, the surface of which is significantly smaller; the ratio between the surfaces is 1–4.7. Drainage area of the River Vistula middle section is laid out parallel to the latitude, with a slight asymmetry in the region of narrow, longitudinal left drainage system of River Pilica. The most significant component-drainage areas of the examined hydrological system include Rivers Wieprz and Pilica. The remaining tributaries have drainage areas far less significant.

The studied river reach stretches from Zawichost to Warsaw-Port Praski gauging stations (Fig. 2). This section includes 226 km of the river, which means that it is a bit shorter than the real length of the middle section of the River Vistula. It includes almost whole mid-section of the river, which is particularly prone to flooding and is characterised by the high riverbed dynamics, caused by numerous bents in the longitudinal profile resulting in increased steepness of the slope. The main geographic features of the studied reach can be divided into two parts: Lubelski (from rivers San to Wieprz) and Mazowiecki (from Wieprz to Narew), (Piskozub 1982; Laskowski 1986; Kondracki 2002). This division originates from the substantial increase of the catchment area and the differences in geology and morphology of both parts of the reach.

Evolution of the valley of the middle part of River Vistula was very complex, which is stressed by Łajczak et al. (2006). High level of complexity and poly-genetic character of the terrain shape, geological structure of landscape and riverbed may be noted in the middle section of the River Vistula, which is a factor which conditions the variety of terrain covering, hypsometry and longitudinal drop, but is also a deciding factor on roughness and absorbability of the riverbed and terrace sediments. Terrain shape and morphology of the Middle River Vistula drainage area consist of three big structural units: Małopolska Upland (Wyżyna Małopolska), Lublelska Upland (Wyżyna Lubelska), and Mazowiecka Lowland (Nizina Mazowiecka) being part of Środkowopolskie Lowlands (Niziny Środkowopolskie). These are characterized by a variety of terrain shapes, with predominance of the convex forms (uplands, hills, elevated areas), while the second group is featuring a larger amount of convex forms (valleys, lowlands, lower sections of the terrain). The main river valleys, including middle section of River Vistula and tributaries such as: Pilica, Wieprz, Kurówka and Okrzejka, have been shaped in the erosion and denudation processes during the Masovian interglacial period (Kosmowska-Suffczyńska 2009).

3 Methodology

In this study, the land cover and land cover change were derived from the CO-oRdination of INformation on Environment (CORINE) Land Cover databases. The CORINE Land Cover inventory program represents a comprehensive approach to provide ongoing inventory for European Union. The first European and Polish CORINE Land Cover map was created for the reference year 1990, then subsequently for 2000, 2006 and recently for 2012. The long heritage of CLC inventories allows the long-term analysis of land cover dynamics at national and European scale to be performed. In this study, the national CORINE Land Cover databases have been used to derive information on the land cover dynamics in the Middle River Vistula catchment for the four years: 1990, 2000, 2006 and 2012.

The CORINE Land Cover (CLC) is a vector dataset with a scale of 1:100,000, a minimum cartographic unit of 25 ha and a geometric accuracy better than 100 m. The CLC nomenclature is a 3 level hierarchical classification system and has 5 classes at the 1st level, 15 classes at the 2nd level and 44 classes at the 3rd, most detailed level. A minimum cartographic unit for CORINE Land Cover Change map is set to 5 ha (Büttner and Kosztra 2011; Büttner et al. 2014). Mapping of CLC-Change 2006–2012 is carried out by applying the 'change mapping first' approach. The changes were interpreted directly, by comparison of reference images. A set of multitemporal images delivered by Landsat TM satellite was used to map land cover in 1990 and 2000, then images delivered by SPOT4 and IRS satellites were used to map land cover changes in 2006 and finally, for CLC2012 inventory images from RapidEye and IRS satellites were applied (Ciołkosz and Bielecka 2005; Bielecka and Ciołkosz 2009; Chościło and Tomaszewska 2015).

In total, 27 CLC classes of level 3 were distinguished within the study area. In order to understand how the land cover change affects hydrological processes the CLC legend (3rd level) was aggregated into six land cover groups: urban, agricultural, meadow, forests, open space and surface water. Aggregation groups are similar in terms of contribution in the development of hydrological processes. Analysis of the land cover for the Middle River Vistula catchment was performed for 1990, 2000, 2006 and 2012; analysis of the land cover change was conducted for the entire period 1990–2012 and separately for the periods 1990–2000, 2000–2006, 2006–2012. The analysis was conducted for (i) the whole study area (the Middle River Vistula catchment), (ii) 10 sub-catchments and (iii) 11 adjacent areas (Fig. 2). The catchment boundaries were derived using hydrological divide of Poland (MPHP version 2010). All datasets were transformed to the national projection PL_EUREF89/1992 (EPSG:2180). In the next step, CLC datasets were clipped to selected catchments and adjacent areas.

Figure 3 presents the example map of state land cover in last time (2012) for the entire catchment area. The applied legend, prepared according to the 3th CLC level (Table 1), shows heterogeneity of the landscape. Thus, the aggregation of land cover classes proposed in this study was essential.

The first type of land cover according to the aggregation is "urban areas". This type includes forms which result from heavy anthropogenic activity—urbanization and industrialization (Table 1—CLC classes: 111, 112, 121, 122, 123, 124, 131, 132, 133, 141 and 142). The second group is "agricultural areas" containing the following CLC classes: 211, 222, 242 and 243. The third group "forest" consists of "forests and bush vegetation" (311, 312, 313 and 324). The next group called "meadows" includes green utility areas—dry and wetland (CLC classes 231, 321, 411, and 412) and the remaining two groups are: "open space" (non-vegetated areas), which includes sandbanks, and burnt areas (CLC classes 331, 334) and "surface waters" (CLC classes 511, 512).

4 Spatial and Temporal Distribution of Land Cover Change

The analyses of land cover change was carried out using the CORINE land cover databases for the years 1990, 2000, 2006 and 2012 for: (i) the entire Middle River Vistula catchment, (ii) 10 tributaries and (iii) 11 adjacent areas. The long heritage

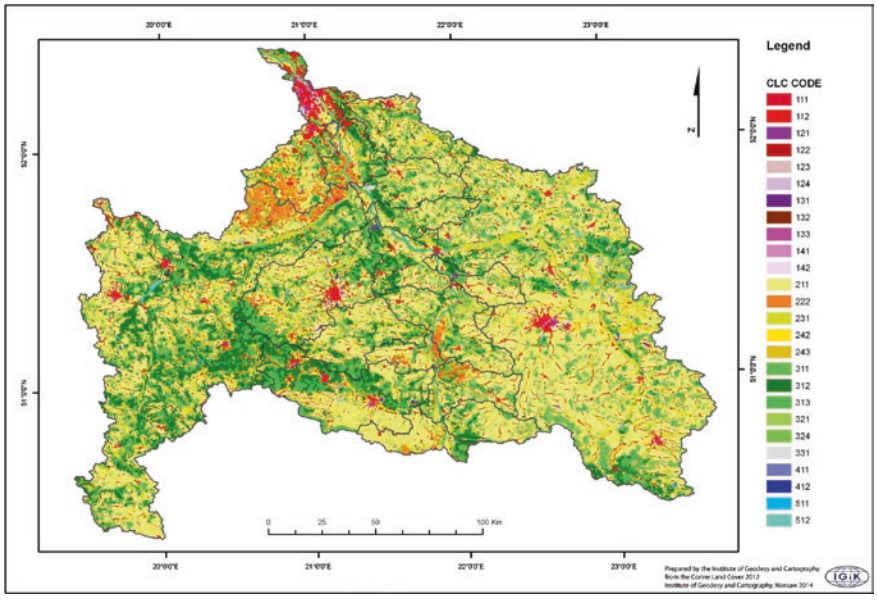

Fig. 3 Map of the CORINE Land Cover 2012 (level 3) for the Middle River Vistula catchment. The legend is presented in Table 1

Table 1 CORINE Land Cover legend (3 levels and aggregation used for the purpose of this study)

Level 1	Level 2	Level 3	Aggregation used in the study
Artificial surfaces	11 urban fabric	111 continuous urban fabric 112 discontinuous urban fabric	Urban areas (111 ÷ 142)
	12 industrial, commercial and transport units	121 industrial or commercial units 122 road and rail networks and associated lands 123 port areas 124 airports	
	13 mine, dump and construction sites	131 mineral extraction sites 132 dump sites 133 construction sites	
	14 artificial, non-agricultural vegetated areas	141 green urban areas 142 sport and leisure facilities	
Agricultural areas	21 arable land	211 non-irrigated arable land	Agricultural areas (211 ÷ 222)
	22 permanent crops	222 fruit trees and berry plantations	
	23 pastures	231 pastures	Meadow areas (231)
	24 heterogeneous agricultural areas	242 complex cultivation patterns 243 land principally occupied by agriculture, with significant areas of natural vegetation	Agricultural areas (242 ÷ 243)
Forest and semi natural areas	31 forests	311 broad-leaved forests 312 coniferous forests 313 mixed forests	Forest areas (311 ÷ 313)
	32 scrub and/or herbaceous vegetation associations	321 natural grasslands	Meadow areas (321)
		324 transitional woodland-shrub	Forest areas (324)
	33 open spaces with little or no vegetation	331 dunes, airy sands 334 burn areas	Open spaces (321 ÷ 334)
Wetlands	41 inland wetlands	411 marshes 412 peat bogs	Meadow areas (411 ÷ 412)
Water bodies	51 inland waters	511 water courses (rivers, lakes) 512 water bodies (reservoirs, ponds)	Surface water (511 ÷ 512)

of CLC inventories allows the long-term analysis of land cover dynamics within the study area to be performed.

4.1 Land Cover Change in the Entire Middle River Vistula Catchment

The results of the assessment for the entire Middle River Vistula catchment are presented in Fig. 4. It is seen that the catchment is predominantly agricultural. At the beginning of analysed period, in 1990, the agricultural areas covered around 20,153 km^2 (57.8 % of the entire catchment). In the following years, a small decrease of agriculture land was observed, and reached the lowest value of 18,859 km^2 in the year 2012.

The second type of land cover—"forests", in 1990 constituted around 8958 km^2 (25.7 % of the area); in the subsequent periods a slow rise is noted, and it exceeded 29 % of the entire catchment in 2012 (10,125 km^2). "Meadows" cover around 9.5 % of the entire catchment and show almost no changes during the investigation period. The urbanized areas covered 2036 km^2 (5.8 %) in 1990 and increased to 2332 km^2 in 2012 (6.7 % of the entire study area). These results confirm the earlier findings by Bański (2003) and those given in the Agricultural Atlas of Poland (2010), regarding the intensity and direction of land cover changes.

There is observed an increase of "urban area" during the entire investigation period. In 1990, this type of land cover occupied more than 2036 km^2 of the Middle River Vistula catchment and 2332 km^2 in the year 2012. Area of the "surface waters" consisting of large rivers (more than 100 m wide) and the

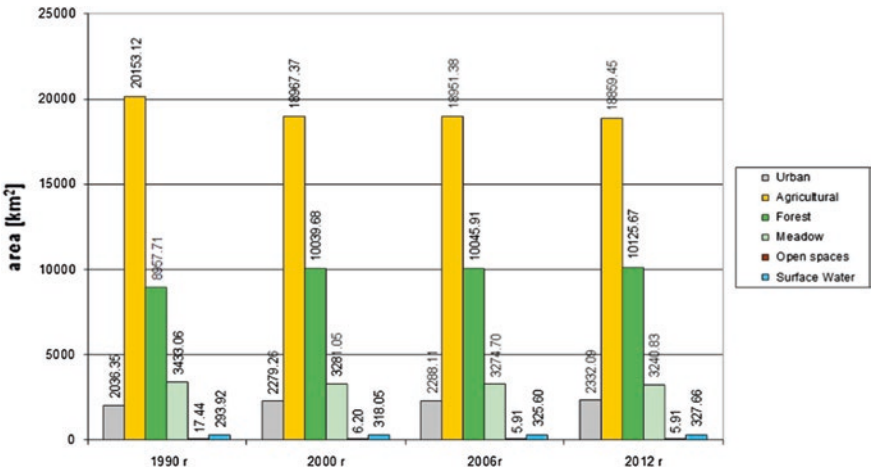

Fig. 4 Analysis of land cover in the entire Middle River Vistula catchment in years 1990, 2000, 2006 and 2012

reservoirs was equal to 293.9 km^2 in 1990 and was slowly increasing up to 327.7 km^2 in 2012. It should be noted that due to the minimum size of the CLC polygons (25 ha) only the larger storage reservoirs are taken into account, including: Sulejowskie Lake, Brody Iłżeckie Lake, system of reservoirs along Rivers Radomka and Świder and ponds on River Jeziorka (in Żabieniec). The "open space" class coved the small area of 17 km^2 in 1990 and around 6 km^2 after 2000. The decrease in the 'open space' is due to the post-fire vegetation regrowth that occurred over the burned area in 2000. The rest of this class is made of sandbanks.

In summary, the estimated changes of land cover over the entire Middle River Vistula catchment indicate a decrease (6.4 %) of the agricultural areas, complemented by an insignificant increase of the forested areas (3.3 %) in comparison with the 1990 status. The spatial distribution of land cover changes between the year 1990 and 2012 is presented in Fig. 5.

4.2 Land Cover Change in Tributaries and Adjacent Areas

The preceding sections presented an analysis of land cover change over the entire Middle River Vistula catchment treated as a single area. In the present section, particular emphasis is placed on an assessment of the land cover change in the 10

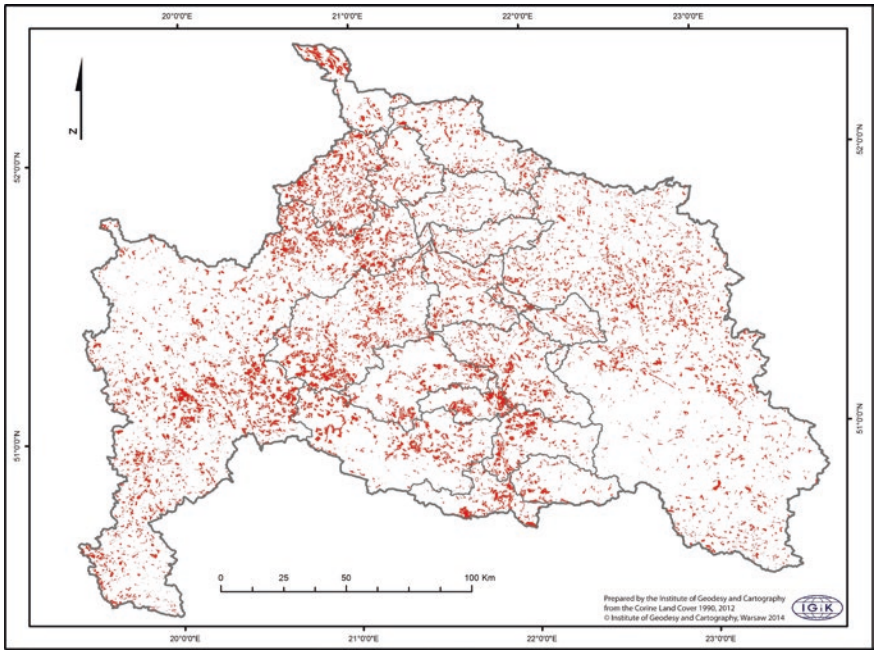

Fig. 5 Map of land cover change between 1990 and 2012 in the Middle River Vistula catchment

tributaries and 11 adjacent areas presented in Fig. 2. Similarly to the previous sections, the analysis was conducted with the aggregated level where six types of land cover are distinguished: agricultural, forests, meadows, open spaces and surface water. The results of the assessment of the land cover in the analysed areas for the year 2012 are presented in Fig. 6. In the case of tributaries, the largest percentage of the "agricultural areas" is located in the following catchments: Sanna (65.5 %), Kurówka (65.3 %) and Wieprz (61.8 %); whereas the River Pilica (47 %) and Kamienna (44.8 %) catchments are characterised by the lowest agricultural land. Right tributaries show a significant predominance of the agricultural areas comparing to the left ones. The percentage of the areas covered by forest is shown in the second row from the top in Fig. 6. A significant asymmetry between the left and right tributaries is observed.

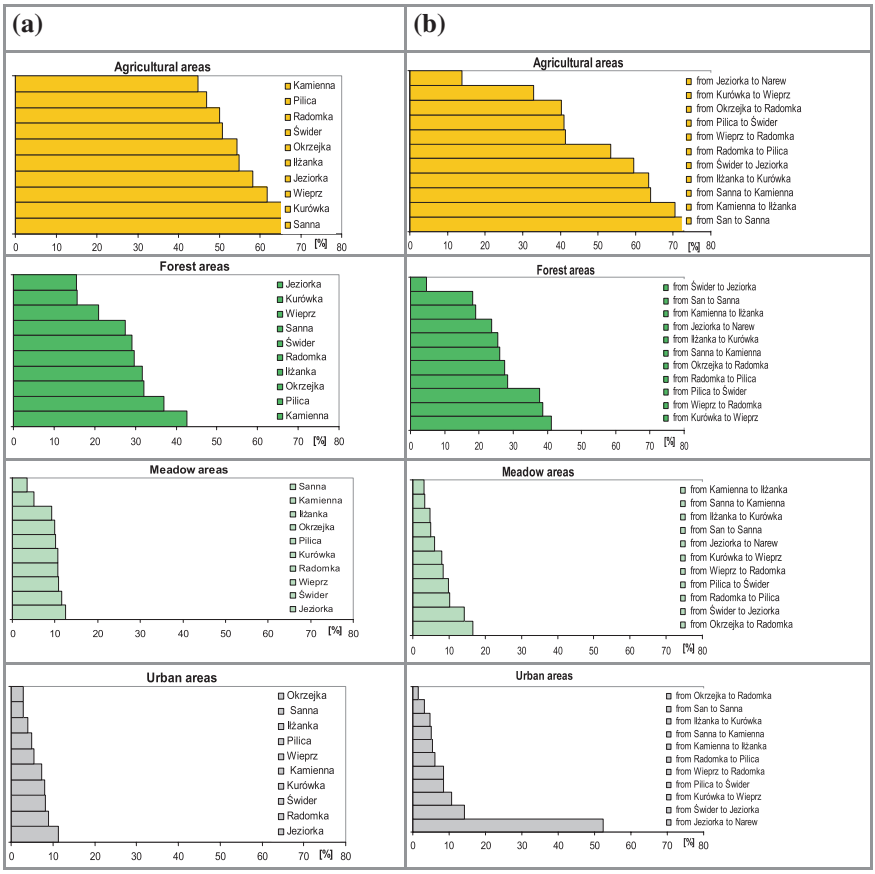

Fig. 6 A comparison of percentage of the four land cover types (acc. to aggregated groups) in the 10 tributaries and 11 adjacent areas for the year 2012. The results were sorted in the ascending order. **a** Tributaries. **b** Rivers adjacents

The analysis of area covered by 'meadows' is presented in the third row from the top in Fig. 6. The largest percentage of the meadow areas is located in the northern part of the study area: Jeziorka (12.4 %), Świder (11.5 %), Wieprz (10.9 %) and Radomka (10.7 %) catchments.

The lowest raw of Fig. 6 presents the percentage of 'urban areas' in the 10 tributaries and 11 adjacent areas. It can be noted that urban development and industrialization occur mostly in the northernmost part of the study area, with the largest contribution of the following rivers: Jeziorka (11.3 %), Radomka (9.0 %) and Świder (8.3 %) catchments. The tributaries located in the south part of the study area are far less urbanized: Iłżanka (4.0 %) or Sanna and Okrzejka (2.9 %).

The results of dominant types of land cover between tributaries and adjacent areas indicate different trends. The classification of the adjacent areas according of the percentage of 'agricultural areas' indicated that this type of land cover is dominant in the lower part of study area (over 70 % of the area from the River San to Sanna). The lowest percentage of 'agricultural area' is observed in the northern part of the catchments (from 13.8 % up to 30 %). The percentage of areas classified as 'forest' differs between analysed adjacent areas. The highest values were estimated for those located on the left part of the catchment. The largest percentage of 'forest' starts from the River Pilica (>40 %), decreases towards north-east, and reaches the minimum in the River Świder neighbouring area (4.7 %). There are differences in the percentage of meadows between analysed adjacent areas.

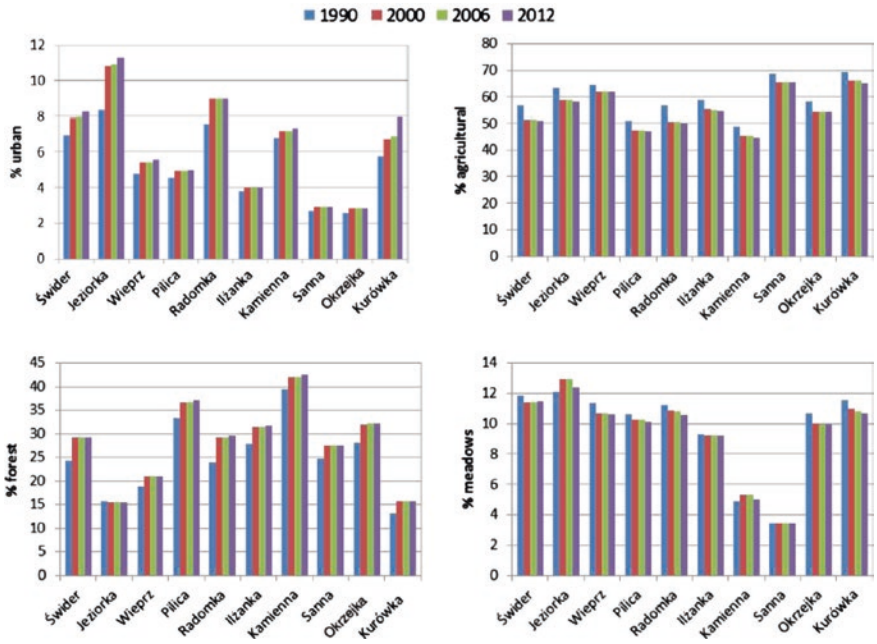

Fig. 7 A comparison of the four dominant land cover types (urban, agricultural, forested and meadows) estimated for years 1990, 2000, 2006 and 2012

The highest percentage (more than 10 %) is observed in the areas from the River Świder to Jeziorka and from the River Okrzejka to Radomka. Urbanized and industrialized areas cover more than 50 % of the river-adjacent areas between Jeziorka and Narew. Southern and middle river neighbouring areas are urbanized to the smallest extent (less than 10 %).

Figure 7 presents a comparison of four dominant land cover types (urban, agricultural, forest and meadows) estimated for the years 1990, 2000, 2006 and 2012. The following three types of trends could be recognized: (a) increase in urban and forested areas, (b) decrease in agricultural areas and meadows, and (c) lack of changes in meadows for the River Sanna catchment. Generally, there is an agreement in the estimated tendency of changes between catchments with few exceptions, however there are differences in the intensity of changes.

5 Summary and Conclusions

In this chapter the land cover and land cover change in the Middle River Vistula catchment were outlined. The description of the catchment was followed by the presentation of land cover changes over the period 1990–2012 derived from the CORINE Land Cover databases. The land cover data were analyzed according to the proposed aggregated level resulting in five classes: urban, agricultural, meadow, forest, open spaces and surface water. The analyses of land cover change were conducted for the entire Middle River Vistula catchment and for the 10 tributaries and 11 adjacent areas. Three different trends in land cover changes were recognized. Namely, it was found that the percentage of urban and forest covered area increased within the analyzed period. The percentage of agricultural areas and meadows decreased whilst no change was observed in some sub-catchments. In general, the trends listed above occur in most of catchments, but the intensity of changes differs between the analyzed regions.

The performed analysis presents changes integrated over the catchment area, but do not provide the specific location of the changes. The available information on the land cover changes can be applied in the study of the human impact on runoff in the catchment. An example of that kind of research was given by Romanowicz and Osuch (2011) for the River Narew. However, the preliminary research shows that the impact of the land cover on the flow in the main channel of the River Vistula might be difficult to recognize due to the disproportion between the amount of water from tributaries and the main channel of the River Vistula at the studied middle part of the river.

Acknowledgments This work was supported by the project “Stochastic flood forecasting system (The Middle River Vistula reach from Zawichost to Warsaw)” carried by the Institute of Geophysics, Polish Academy of Sciences, on the order of the National Science Centre (contract No. 2011/01/B/ST10/06866).

The authors thank the Chief Inspectorate of Environmental Protection (GIOŚ) for providing the national CLC2006 and CLC2012 databases. The national CLC2012 project was conducted in the framework of the Copernicus GIO Land Monitoring (with funding from the European Union) by the Institute of Geodesy and Cartography in Warsaw.
The Map of Hydrographical Divide of Poland (MPHP) was prepared by the Institute of Meteorology and Water Management and was funded by the National Fund for Environmental Protection and Water Management.

References

Bański J (2003) Selected aspects of present-day changes in Polish rural space. Geographia Polonica 76(1):73–76

Bański J (ed) (2010) Atlas rolnictwa Polski. Wyd. IGiPZ PAN, Warszawa

Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology. Hydrol Sci Bull 24:43–69

Bielecka E, Ciołkosz A (2009) Land cover database in Poland CLC-2006 (in Polish). Polski Przegląd Kartograficzny 41(3):227–236

Büttner G, Kosztra B (2011) Manual of CORINE Land Cover changes. Final Draft, Copenhagen (EEA)

Büttner G, Soukup T, Kosztra B (2014) CLC2012 addendum to CLC2006 technical guidelines. Final Draft, Copenhagen (EEA)

Ciołkosz A, Bielecka E (2005) Land Cover in Poland—CORINE Land Cover data base (in Polish). Biblioteka Monitoringu Środowiska, Inspektorat Ochrony Środowiska, Warszawa, pp 1–76

Crave A, Gascuel-Odoux C (1997) The influence of topography on time and space distribution of soil surface water content. Hydrol Process 11(2):203–210

Dąbrowska-Zielińska K, Ciołkosz A, Malińska A, Bartold M (2011) Monitoring of agricultural drought In Poland using data derived from environmental satellite images. Geoinf Issues 3(3):87

Dąbrowska-Zielińska K, Budzyńska M, Kowalik W, Małek I, Gatkowska M, Bartold M, Turlej K (2012) Biophysical parameters assessed from mircrowave and optical data. Int J Electron Telecommun 58(2):99–104

Chościło A, Tomaszewska M (2015) CORINE Land Cover 2012-4th CLC inventory completed in Poland, Geoinf Issues (in press)

Jarsjö J et al (2012) Hydrological responses to climate change conditioned by historic alternations of land-use and water-use. Hydrol Earth Syst Sci 16:1335–1347

Kondracki J (2002) Geografia regionalna Polski. PWN, Warszawa

Kosmowska-Suffczyńska D (2009) Geneza i rozwój rzeźby w czwartorzędzie. In: Richling A, Ostaszewska K (eds) Geografia fizyczna Polski. WN PWN, Warszawa, pp 32–67

Łajczak A, Plit J, Soja R, Starkel L, Warowna J (2006) Changes of the vistula river channel and floodplain in the last 200 years. Geogr Polon 79:65–87

Laskowski K (1986) Przebieg erozji w dolinie Wisły na odcinku Puławy-Warszawa, Biul. Geol. Uniwersytetu Warszawskiego, No. 30

Mapa Podziału Hydrograficznego Polski (MPHP) (2010) Ośrodek Zasobów Wodnych IMGW-PIB. Warszawa

Montanari AG, Young G, Savenije HHG, Hughes D, Wagener T, Ren LL, Koutsoyiannis D, Cudennec C, Toth E, Grimaldi S, Blöschl G, Sivapalan M, Beven K, Gupta H, Hipsey M, Schaefli B, Arheimer B, Boegh E, Schymanski SJ, Di Baldassarre G, Yu B, Hubert P, Huang Y, Schumann A, Post D, Srinivasan V, Harman C, Thompson S, Rogger M, Viglione A, McMillan H, Characklis G, Pang Z, Belyaev V (2013) Panta Rhei – everything flows: change in hydrology and society – the IAHS scientific decade 2013–2022. Hydrol Sci J 58(6):1256–1275

Piskozub A (1982) Wisła. Monografia rzeki, Wydawn. Komunikacji i Łączności. Warszawa
Romanowicz RJ, Osuch M (2011) Assessment of land use and water management induced changes in flow regime of the Upper Narew. Phys Chem Earth. doi:10.1016/j.pce.2011.04.012
Schulze RE (2000) Modelling hydrological responses to land use and climate change: a Southern African perspective. Ambio 29(1):12–22
Sivapalan M, Savenije HHC, Blöschl G (2012) Sociohydrology a new science of people and water. Hydrol Process 26:1270–1276
Vachaud G, Passerat de Silans A, Balabanis P, Vauclin M (1985) Temporal stability of spatially measured soil water probability density function. Soil Sci Soc Am J 49:822–828

Influence of Land-Use and Water Management Practices on Water Levels in the Middle River Vistula

E. Karamuz, R.J. Romanowicz and M.J. Booij

Abstract The present paper addresses the influence of land use changes and water management on flow along the Middle River Vistula reach. We adopt a methodology tested at the catchment scale, which consists of an optimisation of a rainfall-runoff model using a moving time horizon and an analysis of the variability of model parameters. In the present application, it consists of an analysis of the changes of residence times of a linear Stochastic Transfer Function (STF) type model optimised using a moving ten-year window. The chosen river reach (between Zawichost and Gusin) has a reported history of land use changes over the last 50 years. The nature of the changes is complex and shows different trends for different plant communities and sections of the valley. The first step of the procedure is to define the river reaches that have recorded information on land use change. The second step consists of analyses of static relationships between water level observations at neighbouring gauging stations along the river. We analyse the changes in observed water level quantiles. The third task is to perform the moving window optimisation of the STF dynamic model for selected river reaches. The final step is to analyse the results and relate the model parameter changes to historical land use changes. The results of the static and dynamic analyses are consistent for the first two reaches (Zawichost-Annopol-Puławy) and can be explained by

E. Karamuz (✉) · R.J. Romanowicz
Institute of Geophysics Polish Academy of Sciences, Ks. Janusza 64,
01-452 Warsaw, Poland
e-mail: emilia_karamuz@igf.edu.pl

R.J. Romanowicz
e-mail: romanowicz@igf.edu.pl

M.J. Booij
Department of Water Engineering and Management, University of Twente,
Enschede, The Netherlands
e-mail: M.J.Booij@utwente.nl

R.J. Romanowicz and M. Osuch (eds.), *Stochastic Flood Forecasting System*,
GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-319-18854-6_2

land-use and water management influences. However, the reaches from Puławy-Dęblin-Gusin show inconsistent behaviour, pointing to a need for further study.

1 Introduction

The aim of the study is an analysis of the changes in flow regime (1975–2003) in the river under different flow conditions and finding possible causes. The time period was selected from the availability of observations of water levels and land-use data. The analysis was conducted using summer (V–X) and winter (XI–IV) months separately, to investigate different origins of the changes. The flow conditions can vary due to changes in the river reach, including changes in water management, channel geomorphology and land use modifications, as well as seasonal changes related to climatic conditions. We are particularly interested in linking land cover changes to changes in flow regime at the scale of a river reach. Vegetation in the channel modifies its capacity and also affects the conveyance, which is very important for both low and high flow events. We expect that due to climate change the frequency and magnitude of extreme flows can change significantly in the near future and for this reason it is very important to examine the long-term response of flow variation to land-use and water management changes within the river bed. Vegetation in direct contact with flowing water affects the flow velocity and water levels by a direct reduction of the available space in the channel (reduction of capacity in a particular section of the river), by using the energy of the flow and by blocking the flow and reducing velocity.

Various approaches have been proposed to investigate how vegetation affects open-channel flow (e.g. Tsujimoto et al. 1993; Darby 1999; Lopez and Garcia 2001; Warmink et al. 2013). Most researchers study the hydraulic properties of the channel and use empirical formulas (Manning, Darcy-Weisbach, Chezy) to assess the effects of changes of flow resistance on open–channel flow.

The influence of the type of channel and flood-plain vegetation on flow has been studied intensively at the micro-scale (cm to meters), in laboratories and in open channels. Many researchers (Kouwen and Unny 1973; Li and Shen 1973; Petryk and Bosmajian 1975; Yen 1991; Bathrust 1993) studied how submerged and non-submerged vegetation influence flow conditions in the channel. Järvelä (2002) investigated how the type, density and placement of vegetation, flow depth and velocity influence friction losses. He studied the influence of plants on flow resistance under non-submerged and submerged conditions in a total of 350 test runs. The results indicated the existence of large fluctuations in the friction factor with respect to depth of flow, velocity, Reynolds number, and vegetative density. He demonstrated that the friction factor depended mostly on the relative roughness in the case of grasses; on the flow velocity in the case of willows and sedges/grasses combined; and on the flow depth in the case of leafless willows on the bare bottom soil. He noticed that the leaves on willows seemed to double or even triple the friction factor compared to the leafless trees. This important information indicates the existence of seasonal differences in channel roughness.

The water level is the primarily observed variable in hydrology and hydrologists still regard it as being a secondary variable, after flow. The reason might lie in the local character of that variable, i.e. its dependence on the geometry of a river cross-section where a gauge is situated. Flows have a more integrated character, as they are derived taking into account the velocity of the flow. In this study we focus on water levels observed along the river reach in order to determine what kind of information we can derive when analysing their dynamic and static changes. The Middle River Vistula reach will be used as a case study.

The study area and a summary of historical changes within it are given in Sect. 2. A statistical analysis of changes of water level quantiles is presented in Sect. 3. Section 4 gives the results of a dynamic analysis of water level changes. The results of the analyses are compared and summarised in Sect. 5 and in Sect. 6 conclusions are drawn.

2 The Middle River Vistula Study Area

As a case study we chose the middle part of the Vistula River, between Zawichost and Gusin. In order to trace changes in the water level regime we divided the reach into four parts according to the location of gauging stations (Zawichost-Annopol, Annopol-Puławy, Puławy-Dęblin, Dęblin-Gusin) for which water level measurements were available (Fig. 1). The whole studied reach is 173.9 km long and it is located within two meso-regions: the Vistula River Gorge and the Middle Vistula River Valley (Kondracki 1994). This results in a different channel character for the sections Zawichost-Annopol and Annopol-Puławy compared to the other two sections, Puławy-Dęblin and Dęblin-Gusin. The river part in the gorge section is narrow, deep and straight. After leaving the gorge, the main channel undergoes a transformation. The valley begins to widen downstream of Puławy and the channel is shallower and more sinuous.

The reach between Puławy and Gusin (86.6 km) has a floodplain valley 13–16 km wide and it separates the Radom Plane (Nizina Radomska), and its lower part—Kozienice Valley (Dolina Kozienicka)—from the Siedlce Upland (Wysoczyzna Sieldecka). The valley bottom is about 200–1200 m wide and has a meandering character, with vegetated islands and sand dunes (oxbows). The slope of the River Vistula in that section is similar to that within the gorge and varies from 0.23 to 0.29 ‰ (Maruszczak 1972; Starkel 2001).

We are particularly interested in changes of vegetation cover in the river valley that periodically comes into direct contact with the flowing water; therefore, the study of land cover changes which affected the flow conditions was limited to the area between the embankments. We use Corine Land Cover data obtained from the EEA web page for the years 1990, 2000 and 2006 to analyse changes in vegetation cover along the reach. A summary of changes between the years 1990 and 2006 for each section is presented in Table 1.

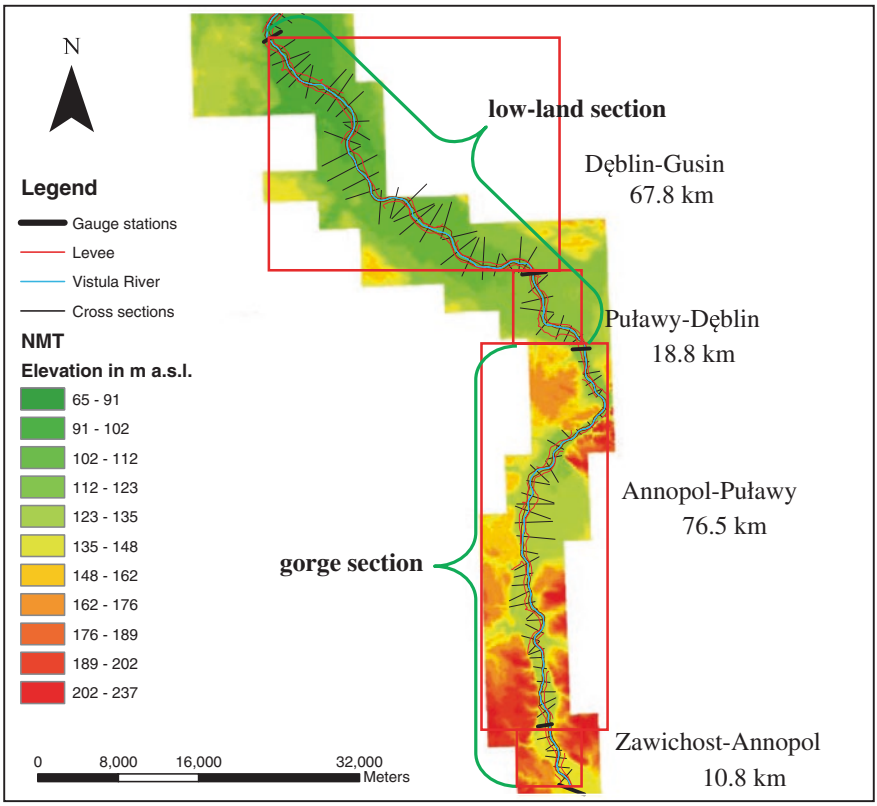

Fig. 1 Study area showing part of the Middle Vistula valley between Zawichost and Gusin

Table 1 Relative changes of different land use types in % for the studied reaches between 1990 and 2006

	Zawichost-Annopol	Annopol-Puławy	Puławy-Dęblin	Dęblin-Gusin
Arable land	−3.6	−0.3	0.0	+0.2
Pastures and grasslands	+4.6	+1.2	0.0	+0.2
Agriculture/natural vegetation	−52.4	−6.3	+2.2	−1.3
Deciduous forest	+50.8	+4.0	−2.1	0.0
Beaches/dunes	0.0	−1.4	0.0	−0.8
Water courses	0.0	+1.0	−0.1	+0.6

The area of the river valley between dikes for the whole studied section (Zawichost-Annopol-Puławy-Dęblin-Gusin) is covered mainly by willow and different kinds of poplar alluvial forests with a significant part also used as grasslands—meadows and pastures. A small area is cultivated; quite often these are orchards and gardens.

Table 2 Number of groynes per km (the total number of groynes is given in parentheses) for the Annopol-Puławy, Puławy-Deblin and Dęblin-Gusin reaches

	Annpol-Puławy	Puławy-Dęblin	Dęblin-Gusin
1960s	0.7 (50)	0.1 (2)	0.3 (19)
1970s	1.3 (97)	0.6 (11)	0.6 (41)
1980s	0.7 (50)	0.1 (2)	0.1 (8)
1990s	0.1 (4)	0.3 (5)	0.2 (12)

The largest changes in land use between river embankments occurred mainly within the Zawichost-Annopol reach, where over 50 % of natural vegetation, agricultural land, meadows and pastures was replaced by deciduous forest. Within the Annopol-Puławy reach, that pattern was repeated to a lesser extent (only 6 %). Also within that reach a small decrease of beach/sand area was noted and a small increase in water courses, which suggests morphometric changes in this section. The other two reaches showed very small changes.

Apart from land-use, changes in hydro-technical constructions along the middle Vistula reach could also have affected flow. The selected part of the river reach is mostly regulated, but hydro-engineering constructions are unevenly located (Kowalski 1997). There are embankments close to the whole stretch studied. They were built mostly after the Second World War, but local dikes were already in place in the 19th century (Łajczak et al. 2006). Channel transformations resulted from accelerated outflow of water and sediment load through the channelized reaches of the Upper Vistula and tributaries (Warowna 2003). Channelization works were performed partly in the 1920s and later between 1960 and 1990, resulting in the narrowing of the main channel (by a system of dams and groynes) to 120–300 m and the cutting-off of side branches by dikes.

Table 2 shows the history of groyne construction along three reaches: Annopol-Puławy, Puławy-Dęblin and Dęblin-Gusin. The history of changes for the Zawichost-Annopol reach is not reported here due to an absence of information. Most affected was the reach between Annopol and Puławy before 1990. The groynes mainly affect the mean and low flows, acting as submerged weirs. However, the channelization might affect the whole range of flows. Low flows are supposed to be affected mostly by the construction of groynes in that section which has resulted in an increase of water levels during low flow periods.

3 Statistical Analysis of Water Levels at Gauging Stations Along the River Reach

The statistical analysis was based on changes in the quantile values of observed water levels for each section of the river. We analysed the differences in quantiles of neighbouring gauging stations for a moving time window of ten years, as shown in Fig. 2. For a comparability of results we used standardized values of quantiles.

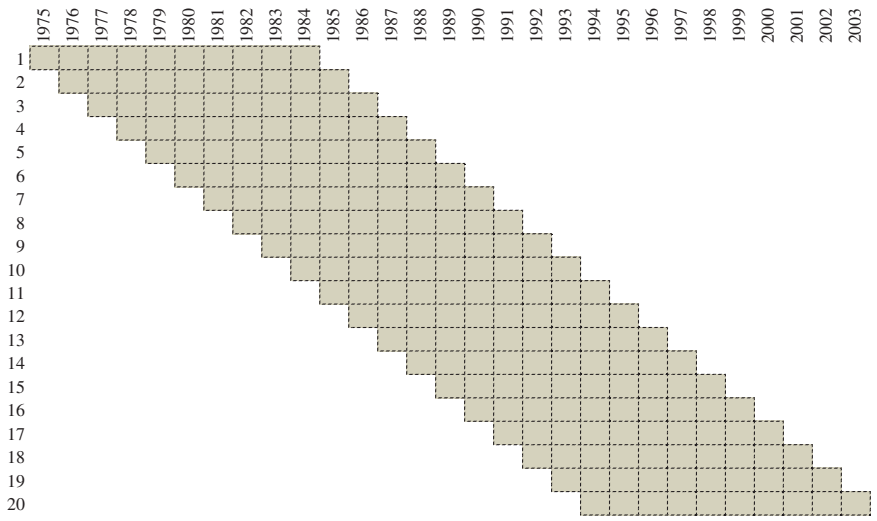

Fig. 2 Illustration of a smoothing ten-year window applied during the static and dynamic analysis of water levels

This part of the work was meant as a preliminary assessment of the changes for the selected sections and was a motivation for further research. Changes in quantile values can be due to various changes within the river reach. Following the study of Ozga-Zielinski et al. (2010), an increase of roughness of the channel due to vegetation affects flow and can cause a rise of water levels.

The approach is based on a comparison of quantiles with a given probability of exceedance p, of daily observed water levels h_p, at the selected i'th gauging station within a 10-year moving window (Fig. 2) with the corresponding quantiles from the $(i-1)$ gauging station upstream. The moving window was introduced in order to smooth the changes (Osuch et al. 2015).

In order to make the comparison independent of the time window, the p'th-quantile at each gauging station is standardised as follows:

$$h'_p = \frac{h_p - h_{\min}}{h_{\max} - h_{\min}} \tag{1}$$

where h'_p is the standardized quantile value for a selected station; h_p is the p'th quantile for the selected 10 year time window; $h_{\min}$ is the minimum h_p quantile value within all time-windows and $h_{\max}$ is the maximum h_p value. The difference in quantiles $\Delta h'_{p,i}$ at the i'th cross-section is defined as follows:

$$\Delta h'_{p,i} = h'_{p,i} - h'_{p,i-1} \tag{2}$$

where $h'_{p,i}$ denotes a standardised p'th quantile at the i'th cross-section and $h'_{p,i-1}$ denotes a standardised quantile at the cross-section upstream.

This fairly simple approach allows us to make inferences about the changes in conveyance in the river reach associated with environmental changes in the immediate neighbourhood of the channel. All quantiles along the river reach are normalised within all 10-year moving windows. The quantile confidence limits are estimated using nonparametric methods developed by Hutson (1999), Kaplan (2011), and Kaplan and Goldman (2012).

To distinguish the possible impacts of various factors, our analysis was based on low and high quantiles for different seasons. The results for the 0.25 quantile are shown in Fig. 3. The two columns describe differences in quantiles between the downstream and upstream gauging station (Eq. 2) for the summer (left panel) and winter (right panel) periods. The x-axis shows the numbers that correspond to the time-windows defined in Fig. 2. Number 1 corresponds to the time-window 1975–1984, number 2 corresponds to the time-window 1976–1985, and so on.

The largest differences in water level quantiles shown in Fig. 3 occur for the Puławy-Dęblin reach. They are positive both during the summer and winter periods, but are more pronounced in the winter, indicating a possible increase in the reach roughness. This negative trend of conveyance is noticeable from the mid 1970's which is consistent with possible geomorphological changes of the channel related to works on river regulation. Positive changes of 0.25 quantile in the winter period can also be explained by ice–jam formation in that part of the river.

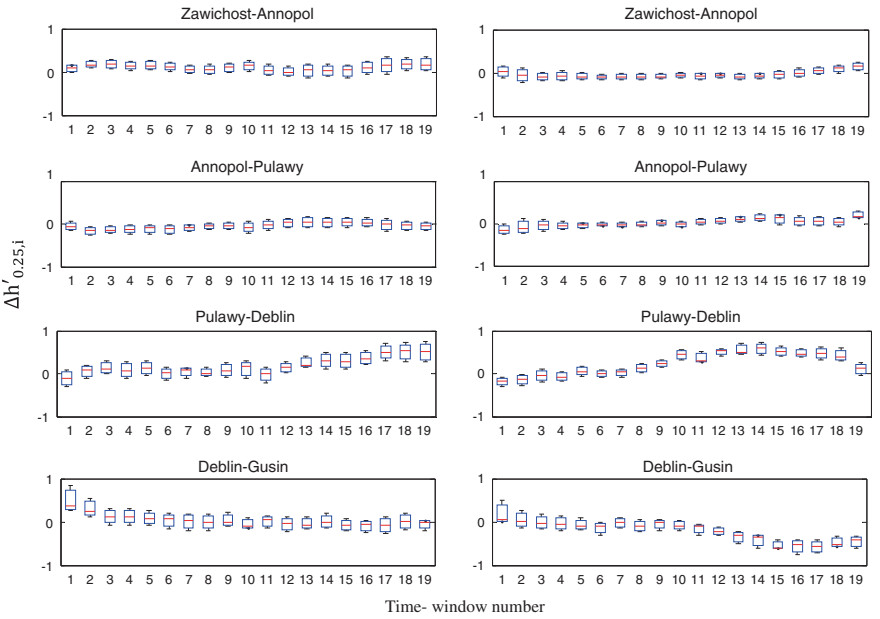

Fig. 3 Differences in water level quantiles between the downstream and upstream gauging stations $\Delta h'_{p,i}$, $i = 2, \ldots, 5$ (Eq. 2) for a probability of exceedance $p = 0.25$; *left column* describes the summer season, *right column* describes the winter season, 0.95 confidence bounds are shown by box-plots, x-axis denotes time-window numbers as explained in Fig. 2

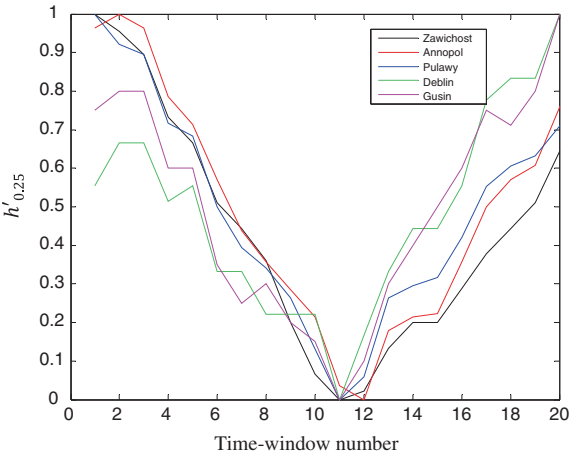

Fig. 4 Standardized 0.25 quantile values for five gauging stations: Zawichost, Annopol, Puławy, Dęblin and Gusin, x-axis denotes time-window numbers as explained in Fig. 2

On the other hand, changes in water level of 0.25 quantiles for the Dęblin-Gusin reach indicate a decrease of roughness coefficients in that part of the reach. These changes might be due to an increase of water levels in the upper cross-section (Dęblin gauging station), without any changes in the downstream water level quantile. Thus we should look also at the time variability of the water level quantiles for each station separately to explain those differences. The time variability of 0.25 quantiles for each gauging station is presented in Fig. 4.

Figure 4 illustrates the variability water level for each gauging station within the period 1975–2003. All quantiles decrease in the period 1975–1985 and increase in the period 1985–2003. Their mutual differences also change, with Gusin and Dęblin changing from the lowest in the first period to the highest in the second period. The latter indicates the existence of a different source of influence than a climate-related source.

The following figures (Figs. 5 and 6) present differences in quantiles with probabilities of exceedance of 0.5 and 0.75.

The differences in water level quantiles with a probability of exceedance of 0.5 for the years 1975–2003 show a positive increase for the Annopol-Puławy reach, starting from 1990, contrasted with a decrease of water levels for the upstream Zawichost-Annopol reach. These changes are more pronounced in the summer period, suggesting a dependence on vegetation. On the other hand, the Puławy-Dęblin and Dęblin-Gusin reaches show stronger changes during the winter periods. In all cases, the summer periods have a larger variability of flow, demonstrated by larger confidence bands of the differences of water level quantiles during those periods.

The differences for the 0.75 water level quantiles show a smaller variability within the Zawichost-Annopol and Annopol-Puławy reaches compared to the other two reaches. In particular, the Dęblin-Gusin reach has a large variability of differences in water level quantiles both in the winter and summer periods.

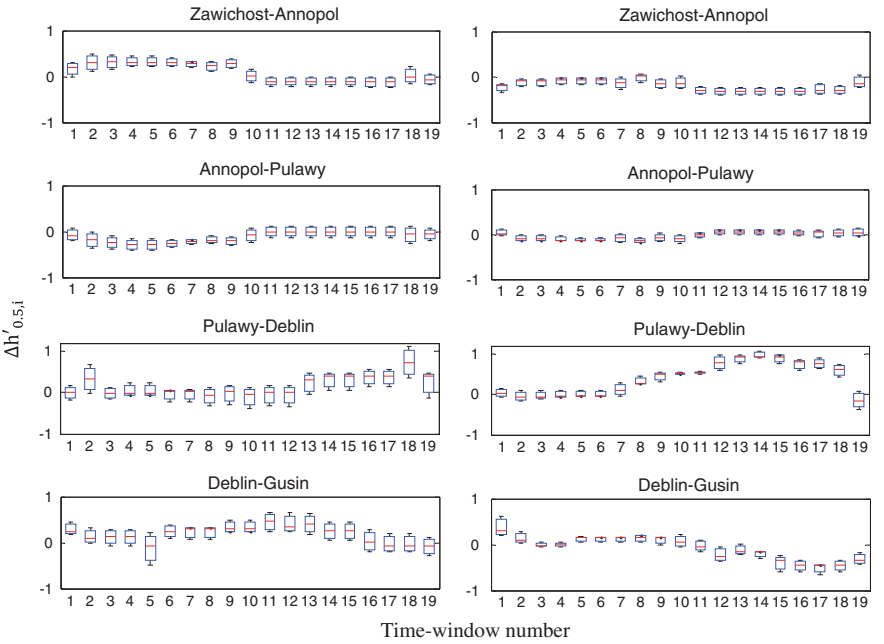

Fig. 5 Differences in water level quantiles between the downstream and upstream gauging stations $\Delta h'_{p,i}$, $\mathrm{i} = 2, \ldots, 5$ (Eq. 2) for a probability of exceedance $p = 0.5$; *left column* describes the summer season, *right column* describes the winter season, 0.95 confidence bounds are shown by box-plots, x-axis denotes time-window numbers as explained in Fig. 2

Assuming that positive differences in water level quantiles indicate an increase of resistance of the channel beds to the flow, and vice versa, negative differences show a decrease of flow resistance, we can make some estimates regarding the changes in flow regime in the reach during the study period for low, medium and high flows.

The Zawichost-Annopol reach shows a decrease of flow resistance during the second half of the period (i.e. years 1990–2003) for medium flows, but an increase of resistance for low flows, as shown in Fig. 3. Differences in high water level quantiles show a small variability for those periods.

The Annopol-Puławy reach shows opposite changes, i.e. an increase of flow resistance for the water level quantiles with a probability of exceedance of 0.5 in the years 1990–2003 compared to the earlier period.

The Puławy-Dęblin reach shows changes similar to the Annopol-Puławy reach, with very pronounced differences in 0.5 water level quantiles in the winter period.

On the other hand, the Dęblin-Gusin reach shows opposite changes, i.e. a decrease of the flow resistance after the year 1990, both in the summer and winter periods.

As mentioned, the high water level quantiles show a much smaller pronounced variability in winter periods than those for medium and low flows. Only the

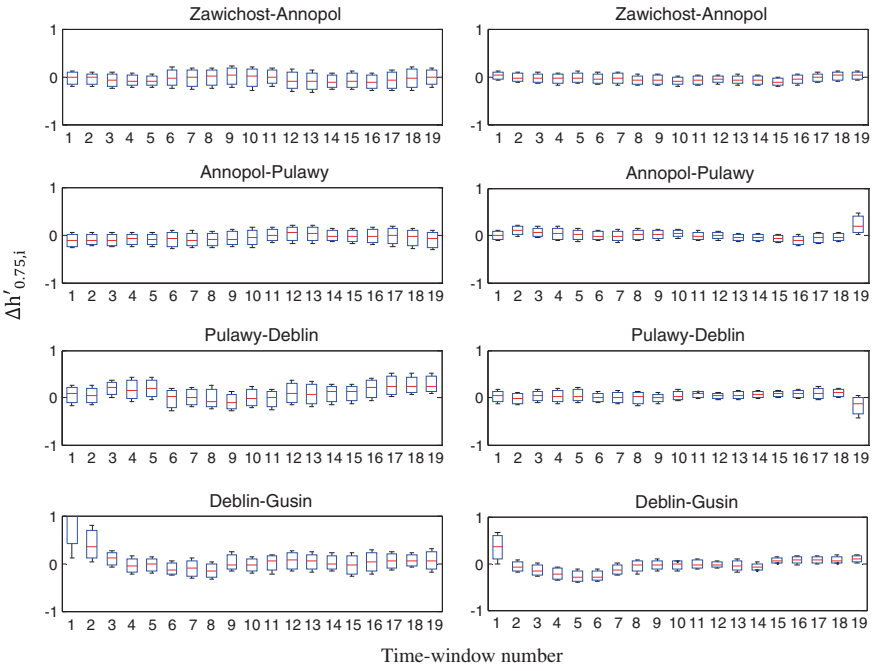

Fig. 6 Differences in water level quantiles between the downstream and upstream gauging stations $\Delta h'_{p,i}$, $\mathrm{i} = 2, \ldots, 5$ (Eq. 2) for a probability of exceedance $p = 0.75$; *left column* describes the summer season, *right column* describes the winter season, 0.95 confidence bounds are shown by box-plots, x-axis denotes time-window numbers as explained in Fig. 2

Deblin-Gusin reach shows relatively large variations. Interestingly, the pattern of changes, expressed by the opposite variability of quantile changes for the reaches Zawichost-Annopol-Puławy and Puławy-Dęblin-Gusin, is not maintained for high water levels compared to the results obtained for low and medium water levels.

4 Linear Routing Model Based on Water Levels: Application of a Stochastic Transfer Function (STF) Model

In order to look for changes in the dynamics of the river regime, linear stochastic transfer (STF) models were applied to the river reaches between the Zawichost-Annopol, Annopol-Puławy, Puławy-Dęblin and Dęblin-Gusin gauging stations (Fig. 1), for the period 1975–2003. The application of STF models to modelling water levels along river reaches was described in Romanowicz et al. (2006) which follows research on the application of the linear transfer function approach

introduced by Nash (1957) and Dooge (1986). The application of Stochastic Transfer Functions in the environmental sciences was described by Young (2001).

The identification of the model structure and estimation of its parameters are performed using the Simplified Refined Instrumental Variable (SRIV) method from the Captain toolbox (Young 1984; Taylor et al. 2007). The model parameters are estimated for ten year periods with moving time windows in order to track the changes in water levels dynamics. The goodness of fit criterion R_T^2 is defined as:

$$R_T^2 = 1 - \mathrm{cov}(h_{sim} - h_{obs}) \big/ \mathrm{cov}(h_{obs}) \tag{3}$$

where h_{sim} stands for the simulated water levels, h_{obs} denotes the observed water levels and cov(.) denotes covariance.

Apart from the estimates of mean values, the uncertainty of the predictions is also evaluated. The transfer function formulation is equivalent to a linear differential equation description of the process dynamics. The advantage of using the STF approach lies mainly in the simplicity of the Captain toolbox application and its statistical efficiency.

The determination coefficient for all calibrated models is in the range of 0.85–0.98, which indicates that the assumption of linearity was justified. The model parameters are used to derive residence times for each year, describing the response of the channel to external forcing during a period of over 29 years. We apply Monte Carlo simulations to estimate the uncertainty of the residence times. The estimated residence times together with 0.95 confidence limits for the Zawichost-Annopol reach, the Annopol–Puławy reach, Puławy-Dęblin reach and Dęblin-Gusin reach for the summer and winter season, are shown in Fig. 7. The red line represents the estimated median values, whilst whisker plots shows the 0.95 confidence limits of the predictions. Both the mean and variance of the predictions show a high variability during the analysed period.

It is worth noting that for the section Puławy-Dęblin we observe the opposite trend of changes compared to the two neighbouring sections, both in the summer and winter seasons. The residence time values for the sections Annopol-Puławy and Dęblin-Gusin are within the range 1–2 days and for the section Puławy-Dęblin, 0.5–1 days, due to the differences in lengths of the corresponding reaches. The variance of residence time estimates depends on the length of the river reach, but also depends on the season. The 10-year calibration periods overlap in most years except for the first and last (10th) year. Therefore any abrupt changes of residence times are a result of changes in those "boundary" years.

There is evidence for a substantial difference in the residence times between summer and winter periods. Those differences are due to icy conditions and snowmelt related flows in winter periods.

The Zawichost-Annopol reach shows a decrease of residence times in the 90s in comparison with earlier periods, more prominent in the winter than in the summer season.

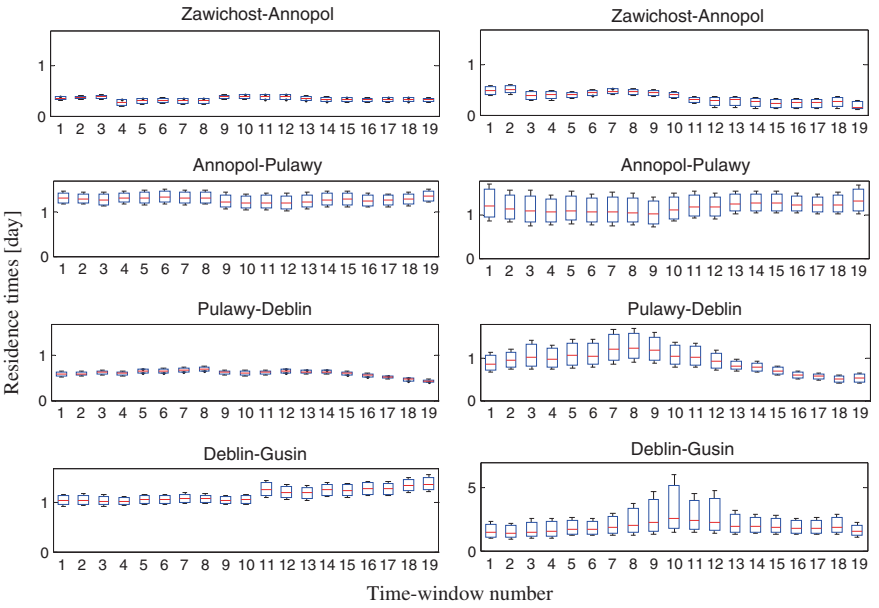

Fig. 7 Changes of residence times (in days) for summer (*first column*) and winter periods (*second column*) obtained from the STF models based on water levels, 0.95 confidence bounds are shown by box-plots, x-axis denotes time-window numbers as explained in Fig. 2

The Annopol-Puławy reach shows the opposite effect, i.e. an increase of residence times in the second part of the study period, once again more pronounced in the winter than in the summer periods.

The Puławy-Dęblin reach shows a decrease of residence times towards the end of the study period (after the mid-90s), more visible in the winter period.

The Dęblin-Gusin reach shows an increase of residence times after 1990 in the summer season. The winter season shows large uncertainties and an increase and subsequent decrease of residence times during the studied time period.

5 Discussion

Both differences in quantiles and STF residence times indicate changes in flow regime. Our aim is to explain those changes. The question is: can we explain the changes in flow regime by land use, including the vegetation and water works, or are they caused by morphological changes, or both?

Dynamic changes in flow regime are affected both by changes in land use (vegetation cover) and by values of water levels (flow velocity increases with water levels). Due to the icy conditions that may occur in winter, it is more difficult to model flow using a simple linear STF model. Moreover, vegetation has a larger

influence on high flows, since vegetation mainly occurs in the floodplains. And low flows are mainly influenced by river dunes and ripples in the main channel determined by the morpho-dynamics in the main river.

The variability of differences in water level quantiles can be explained, to some extent, by changes in the flow regime. The negative differences indicate a possible decrease of flow resistance (Ozga-Zielinski et al. 2010), whilst positive differences may indicate an increase of resistance to flow. This can be explained by a "back-water" type effect. However, these changes cannot be uniquely explained, as the differences in quantiles depend also on the quantile variability, as explained in Sect. 3 (Fig. 4). Those static changes correspond to changes in water residence times. We expect that residence times increase with an increase of flow resistance and decrease when water flow velocity increases.

In this respect, both static and dynamic analyses show consistent results, indicating a decrease of flow resistance in the Zawichost-Annopol reach during the period after 1990.

The Annopol-Puławy reach shows an increase of flow resistance, confirmed by both static and dynamic analyses, in particular in the winter season. These changes could possibly be attributed to changes in the channel, and in particular to the building of groynes along that reach between 1975 and 1990. Those changes should have affected only low and medium flows, which is confirmed by the analysis of differences in water quantiles between Annopol and Puławy.

The static and dynamic analyses for the Puławy-Dęblin reach give different results. Dynamic analysis indicates a decrease of residence times after 1990, whilst the static analysis shows an increase of flow resistance.

Also the results of the static and dynamic analyses for the Dęblin-Gusin reach are not consistent, where the static analysis indicates a decrease of flow resistance and the dynamic analysis shows the opposite pattern.

6 Conclusions

This paper has presented an analysis of changes in the flow regime in the River Vistula reach between Zawichost and Gusin in the years 1975–2003. Changes in the characteristics of the river regime have been analysed in a comprehensive way. First, the static relationships between water levels observed at the gauging stations have been examined through an analysis of differences between normalised water level quantiles, and second, a dynamic analysis using a water-level routing model based on the STF approach has been carried out. The static approach provides differences in normalised water level quantiles that can be used to explain changes in flow resistance. Vegetation has a larger influence on high flows, since vegetation will mainly occur in the floodplains. And low flows will be influenced mainly by river dunes and ripples in the main channel determined by the morpho-dynamics in the main river. The dynamic approach provides estimates of residence times of the flow between the gauging stations. Both static and dynamic approaches

provide uncertainty bands of their estimates. In this way the problem is being solved stepwise, gaining at each stage more detailed information on changes in the river regime within selected sections of the Vistula River.

The results indicate a complex pattern of changes in flow regime that is difficult to describe. Both static and dynamic analyses give consistent results regarding the variations in flow resistance of the first two reaches (Zawichost-Annopol and Annopol-Puławy) and opposite results for the next two reaches (Puławy-Dęblin and Dęblin-Gusin). However, it is difficult to attribute those changes to changes in land-use and water works in an unequivocal way; neither can we explain the inconsistency of the results for the Puławy-Dęblin and Dęblin-Gusin reaches. Further work is undergoing, including the application of flows instead of water levels as a driving variable and the use of event-based, rather than continuous, simulations together with a classification into low, medium and high flow events. In the next step we plan to apply a distributed model within a similar set-up as the STF approach to study the temporal and spatial variation of resistance along the river reach.

Acknowledgments This work was partly supported by the project "Stochastic flood forecasting system (The River Vistula reach from Zawichost to Warsaw)" carried by the Institute of Geophysics, Polish Academy of Sciences on the order of the National Science Centre (contract No. 2011/01/B/ST10/06866) and partly was done within statutory activities No 3841/E-41/S/2015 of the Ministry of Science and Higher Education of Poland. The water level and flow data were provided by the Institute of Meteorology and Water Management (IMGW), Poland.

References

Bathurst JC (1993) Flow resistance through the channel network. In: Beven K, Kirkby MJ (eds) Channel network hydrology. Wiley, New York, pp 69–98

Darby SE (1999) Effect of riparian vegetation on flow resistance and flood potential. J Hydraul Eng 125(5):443–454

Dooge JCI (1986) Theory of flood routing. In: Kraijenhoff DA, Moll JR (eds) River flow modelling and forecasting. Reidel Publishing Company, Dordrecht

Hutson AD (1999) Calculating nonparametric confidence intervals for quantiles using fractional order statistics. J Appl Stat 26:343–354

Järvelä J (2002) Flow resistance of flexible and stiff vegetation: a flume study with natural plants. J Hydrol 269(1/2):44–54

Kaplan DM (2011) Inference on quantiles: confidence intervals, p-values, and testing. MATLAB central file exchange, http://www.mathworks.com/matlabcentral/fileexchange/

Kaplan DM, Goldman M (2012) Two-sample nonparametric quantile inference using fractional order statistics. MATLAB central file exchange, http://www.mathworks.com/matlabcentral/fileexchange/

Kondracki J (1994) Geografia Polski. Mezoregiony fizyczno-geograficzne [Geography of Poland. Physico-geographical mesoregions]. Wydawnictwo Naukowe PWN, Warszawa

Kouwen N, Unny T (1973) Flexible roughness in open channels. J Hydraul Div Am Soc Civ Eng 99(HY5):713–728

Kowalski Cz (1997) Zabudowa hydrotechniczna i konieczne roboty zabezpieczające przed powodzią na odcinku Wisły od Annopola do rejonu Płocka [Hydro-engineering constructions and

works protecting from floods along the Vistula stretch between Annopol and Płock]. Gosp Wodna 1:18–23
Łajczak A, Plit J, Soja R, Starkel L, Warowna J (2006) Changes of the Vistula River channel and floodplain in the last 200 years. Geogr Polon 79:65–87
Li RM, Shen HW (1973) Effect of tall vegetations on flow and sediment. In: Journal of the Hydraulics Division, Proceedings of the American society of civil engineers, vol 99, No. HY5
Lopez F, Garcia M (2001) Mean flow and turbulence structure of open-channel flow through non-emergent vegetation. J Hydraul Eng 127:392–402
Maruszczak H (1972) Wyżyny Lubelsko-Wołyńskie. In: Klimaszewski M (eds) Geomorfologia Polski. Wydawnictwo PWN. Warszawa
Nash JE (1957) The form of instantaneous unit hydrograph. Surf Water Prevision Evap 45:114–121 International Association of Sciences and Hydrological Publications
Osuch M, Romanowicz RJ, Booij MJ (2015) The influence of parametric uncertainty on the relationships between HBV model parameters and climatic characteristics. Hydrol Sci J. doi:10.1080/02626667.2014.967694
Ozga-Zieliński B, Szkutnicki J, Kadłubowski A, Chudy Ł (2010) Wisła w Warszawie – wybrane problemy hydrologiczne. Gospodarka Wodna 12:490–493
Petryk S., and Bosmajian III, G. 1975, Analysis of flow through vegetation, J. Hydr. Div., ASCE, 101(7)
Romanowicz RJ, Young PC, Beven KJ (2006) Data assimilation and adaptive forecasting of water levels in the river Severn catchment. UK Water Resour Res 42:W06407. doi:10.1029/2005WR004373
Starkel L (2001) Historia Doliny Wisły od ostatniego zlodowacenia do dziś. Monografie IGiPZ, Warszawa
Taylor CJ, Pedregal DJ, Young PC, Tych W (2007) Environmental time series analysis and forecasting with the captain toolbox. Environ Model Softw 22(6):797–814
Tsujimoto T, Okada T, Kontani K (1993) Turbulent structure of open-channel flow over flexible vegetation. Progressive Report, December 1993. Hydr Lab, Kanazawa University
Warmink JJ, Straatsma MW, Huthoff F, Booij MJ, Hulscher SJMH (2013) Uncertainty of design water levels due to combined bed form and vegetation roughness in the Dutch River Waal. J Flood Risk Manage 6:302–318
Warowna J (2003) Wpływ zabudowy hydrotechnicznej na warunki sedymentacji w korycie powodziowym Wisły na odcinku Zawichost-Puławy [Impact of hydro-engineering construction on sedimentation conditions in the Vistula flood channel in the stretch between Zawichost and Puławy]. Wydawnictwo UMCS, Lublin (in Polish)
Yen BC (1991) Hydraulic resistance in open channels. In: Yen BC (ed) Channel flow resistance: centennial of manning's formula. Water Resources Publications, Littleton, pp 1–135
Young PC (1984) Recursive estimation and time series analysis. Springer, Heidelberg
Young PC (2001) The identification and estimation of nonlinear stochastic Systems. In: Mees AI (ed) Nonlinear dynamics and statistics. Springer, Heidelberg, pp 127–166

On the Run Length in Annual Maximum Flow Series in the Middle Vistula Basin in the Context of Climate Change Impact

An Attempt to Determine the Indicator for the Assessment of Changes in the Regime of High Waters

Ewa Bogdanowicz, Witold G. Strupczewski and Krzysztof Kochanek

Abstract This chapter presents an application of the 1st order Markov model to the description of flood-poor and flood-rich period length. The flood-poor and flood-rich periods are defined as a sequence of successive years with maximum flows lower (for flood-poor) or greater (for flood-rich) than the specified threshold. The theoretical distributions for two chosen thresholds (median and upper quartile) are compared with empirical distributions for the series of annual flow maxima at Warsaw gauging station, representative for the Middle Vistula River reach, in order to determine an indicator for the assessment of changes in flood regime under the future climate. The results have shown that serial dependency in Warsaw maximum flow series—the longest series in Poland—is rather weak and the run length distributions are not significantly affected by this dependency. The longest flood-poor period of 19 years has been found for the upper quartile threshold.

Keywords Flood frequency analysis · Flood-poor and flood rich periods length distribution

W.G. Strupczewski · K. Kochanek
Institute of Geophysics, Polish Academy of Sciences,
Ksiecia Janusza 64, 01-452 Warsaw, Poland
e-mail: wgs@igf.edu.pl

K. Kochanek
e-mail: kochanek@igf.edu.pl

E. Bogdanowicz (✉)
Project CHIHE, Institute of Geophysics Polish Academy of Sciences,
Ksiecia Janusza 64, 01-452 Warsaw, Poland
e-mail: ewabgd@igf.edu.pl

R.J. Romanowicz and M. Osuch (eds.), *Stochastic Flood Forecasting System*,
GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-319-18854-6_3

1 Introduction

The majority of research in the frame of Flood Frequency Analysis (FFA) is concerned with the choice of flood peak flows probability distribution and its estimation in order to evaluate upper quantiles under stationary and non-stationary conditions, which are commonly used as design characteristics for structures prone to high waters. The problem of the so-called 'flood-rich' and 'flood-poor' periods remains somewhat off the mainstream of investigations. However, these periods are important in the context of flood risk assessment and hydrological impacts of climate change. In Poland, similar as across all climate zones of Europe (e.g. Wilson et al. 2010; Pattison and Lane 2012; Macdonald 2014), flood events tend to cluster into periods that are relatively flood-rich and relatively flood-poor and these phenomena can be exacerbated by possible climate change. This chapter deals with the problem of distribution of lengths of these clusters by means of discrete Markov chains of order 1 on the Middle Vistula River represented by hydrological gauging station at Warsaw. The run lengths distribution assessed for observation period can be compared with the distribution obtained for hydrological response modeled on the basis of future climate predictions. The long and medium range projection of the observed structure of flood-rich and flood-poor years can also serve as an indicator to verification of climate change impact on annual maxima characteristics. The problem arises: are the proposed indicators stable enough in natural variability of observation series and sensitive enough to capture the changes in floods "reoccurrence" under the future climate?

The chapter is organized as follows. The first part introduces the basic definitions and run length distribution derived from 1st order Markov model. The next three parts give the descriptions of the Middle Vistula Basin, Warsaw annual maxima flow series (1799–2011) and its statistical properties. Then the theoretical run length distributions of annual maxima are compared with empirical distributions for Warsaw series. Finally, the conclusions and recommendation are presented.

2 Model of Run Length and Its Distribution

Given the stochastic process X of random variables X_t, $X = (X_1, X_2, \ldots, X_N, \ldots)$ and its realization in the form of the time series $x_1, x_2, \ldots, x_N$ of annual maximum flows, the problem is to define the flood-poor and flood-rich periods (runs, spells) as the uninterrupted sequences of years in which annual maximum flow is lower or equal (for flood-poor) or greater (for flood-rich) than specified threshold and find the probability distribution and statistical measures of run length. The value of the threshold can be chosen with regard to the aim of the study, i.e., severity of floods, and can be fixed on median or other quantile rank level.

This type of analysis is generally done on stationary processes in the strict sense, when understanding and application of the correlation structure in the time

series is of primary importance. The correlation structure of a time series can be described by autoregressive random (AR) model when the variable in time t depends linearly on its own previous values, autoregressive-moving-average (ARMA) models, Markov chains, wavelets and others. In this paper the first order Markov chain has been applied.

We will assume here below that the process X is stationary in a strict sense, so all random variables X_t are of the same probability distribution (identically distributed).

Let the cumulative probability distribution function of any X_t be $F(x) = P(X_t \leq x)$ and the chosen threshold be equal δ. It is convenient to introduce binary notation for two stages of x. Let "0" represent $X_t \leq \delta$ and "1" values of $X_t > \delta$. Then the absolute probability that in time t the process is in stage "0" can be written as:

$$p_0 = P(X_t \leq \delta) = F(\delta) \tag{1}$$

and the absolute probability of the stage "1" is given by:

$$p_1 = P(X_t > \delta) = 1 - F(\delta) \tag{2}$$

Four transition probabilities are needed to model the chain, namely p_{ij} $(i, j = 0, 1)$ corresponding to conditional probabilities that the process being at stage i in time t will be in stage j in time $t + 1$. Taking into account the Markov property of data dependence (i.e., the probability of the next stage depends only of the current stage) and denoting by $\boldsymbol{F}(\cdot, \cdot)$ the bivariate cumulative distribution function of any pair of successive random variables X_t, X_{t+1} one can obtain:

$$\begin{aligned}
p_{00} &= P(X_{t+1} \leq \delta | X_t \leq \delta) = \frac{\boldsymbol{F}(\delta, \delta)}{F(\delta)} \\
p_{01} &= P(X_{t+1} > \delta | X_t \leq \delta) = \frac{F(\delta) - \boldsymbol{F}(\delta, \delta)}{F(\delta)} \\
p_{10} &= P(X_{t+1} \leq \delta | X_t > \delta) = \frac{F(\delta) - \boldsymbol{F}(\delta, \delta)}{1 - F(\delta)} \\
p_{11} &= P(X_{t+1} > \delta | X_t > \delta) = \frac{1 - 2F(\delta) + \boldsymbol{F}(\delta, \delta)}{1 - F(\delta)}
\end{aligned} \tag{3}$$

The distribution of the run length D_0 when the process is in the state "0" with regard to the chosen threshold δ (flood-poor period) can be easily derived from the expression for the probability that the process remains d years in the state "0":

$$P(D_0 = d) = p_1 p_{10} p_{00}^{d-1} p_{01} \tag{4}$$

Thus, the joint probability mass function of D_0 can be expressed by:

$$P(D_0 = d, X_t = 1, X_{t+1} = 0) = \frac{(F(\delta) - \boldsymbol{F}(\delta, \delta))^2}{F(\delta)} \left(\frac{\boldsymbol{F}(\delta, \delta)}{F(\delta)} \right)^{d-1}, \quad d \geq 1$$

or in the form of conditional probability useful to compare theoretical and observed run length

$$P(D_0 = d|X_t = 1, X_{t+1} = 0) = \frac{F(\delta) - \boldsymbol{F}(\delta,\delta)}{F(\delta)} \left(\frac{\boldsymbol{F}(\delta,\delta)}{F(\delta)}\right)^{d-1}, \quad d \geq 1 \tag{5}$$

and the resulting cumulative distribution function (cdf) is:

$$F_{D_0}(d) = 1 - \left(\frac{\boldsymbol{F}(\delta,\delta)}{F(\delta)}\right)^{d} = 1 - p_{00}^{d}, \quad d \geq 1 \tag{6}$$

Similarly, for the length of the sequences of the state "1" we obtain:

$$\begin{aligned} P(D_1 = d|X_t = 0, X_{t+1} = 1) &= \tfrac{F(\delta)-\boldsymbol{F}(\delta,\delta)}{1-F(\delta)} \left(\tfrac{1-2F(\delta)+\boldsymbol{F}(\delta,\delta)}{1-F(\delta)}\right)^{d-1}, \quad d \geq 1 \\ F_{D_1}(d) &= 1 - \left(\tfrac{1-2F(\delta)+\boldsymbol{F}(\delta,\delta)}{1-F(\delta)}\right)^{d} = 1 - p_{11}^{d}, \quad d \geq 1 \end{aligned} \tag{7}$$

Note that in the case of independent data Eqs. 1–7 can be easily transformed by putting $\boldsymbol{F}(\delta, \delta) = \mathrm{F}(\delta) \cdot \mathrm{F}(\delta)$.

In the context of assessing the impact of variability and climate change on the distribution of the annual maximum flows, the long- and medium-range prediction of the structure of the occurrence of states "0" and "1" will be useful (Mitosek 2009; Zelias et al. 2008). Such a projection could be compared with the results of hydrological models defining the flood regime under the future climate. The medium-range prediction will be obtained by:

$$\boldsymbol{S}_t = \boldsymbol{S}_0 P^t \tag{8}$$

where $\boldsymbol{S}_0 = (p_0, p_1)$ is the vector of absolute probabilities of stages "0" and "1", the matrix P represent the transition matrix of the process $P = [p_{ij}]$, $i, j = 0,1$, and $\boldsymbol{S}_t$ is the vector of probabilities of the states in the time t.

The long-range projection in regular Markov chains, as described above, is based on the fact that there exists the limit:

$$\lim_{t\to\infty} \boldsymbol{S}_t = \lim_{t\to\infty} \boldsymbol{S}_0 P^t = \boldsymbol{S} \tag{9}$$

where $S = (s_1, s_2)$, called ergodic probabilities, does not depend on S_0 and can be find by solving the equations:

$$\begin{cases} \boldsymbol{S} = \boldsymbol{S}P \\ s_1 + s_2 = 1 \end{cases} \tag{10}$$

The values of s_i are interpreted as mean probabilities that the process will be in the state i in the infinite time span.

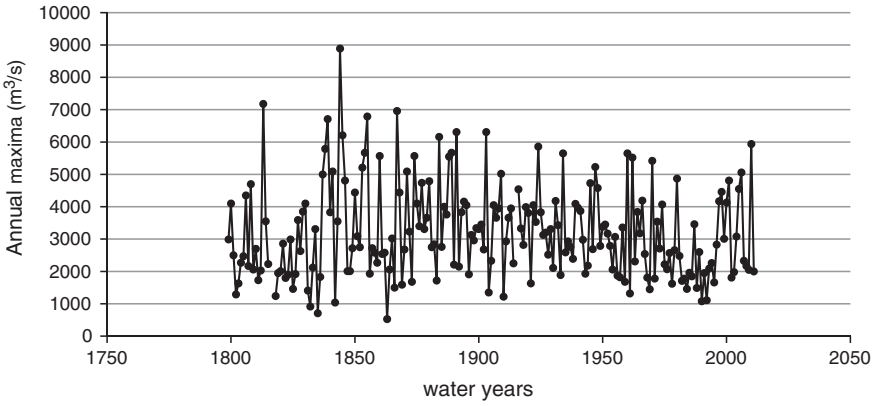

Fig. 1 Annual maximum flows at Warsaw in the period 1799–2010

3 The Middle Vistula Basin

The Middle Vistula basin (the catchment area from the San to the Narew confluence) lies entirely within the Polish borders and covers the area of 34 767.3 km^2 (Fig. 1, Chap. 1 of this book). The river (from Rkm 280 to Rkm 550) flows through the Małopolska Gorge, Mazowsze Lowland and Warszawa Dale. Average valley slope is about 0.27 %. About 50 % of the catchment area is covered by agriculture. The river has a natural braided appearance with numerous sand bars and vegetated islands. The natural character of the river, along with unique ecosystems, is protected by the NATURA 2000 areas (11 % of the basin). The Vistula is ecological corridor of pivotal importance connecting the Southern Carpathian Region with the Baltic Sea and the Eastern European regions. The city of Warsaw is the most populated and industrialized area in the basin. The Middle Vistula river is an allochtonous reach where the flood regime is formed in the conditions different from those through which it flows. The largest summer floods coming from the upper Vistula—mountainous part of the Vistula basin—are only slightly supplied here by tributaries (Kamienna, Radomka, Wieprz, Pilica et al.). Their peak flows usually decrease along the river course. Spring floods of snowmelt origin also reach here large sizes.

4 Warsaw Data Series

In Poland, first instrumental hydrological observations (river stages) had been started up at the beginning of the 18th century and the first flow measurements were performed in 1833. National hydrological and meteorological services were established in 1919, after the World War I and the age of partitions, when Poland regained its independence.

Due to the complex history of Polish state and destructions during both World Wars it is commonly accepted that only the flow data since 1951 can be regarded as sufficiently reliable (from the point of view of technology and organization of flow measurement and data processing procedures) for statistical inference in general, and for flow maxima and upper quantiles assessment in particular.

Well known problems in the statistical analysis of flow maxima are of both statistical and hydrological nature, which are rarely treated jointly. Statistical aspects of uncertainty of estimates of maxima characteristics result from short time series, unknown probability distribution function of annual peaks, the simplifying assumptions of identical independently distributed (i.i.d.) data and, in particular, the assumption of the homogeneity and stationarity of data series. The main hydrological aspect relates to uncertainty of hydrological measurements and data processing procedures.

The relatively simplest way to improve the accuracy of flood estimates is to lengthen the observation series using all possible sources of additional information primarily by both: regional approach and by using historical and paleohydrologic floods data. Historical data are recorded, non-regular observations of floods (mainly the highest water level), made before systematic data were collected. Although the historical extreme flood level records can be obtained from the reliable local chronicles, parish registers, high-water marks, boards and stones, the newspapers and further validated by field investigations, their accuracy is likely to be much lower than those of systematic records, which in fact are also not free from measurement error. Serious difficulties are related to an estimated stage-discharge relationship while historical river bed characteristics remains unknown, especially in alluvial channels. In addition, gathering and processing the historical information about floods remains extremely tedious work requiring to take into account a broad historical background and many aspects of language, ancient units of measure, etc., and their assembly and synthesis are time-consuming and labor-intensive (Girguś and Strupczewski 1965; Bielański 1984).

For Warsaw hydrological station, some interesting historical floods descriptions, research and analysis were published (Gutry-Korycka et al. 2006; Gutry-Korycka 2007; Magnuszewski and Gutry-Korycka 2009; Kuzniar 1997). The enormous work had been performed by Fal and Dabrowski (2001a, b) and, as a result, the series of annual maxima of water levels and flows starting from 1799 have been prepared. Updated and supplemented by recent flood data the series covers now 210 years of annual maxima in the period 1799–2011 with 2 gaps in the years 1816–1817 and 1915. The time series is shown in Fig. 1.

5 Some Statistical Properties of Warsaw Annual Maximum Flows Series

The essential question to be answered while analyzing long hydrological series with historical data is the question about their homogeneity and stationarity.

Fig. 2 Warsaw in the late 18th century (*left*) and today

Have changes in land use in the catchment such as urbanization (Fig. 2), deforestation or construction of water management structures affected flood regime to the extent that the assumption of stationarity is violated? Kaczmarek (1970) gives a simple rule. In the approximation sufficient for practical purposes it is generally assumed that, if the hydrological or meteorological time series exhibit no cyclic change, it can be considered stationary. Strupczewski et al. (2014) state that the utility of the pre-recorded flood information to FFA depends on whether it enables a worthwhile improvement in flood estimates to be obtained.

However, using various types of data (instrumental records, documentary information, and indirect data) of largely differing accuracy and reliability is not a trivial matter (Gutry-Korycka et al. 2006). The statistical tests of homogeneity/stationarity can be helpful. The power of statistical trend tests is improved when longer time series data are used, but problems of heterogeneity resulting from the measurement methods and measuring instruments will affect the results, which ought to be regarded in view of rather low data accuracy, especially for historical information and old records. And, if the existence of a trend cannot be rejected it should be decided what form of a trend will be applied. It is worth to note that free available software offers a plethora of distributions and methods of estimating trend in their parameters. However, looking for trends in parameters of distributions makes these trends dependent on the distribution and it happens that for different distribution models the recalculated trends in mean and variance differ significantly in magnitude and even direction.

As pointed out by Strupczewski et al. (2013), if the probability distribution of the population from which the data sequence is generated remains unknown (that is always the case), the method of simultaneous trend detection in the first two moments (mean and variance) by weighted least squares method (Strupczewski and Kaczmarek 2001) and then estimation of chosen distribution for detrended data and re-imposition of the trend gives more accurate time-dependent flood quantiles than the maximum likelihood estimation of trends in the parameters of pre-selected distribution, regardless of the size of the random sample, moment of time and the probability of non-exceedance.

In the case of unknown population distribution, the use of Mann-Kendall or Spearman tests is recommended (Report of COST Action ES0901 2013). The stationarity of mean and variance of Warsaw series was tested by non-parametric Spearman test. The hypothesis of stationarity in mean and variance cannot be rejected on the significance level 0.05 (p-values of the tests being 0.557 for the mean and 0.382 for variance).

The autocorrelation function of annual maxima is shown in Fig. 3.

All considered autocorrelation coefficients are almost constant for studied lags ranging from 1 to 15. Their values lie within the limits of confidence interval, so they can be assumed to be constant and equal zero. Thus, taking into account the results of Spearman test, Warsaw time series can be treated as stationary in weaker sense (second order stationarity).

The quantile regression method applied for annual maxima with linear trend functions gives the overall view of distribution form of the series.

Here, the τth regression quantile, $0 < \tau < 1$, is defined as any solution to the minimization problem (Koenker and Bassett 1978):

$$\min_{a,b\in R}\left(\sum_{t\in\{t:x_t\geq at+b\}}\tau\cdot|x_t-(at+b)|+\sum_{t\in\{t:x_t<at+b\}}(1-\tau)\cdot|x_t-(at+b)|\right) \quad (11)$$

which is a simple extension of the sample quantile definition. The notation $|\cdot|$ stands for absolute value of $\cdot$.

The regression lines found for chosen quantile ranks are presented in Fig. 4.

As it is visible in the Fig. 4, the quantiles of rank $\tau = 25$, 50 and 75 % show week and negligible trend, but 10 and 90 % quantiles exhibit increasing (quantile 10 %) and decreasing (quantile 90 %) trend suggesting that in contrast to the Spearman test result the variance decreases in time. Please note, however, that this effect is caused by the concentration of 5 highest and 3 minor floods during the first 70 years of the record and maybe application of the simplest form of quantiles trend. For the purposes of this study the results of Spearman test were accepted as more reliable.

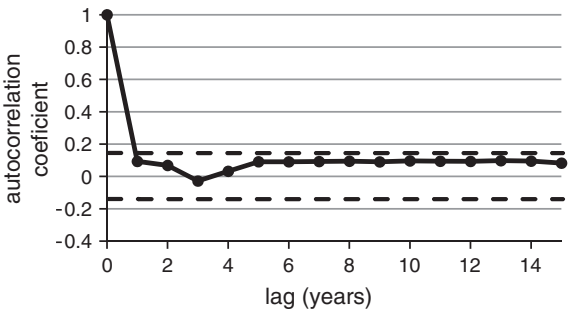

Fig. 3 Correlogram of annual maxima with 95 % confidence interval limits (*dashed lines*)

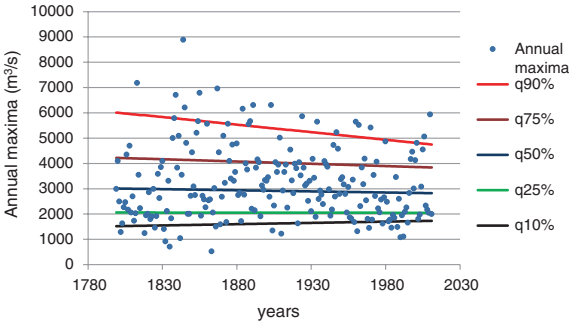

Fig. 4 Quantile regression for chosen quantiles of Warsaw annual maxima series

The three-parameter Weibul probability distribution with the pdf and cdf functions

$$f(x) = \frac{\lambda}{\alpha}\left(\frac{x-\varepsilon}{\alpha}\right)^{\lambda-1} \exp\left(-\left(\frac{x-\varepsilon}{\alpha}\right)^{\lambda}\right)$$
$$F(x) = 1 - \exp\left(-\left(\frac{x-\varepsilon}{\alpha}\right)^{\lambda}\right) \tag{12}$$

where α, ε and, $\lambda \in R_+$, was fitted to the data by means of the maximum likelihood method. The goodness of fit can be inspected visually in Figs. 5 and 6.

The choice of the probability distribution was made subjectively on the basis of numerous experiments with a choice of distribution type for the maximum rainfall and peak flows of Polish rivers. It should be noted that the real distribution in the population of the maximum annual flows is not known, the adopted distribution is only a model of it. Widely used procedure for selecting the best distribution on the basis of the AIC (Akaike criterion) gives the illusion of true distribution choice, and paradoxically, this "true" distribution after adding a few sample items may change.

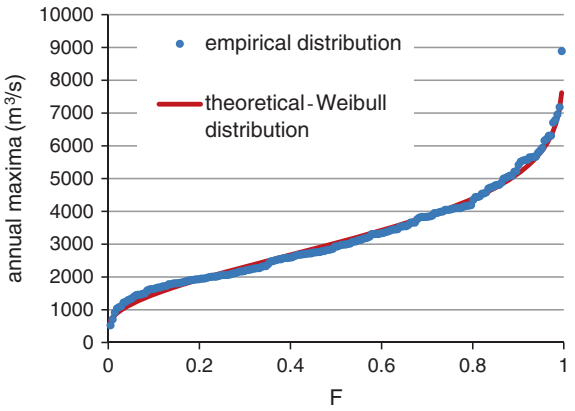

Fig. 5 Empirical cumulative distribution function and fitted Weibull distribution

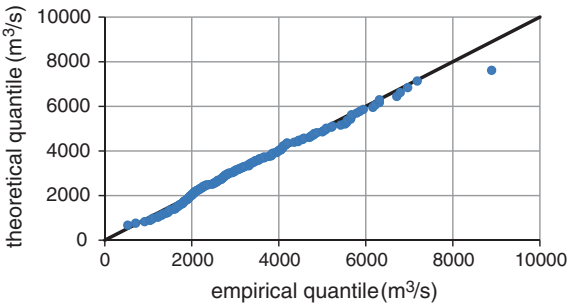

Fig. 6 The Q–Q plot. Fitted versus empirical quantiles

To model the bivariate distribution function of the pair of successive dependent random responses X_i, X_{i+1} the form of archimedean copula function with the Gumbel-Hougaard generator, given by:

$$C(u,v) = \exp\left(-\left((-\ln(u))^{\theta} + (-\ln(v))^{\theta}\right)^{1/\theta}\right) \tag{13}$$

was adopted.

The bivariate distribution $\boldsymbol{F}(x, y)$ can be expressed by:

$$\boldsymbol{F}(x,y) = C(F(x), F(y)) \tag{14}$$

where $F(\cdot)$ is the cdf function from Eq. 12.

A copula approach separates the dependence structure of a multivariate (in the general case) distribution from its marginal distributions. The IMF (Inference Functions for Margins) method was used for copula parameter θ estimation. When dealing with dependence in univariate time series, the corresponding copula is often called the univariate copula. In the case analysed here—the univariate copula with dependence in the form of 1st order Markov model—the log-likelihood function can be written as:

$$\ln L = \sum_{i=1}^{n-1} \ln\left(c(u_{i+1}, u_i, \theta)\right) + \sum_{i=1}^{n} \ln\left(f\left(x_i, \vec{\Psi}\right)\right) \tag{15}$$

where $c(.,.)$ is the copula density function, n is the time series length, and $\vec{\Psi}$ is a vector of parameters of pdf f.

It is assumed here that the probability of the flow value occurring at the i-th position, when nothing is known about the results of the preceding states is independent of i. This implies that the start of the sequence of observations is a randomly selected point in a longer sequence following the same probability law.

The copula parameter θ estimated by IFM takes the value close to 1, i.e., 1.0517, which confirms weak dependence between variables.

Table 1 Theoretical absolute and transition probabilities for two threshold values

Probability	Threshold δ Median (50 %)		Threshold δ Upper quartile (75 %)	
	Dependent	Independent	Dependent	Independent
$\boldsymbol{F}(\delta, \delta)$	0.2619	0.25	0.5735	0.5625
p_0	0.5	0.5	0.75	0.75
p_1	0.5	0.5	0.25	0.25
p_{00}	0.5238	0.5	0.7646	0.75
p_{01}	0.4762	0.5	0.2354	0.25
p_{10}	0.4762	0.5	0.7062	0.75
p_{11}	0.5238	0.5	0.2938	0.25

This weak dependence of the successive annual flows is reflected in close values of probabilities obtained from copula (14) and product model $\boldsymbol{F}(x, y) = F(x)F(y)$ (Table 1).

6 Run Length Distribution

Run length distributions were found for two threshold values corresponding to median (50 % quantile) and upper quartile (75 % quantile) and in two versions. The first version takes into account the first order autocorrelation; in the second it is assumed that the data are independent.

Elements of absolute probability vectors and transition matrices are presented in Table 1. In the case of dependency they are given by Eqs. 1, 2 and 3. For the independency case the same equations were transformed by putting $\boldsymbol{F}(\delta, \delta) = F(\delta) \cdot F(\delta)$.

The comparison of runs length distribution is shown in Figs. 7, 8, 9 and 10.

Analysis of the empirical distributions confirms the thesis about the existence of flood-rich and flood-poor periods. The longest observed sequence of years when the annual maxima are lower than the median value is 9 years (there were 2

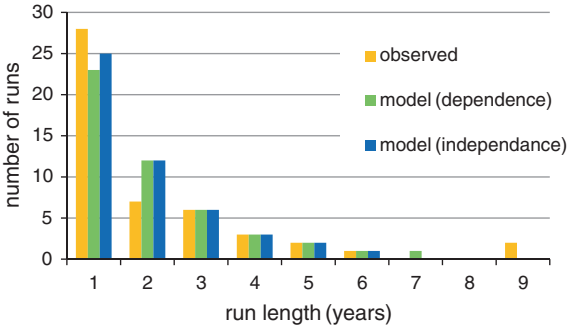

Fig. 7 Theoretical and empirical distributions of run length for the threshold equal to the median (lower values)

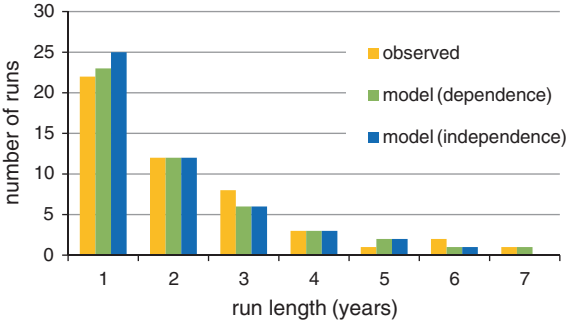

Fig. 8 Theoretical and empirical distributions of run length for the threshold equal to the median (upper values)

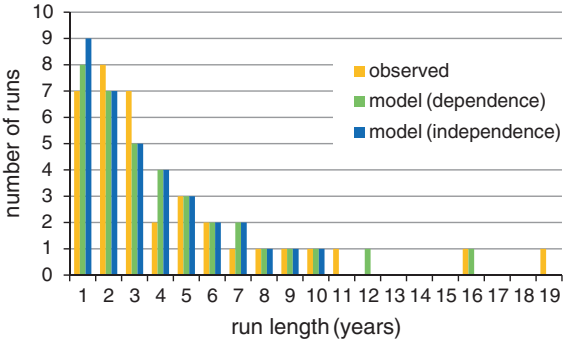

Fig. 9 Theoretical and empirical distributions of run length for the threshold equal to the upper quartile (75 %) (lower values)

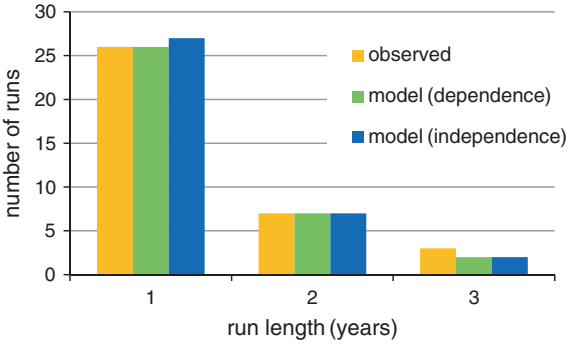

Fig. 10 Theoretical and empirical distributions of run length for the threshold equal to the upper quartile (75 %) (upper values)

runs of this length) delineating the two longest flood-poor periods, 1818–1826 and 1988–1996. The mean run length is equal to 2.2 with standard deviation 1.9. The longest run of values greater than the median is equal to 7, (the longest flood-rich period 1874–1880), the mean value being 2.2 with standard deviation 4.9 years.

The corresponding values for upper quartile are mean 4.4, standard deviation 4.1 for upper values and respectively 1.4 and 0.4 for lower. The longest run of flood-poor period for this threshold is 19 years (1818–1836) and, respectively, 3 years of flood-rich longest period occurred 3 times in the period 1799–1859.

Models describe well the general course of runs length frequency. The conformity between modeled and observed data was tested by chi-square test. At the significance level 0.05 the test rejected the model adequacy only in one case of run length for the threshold equal to the median (lower values) showed in the Fig. 7. Gottschalk (2004) interprets deviations of empirical and theoretical frequencies as a sign of the so-called Joseph effect (biblical cycle of lean and fat years), i.e., the tendency towards long sequences of dry and wet years. Furthermore, it has been questioned by hydrologists whether a Markov model with finite memory can describe the true nature of hydrologic time series at large scales.

In 1975 Erdos and Révész (1975) proved the theorem about the almost sure asymptotic behaviour of the length of the longest run of ones from a coin-tossing sequence (i.e., the X_i are i.i.d. random variables with $P(X_\mathrm{i} = 0) = P(X_\mathrm{i} = 1) = 1/2$). The probability that in n tosses a series of length $\log_2(n)$ will appear tends to 1 as n tends to infinity. In other words, when we are tossing long enough, we can be almost sure that longer and longer series of ones will appear. Thus, in a sequence of 210 tosses, we can expect a series of 7–8 ones. Note that the maximum observed lengths for the threshold equal to the median confirm their statement, the dependency measured by autocorrelation being close to zero.

Dividing the Warsaw series into two almost equal subseries and comparing by means of χ^2 consistency test with 0.05 significance level the runs length distribution in both periods, no statistically significant differences have been found. It seems, however, that typical length of hydrological series of about 60 elements or shorter and despite stationarity of the series confirmed by tests, the distribution of the length of runs in subseries may show important differences that puts in doubt the cognition value of indices based on comparison of runs length distributions of observed and projected series in future climate conditions.

7 Projection of the Structure of Flood and No-Flood Years

The medium-range forecast implying from Eq. 8 in the case of auto-correlated data was calculated for different time horizons. The successive powers of the matrix of transition quickly converge to the transition matrix for independent data. Thus, the future structures of flood and no-flood years are: for the threshold equal to the median—50 % for lower values and 50 % for upper and, respectively, for the threshold equal to the upper quartile—75 % for lower values and 25 % for upper.

The ergodic probabilities $\boldsymbol{S} = (s_1, s_2)$ calculated from Eqs. 10 for long term projections are the same as presented above, i.e., $\boldsymbol{S} =$ (50 %, 50 %) for the threshold equal to the median and (75 %, 25 %) for the upper quartile threshold. Thus, the projection is that 50 % of future annual maxima will be greater than the median and 25 % greater than the upper quartile threshold. These results are due to weak dependency of successive annual maxima and therefore they are compatible with assessments made on the base of maxima distribution function F, provided the process continues to be stationary.

8 Conclusions and Recommendations

- This analysis confirmed the tendency of annual maxima to group into flood-rich and flood-poor periods even in the case of independency of successive elements of the series.
- The maximum observed length of flood-poor run for the upper quartile threshold is 19 years. It is enough to suppress perception of risk of communities, planners and decision-makers and, consequently, their preparedness remains low, which can be dangerous when after years the flood is coming.
- The results obtained in this chapter may be biased by use of annual series which masks trends and statistical properties of seasonal maxima more sensitive to climate change and variability. Analysis of seasonal maxima can give deeper insight on correlation structure of the flood generation processes and thus enables the development of more adequate indices for regime changes assessment. However the seasonal series for Warsaw are much shorter than the annual maxima series analysed here.
- The choice of copula and generating functions should be analysed in more detail as they are of essential importance for determination of transition probabilities.
- At this stage of the study it is too early to formulate an appropriate and unique indicator to compare structures and length distributions of flood-rich and flood-poor runs during the observation period and the projection of the hydrological regime under future climate.
- Further research should include typical hydrological series length to take into account the sampling properties on the runs length results.
- The recommendation for comparison of observed and projected data periods in terms of changes in runs length distribution is to calculate and compare statistical characteristics of runs, but the inference about significant changes ought to be made with much precaution, especially in the case of the longest run length.

Acknowledgments This work was partly supported by the project "Stochastic flood forecasting system (The River Vistula reach from Zawichost to Warsaw)" on the order of the National Science Centre (contract No. 2011/01/B/ST10/06866) and partly by the Polish Norwegian Research Programme project CHIHE (Pol-Nor/196243/80/2013), both carried by the Institute of Geophysics, Polish Academy of Sciences. The water level and flow data were provided by the Institute of Meteorology and Water Management (IMGW), Poland.

References

Bielanski AK (1984) Materiały do historii powodzi w dorzeczu górnej Wisły (Materials for the history of flooding in the upper basin of the Vistula). Monografie Politechniki Krakowskiej, vol 30 (in Polish)

COST Action ES0901 (2013) European procedures for flood frequency estimation (FloodFreq). A review of applied methods in Europe for flood frequency analysis in a changing environment. Center for Ecology & Hydrology

Erdős P, Révész P (1975) On the length of the longest head-run. In: Csiszár I (ed) Topics in information theory. Coll Math Soc, János Bolyai, Budapest, pp 219–228

Fal B, Dąbrowski P (2001a) Dwieście lat obserwacji i pomiarów hydrologicznych Wisły w Warszawie: Obserwacje stanów wody (Two hundred years of hydrological observations and measurements of the Vistula River at Warsaw: Observations of water stage). Gospodarka Wodna, vol 11, p 461–467 (in Polish)

Fal B, Dąbrowski P (2001b) Dwieście lat obserwacji i pomiarów hydrologicznych Wisły w Warszawie: Przepływy Wisły w Warszawie (Two hundred years of hydrological observations and measurements of the Vistula River at Warsaw: the Vistula flows at Warsaw). Gospodarka Wodna, vol 12, pp 503–510 (in Polish)

Girguś R, Strupczewski W (1965) Excerpts from the historical sources dealing with extraordinary hydro-meteorological phenomena on the Polish territories from X–XVI c. Instr i Podr 87(165); Wyd Kom i Laczn, pp 216, Warsaw (in Polish)

Gottschalk L (2004) Time series modeling. In: Tallaksen LM, Van Lanen HAJ (eds) Hydrological drought, processes and estimation methods for steamflow and grounwater. Development in water science, vol 48

Gutry-Korycka M (2007) Wielkie wody Wisły środkowej w ujęciu historycznym (High waters of the Middle Vistula River in the historical perspective). Prace i Studia Geografi czne, vol 38, pp 85–103, (in Polish)

Gurty-Korycka M, Cyberski J, Grzes M, Nachlik E, Kundzewicz Z (2006) History of foods on the River Vistula. Hydrol Sci J 51/5

Kaczmarek Z (1970) Metody statystyczne w hydrologii i meteorologii. Statistical methods in hydrology and meteorology. Wydawnictwa Komunikacji i Lacznosci, Warszawa (in Polish)

Koenker R, Bassett G Jr (1978) Regression quantiles. Econometrica 46(1):33–50

Kuzniar P (1997) Woda 500-letnia w Warszawie w świetle materiałów historycznych I symulacji komputerowych. 500-year water in Warsaw, in the light of historical materials and computer simulation. In: Powodz 1997. Forum Naukowo-Techniczne. Ustron k.Wisly 10-12 wrzesnia 1997, t.2. Warszawa (in Polish)

Macdonald N (2014) Millennial scale variability in high magnitude flooding across Britain. Hydrol Earth Syst Sci Discuss 11:10157–10178. doi:10.5194/hessd-11-10157-2014

Magnuszewski A, Gutry-Korycka M (2009) Rekonstrukcja przepływu wielkich wód Wisły w Warszawie w warunkach naturalnych. Prace i studia geograficzne 43:141–151

Mitosek HT (2009) Metody statystyczne w hydrologii (Statistical methods in hydrology). Wydawnictwo Uniwersytetu Humanistyczno-Przyrodniczego Jana Kochanowskiego, Kielce

Pattison I, Lane SN (2012) The relationship between Lamb weather types and long-term changes in flood frequency, River Eden, UK. Int J Climatol 32:1971–1989. doi:10.1002/joc.2415

Strupczewski WG, Kaczmarek Z (2001) Non-stationary approach to at-site flood-frequency modelling, Part II. Weighted least squares estimation. J Hydrol 248:143–151

Strupczewski WG, Kochanek K, Bogdanowicz E, Feluch W, Markiewicz I (2013) Dwuetapowa metoda estymacji kwantyli powodziowych w warunkach niestacjonarności. Two-stage method of flood quantile estimation under non-stationary conditions. Monografie Komitetu Gospodarki Wodnej PAN z.35 (in Polish)

Strupczewski WG, Kochanek K, Bogdanowicz E (2014) Flood frequency analysis supported by the largest historical flood. NHESS

Wilson D, Hisdal H, Lawrence D (2010) Has streamflow changed in the Nordic countries? Recent trends and comparisons to hydrological projection. J Hydrol 394(3-4):334–346

Zelias A, Pawelek B, Wanat S (2008) Prognozowanie ekonomiczne : teoria, przykłady, zadania. Eeconomic forecasting theory, examples, problems. Wydawnictwo Naukowe PWN, Warszawa

Bed Topography and Discharge Measurements in the Świderskie Islands Nature Reserve, River Vistula, Poland

Robert J. Bialik, Joanna Sziło, Mikołaj Karpiński, Agnieszka Rajwa-Kuligiewicz and Oskar Głowacki

Abstract Field tests were carried out at a section of the River Vistula located within the boundaries of the Świderskie Islands nature reserve in order to analyse bathymetric changes at different discharge values during the period of 2012–2014. The paper shows that the dynamics and transport of sand-waves and sandbanks strongly influence the bathymetry of the riverbed at the selected cross-sections. Moreover, the impact of groynes on the bed topography is briefly discussed. Discharge measurements completed with the use of an acoustic Doppler current profiler (ADCP RiverSurveyor S5, SonTek) are also presented. Results support the values given by the Institute of Meteorology and Water Management (IMGW) at the gauging station in Gusin (461 km from the source) solely for the mean water levels. It is shown that the differences between measured values and values given by IMGW of the flow

R.J. Bialik (✉) · M. Karpiński · A. Rajwa-Kuligiewicz
Department of Hydrology and Hydrodynamics,
Institute of Geophysics, PAS, ul. Ks. Janusza 64, 01-452 Warsaw, Poland
e-mail: rbialik@igf.edu.pl

M. Karpiński
e-mail: mkarpin@igf.edu.pl

A. Rajwa-Kuligiewicz
e-mail: arajwa@igf.edu.pl

J. Sziło
Centre for Polar Studies, Institute of Geophysics, PAS,
ul. Ks. Janusza 64, 01-452 Warsaw, Poland
e-mail: jszilo@igf.edu.pl

O. Głowacki
Department of Polar and Marine Research,
Institute of Geophysics, PAS, ul. Ks. Janusza 64, 01-452 Warsaw, Poland
e-mail: oglowacki@igf.edu.pl

R.J. Romanowicz and M. Osuch (eds.), *Stochastic Flood Forecasting System*,
GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-319-18854-6_4

discharge ranged from 3 to 30 % of measuring values depending on the water level and discharge itself. The findings of this study may provide essential tutorial material for education in hydrology, civil engineering and environmental hydraulics as they emphasize large uncertainties inherent in hydrological modelling resulting from high dynamics of the bedload transport process.

Keywords Acoustic Doppler current profiler · Field measurements · Discharge measurements · River bathymetry · Hydrology · Sediment transport · Groynes

1 Introduction

The monitoring of rivers' bathymetry using boat-based surveys, particularly in large rivers such as the River Vistula, is an expensive and time-consuming endeavour; therefore, it is not performed on a regular basis. More often, the research on rivers with huge bedload transported by the moving fluvial bedforms (see Aberle et al. 2012) has mainly focused on the statistical characterization of such bedforms (i.e. Ashley 1990; Bialik et al. 2014b, c; Parsons et al. 2005) or on the way in which these bedforms are influenced by (1) natural processes such as floods (i.e. Chen et al. 2012); or (2) artificial, such as the existence of groynes in the river channel (i.e. Aberle et al. 2010). However, this approach represents only one method of analysis. For instance, from the pragmatic point of view, water discharge and the bathymetry of cross-sections are the most fundamental inputs in flood wave propagation models.

The purpose of this work is therefore to present the data obtained from field campaigns carried out at the River Vistula, Poland, which were associated with the measurement of the riverbed topography, velocity distribution and discharges at the selected cross-sections.

2 Study Sites

The section of the River Vistula, from Zawichost to Warsaw, which is located between two of its right tributaries: the River San (279.7 km) and the River Narew (550.5 km), is characterized by a unique ecosystem, which is protected by seven nature reserves, namely: Vistula near Zawichost; Cow Island; Brzeskie Sandbanks; Świderskie Islands; Zawadowskie Islands; Kiełpińskie Sandbanks; and Kazuńskie Clumps (starting from the most upstream located, see Fig. 1). The Świderskie and Zawadowskie Islands nature reserves are located south of Warsaw, partly within the administrative boundaries of the city. Both were created in 1998 in order to "preserve nesting sites of rare and endangered bird species in the area of the River Vistula" (Regulation of the Minister of Environment and Natural Resources). Apart from the wildlife conservation, these reserves preserve a unique area characterized by a wide variety of fluvial bedforms such as point bars, islands and submerged sand waves and dunes.

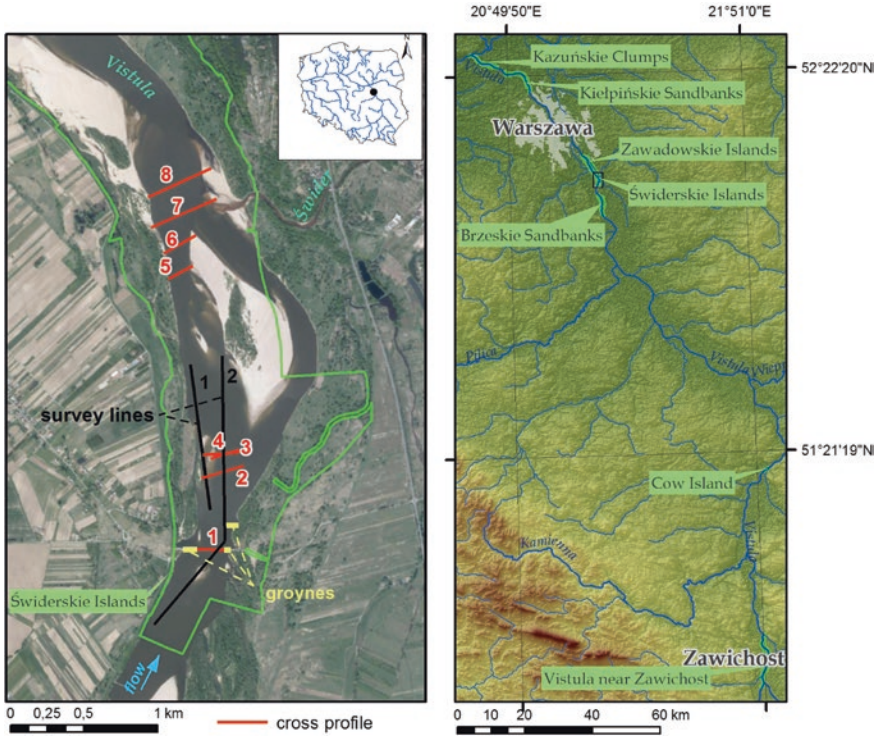

Fig. 1 The location of cross-sections and survey lines at the Świderskie Islands nature reserve (*left*) and location of all the nature reserves on the River Vistula at the section from Zawichost to Warsaw (*right*)

Measurements were carried out within the boundaries of the Świderskie Islands nature reserve (see Fig. 1) and field campaigns were performed during the autumn of 2012, 2013 and 2014. Data were collected using a small vessel equipped with the Structure Scanner and ADCP. The discharge and water velocities were measured with the ADCP (RiverSurveyor S5, manufactured by SonTek), whereas the river bathymetry was obtained from the structure scan, working at the frequencies of 455 and 800 kHz and echo-sounding system, working at the frequency of 1 MHz and accuracy 1 % of measuring values, which is a part of the ADCP. Figure 2 shows some views of the study site for better visualization of this part of the River Vistula.

3 Results

Figure 3 shows the bathymetric map of the study area calculated based on the data obtained with use of the structure scan working at frequency of 800 kHz. The influence of groynes on the bathymetry, which causes the scouring effect, is clearly visible at the distance of almost two kilometres downstream of the groynes' location.

Fig. 2 Views of the study site, **a** cross-section no. 1; **b** single damaged groyne at the right bank; **c** large sand bar in the middle part of the Świderskie Islands nature reserve; and **d** right bank during the low water level on 30-10-2013, which was below the mean low-water level of the River Vistula

Moreover, both banks are relatively shallow, characterized by the occurrence of mega-ripples and ripples on the surface of the riverbed. In order to measure the statistical characteristics of dunes, the bed elevation profiles made in the streamwise direction were carried out (see Figs. 1 and 4). Bialik et al. (2014a, b, c) noted that bedforms developing in the Świderskie Islands nature reserve are characterized by relatively high migration rates, with the approximated speed of the large dune crest about 4 m/h. Moreover, it was shown that large dunes propagate as a result of the movement of secondary dunes over the stoss side of the dune (see Fig. 4a) and between length (L) and height (H) of the dunes the $H = 0.09L^{0.31}$ relationships is fulfilled. This relationship can be utilized for inclusion in flow resistance formulae and helps in the modelling of flood propagation waves.

Due to the intensive sediment transport in the study area, significant changes in the depths of individual profiles could be expected. Figure 5a–i shows the spatial distribution of velocity and topography of some cross-sections, chosen arbitrarily (see Fig. 1). In Fig. 5a–b the cross-section no. 1, which was done between two groynes located on the both banks, is presented during two different seasons and the profile does not show significant differences regardless of the period of measurement. Moreover, in this cross-section two local scour holes at groynes' heads can be easily distinguished. To compare, sections no. 2–4 were made in the central part of the study area. It should be noted that the right bank of the river was excluded from our investigations due to the very low water depths, making this part of the

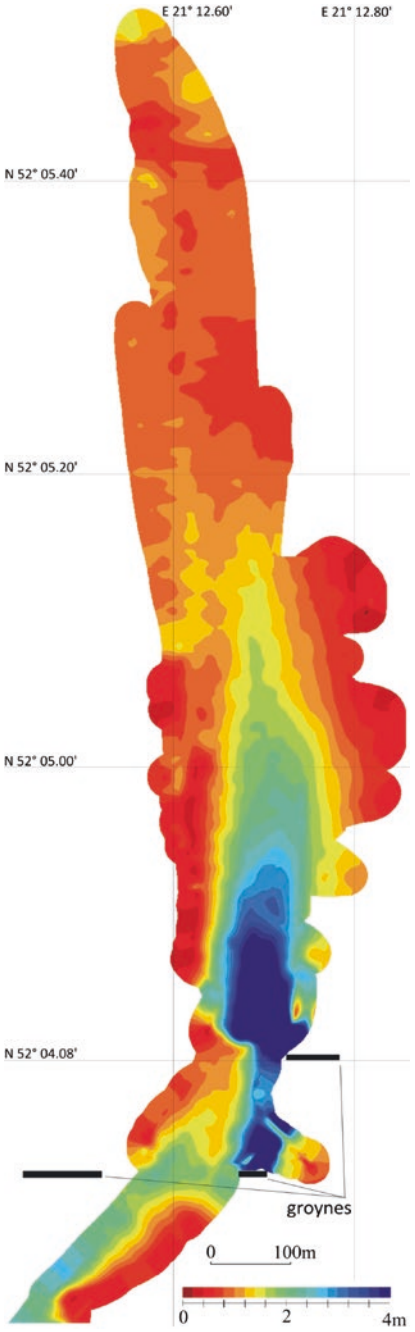

Fig. 3 Bathymetric map of the study site

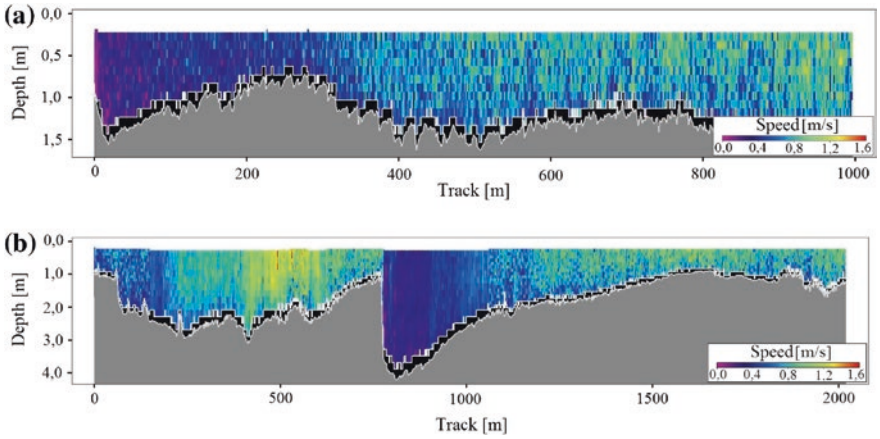

Fig. 4 Bed elevation profiles made in the stream-wise direction along the survey lines measured on 23-10-2013 and 30-10-2013. **a** Survey line 1 taken on 23-10-2013. **b** Survey line 2 taken on 30-10-2013

river impossible for carrying out the measurement. The location of these indicates that during the period of measurement, only the left bank of the river was changing, which is confirmed in the shape of the stream-wise profile, where the long bedform was transported. Furthermore, the higher speeds in the middle of the section, which results from the groynes' location, are visible. On the other hand, a very interesting phenomenon is noticeable in profiles 5 and 6. Water flowing along the right sides of the islands generates the vortex at the tail of the island/sandbar, which further results in the scouring of the sediment material. Therefore, on the right-hand side of these profiles the large depths are visible, which are greater in the vicinity of the vortex (profile no. 6). The last two sections (no. 7 and 8) were done at a short distance, about 50 m from the inflow of the River Świder, which in both periods had a different location. The influence of the Świder on the right side of the profiles, which results from the additional water flux, causes the erosion of the riverbed and local troughs. In addition, two profiles taken at different periods differ significantly, which may be related to the movement of large islands upstream from the River Świder inflow, which almost disappeared during the measurement campaign in 2014.

Concurrently with the bathymetric imaging, the measurements of water velocity and discharge were carried out. Table 1 compares the discharges in the Vistula

Table 1 Discharges of the River Vistula at the cross-section no. 1 measured using ADCP and given by the Institute of Meteorology and Water Management (IMGW) at the gauging station in Gusin

Data	Q [m^3/s] (measured)	Q [m^3/s] (IMGW)
08-11-2012	364	383
15-11-2012	369	399
23-10-2013	294	223
30-10-2013	285	201
4-11-2014	495	481

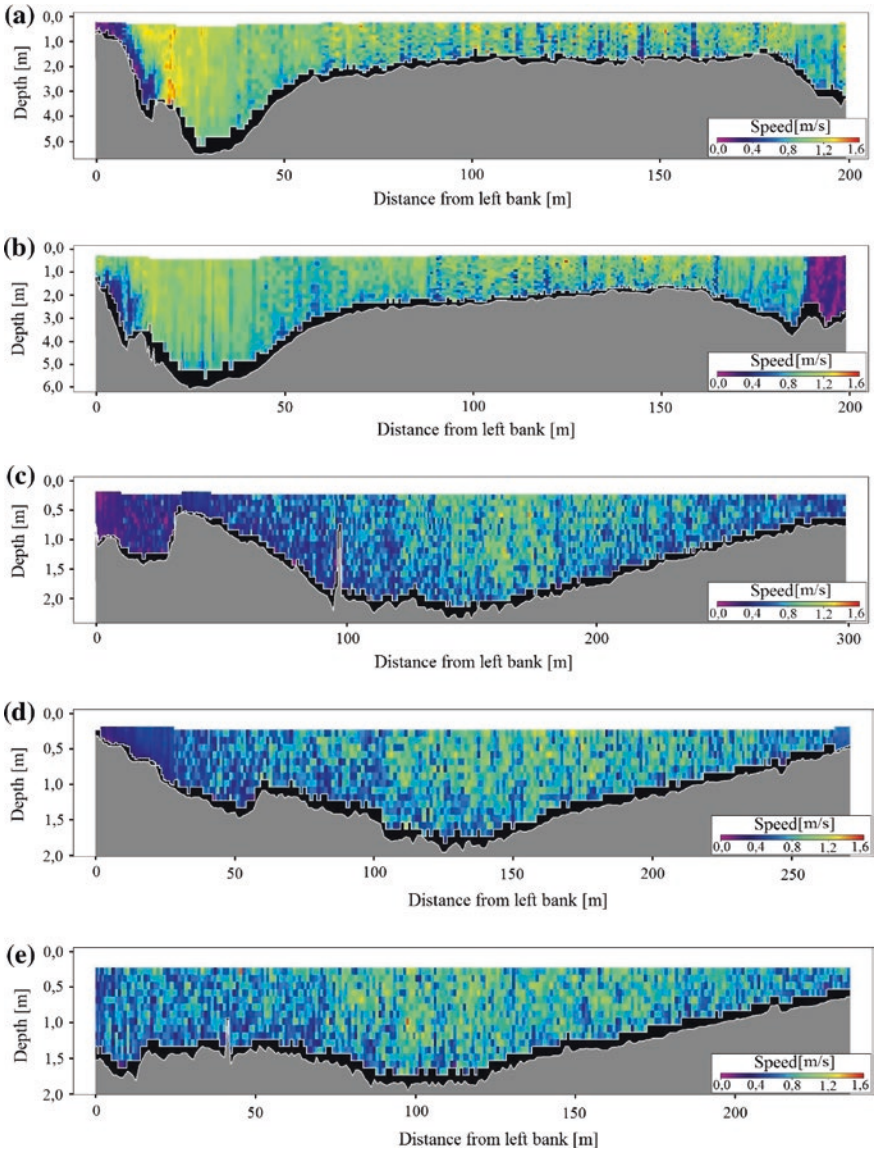

Fig. 5 Spatial distribution of velocity and the bathymetry in the cross-sections no. 1–8. **a** Cross-section no. 1 taken on 08-11-2012. **b** Cross-section no 1. taken on 23-10-2013. **c** Cross-section no. 2 taken on 23-10-2013. **d** Cross-section no. 3 taken on 30-10-2013. **e** Cross-section no. 4 taken on 23-10-2013. **f** Cross-section no. 5 taken on 15-11-2012. **g** Cross-section no. 6 taken on 15-11-2012. **h** Cross-section no. 7 taken on 08-11-2012. **i** Cross-section no. 8 taken on 04-11-2014

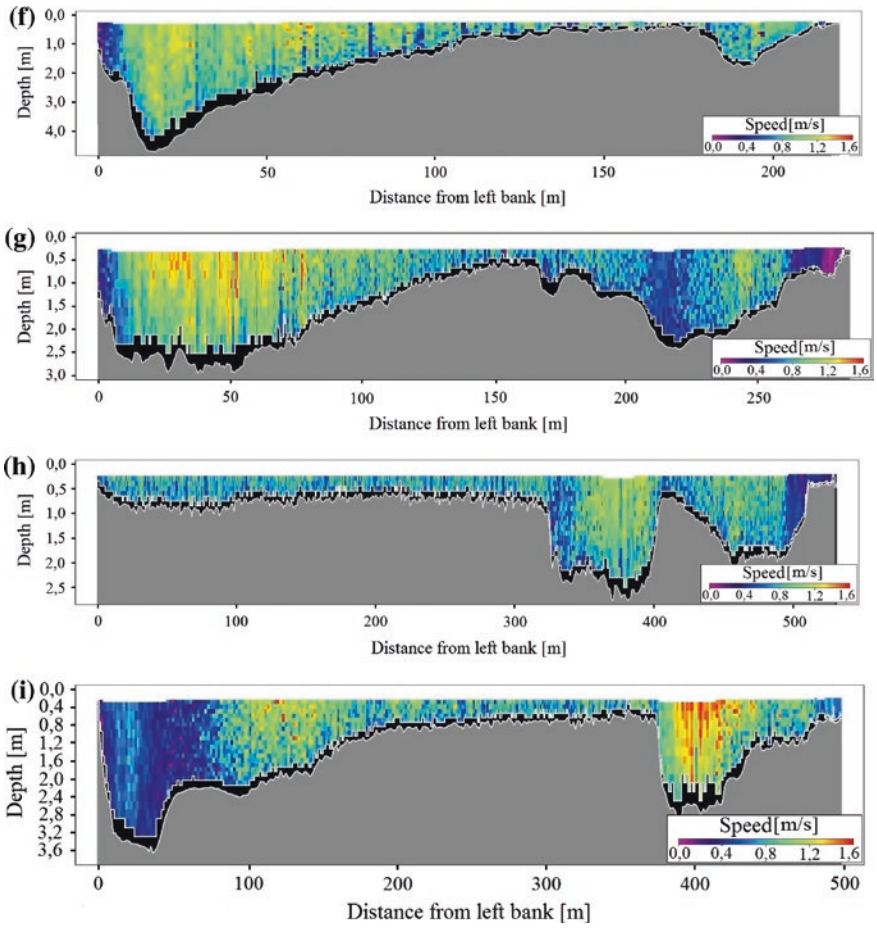

Fig. 5 (continued)

River at cross-section no. 1 measured using ADCP and given by IMGW at the water gauge in Gusin. It is demonstrated that below the low threshold value of the water discharge, which is equal to 338 [m^3/s] for Gusin, the differences between measured and given by the IMGW values are substantial, ranging up to 30 %. Moreover, these differences were decreasing with increasing water discharges and for the mean values of the water flow based on the multi-year measurement, were only 2–3 % of the measured values. This implies that the rating curve was probably determined incorrectly; in particular, large differences are visible for small flow rates, which can have its further consequences, such as the difficulties in proper stochastic modelling of the water discharges. It is worth mentioning here that, in contrast to the classical methods of discharge measurement, i.e. with use of an electromagnetic open channel flow meter, as indicated by Bialik et al. (2014a), ADCP allows to obtain more accurate results, which are not affected, for example, by the duration of the measurement.

4 Conclusions

In this paper, the flow velocity, discharge and topography measurements made using an acoustic Doppler current profiler (ADCP RiverSurveyor S5) and the structure scan were presented in order to show morphological changes in particular cross-sections for sandy bed rivers. Measurements were done on the River Vistula, within the Świderskie Islands nature reserve. It has been shown that at the relatively short distance of the river, the variability in the bathymetry of individual profiles can be very large, particularly when there is a very intense sediment transport, for example, in the form of the migrating fluvial dunes. In addition, the influence of the groynes on the topography of the cross-section was also observed. Moreover, the impact of islands as well as the impact of the tributaries of rivers on the shape of the profiles is noticeable. Attention has also been paid to the measurement of the water discharge. The results of measurements compared with the values given by IMGW at the gauge in Gusin have shown significant differences, ranging from 3 to 30 % of measuring values depending on the water level and discharge itself. Moreover, the obtained results clearly demonstrated that the proper selection of cross-section is necessary, for example, in the modelling of the transformation of flood waves. We suggest that the rating curve should be calibrated more frequently, especially in areas characterized by the intensive sediment transport and the ADCP seems to be the ideal device for this purpose.

Acknowledgment The study was supported by the Ministry of Science and Higher Education "IUVENTUS PLUS II" project No. 0028/IP2/2011/71 and within the Polish National Science Centre project "Stochastic flood forecasting system (River Vistula reach from Zawichost to Warsaw)". This work was also partially supported within statutory activities No 3841/E-41/S/2015 of the Ministry of Science and Higher Education of Poland. Joanna Szało gratefully acknowledges the financial support from the Interdisciplinary Polar Studies. Finally, the authors are very grateful to the Editors of this book: Professor Renata Romanowicz and Dr. Marzena Osuch for their invitation to write this chapter.

References

Aberle J, Nikora V, Henning M, Ettmer B, Hentschel B (2010) Statistical characterization of bed roughness due to bed forms: a field study in Elbe River at Aken, Germany. Water Resour Res 46:W03521. doi:10.1029/2008WR007406

Aberle J, Coleman SE, Nikora VI (2012) Bed load transport by bed form migration. Acta Geophys 60(6):1720–1743. doi:10.2478/s11600-012-0076-y

Ashley GM (1990) Classification of large-scale subaqueous bedforms: a new look at an old problem—SEPM bedforms and bedding structures. J Sediment Pet 60(1):160–172. doi:10.2110/jsr.60.160

Bialik RJ, Karpiński M, Rajwa A (2014a) Discharge measurements in lowland rivers: field comparison between an electromagnetic open channel flow meter (EOCFM) and an acoustic Doppler current profiler (ADCP). In: Bialik RJ, Majdański M, Moskalik M (eds) GeoPlanet: earth and planetary sciences: achievements, history and challenges in geophysics, part II. pp 213–222. doi:10.1007/978-3-319-07599-0_12

Bialik RJ, Karpiński M, Rajwa A, Głowacki O (2014b) Characterization of basic fluvial dunes parameters: a field study in the Vistula River, Poland. In: Proceedings of the 3rd IAHR Europe congress. Porto, Portugal

Bialik RJ, Karpiński M, Rajwa A, Luks B, Rowiński PM (2014c) Bedform characteristics in natural and regulated channels: a comparative field study on the Wilga River, Poland. Acta Geophys 62(6):1413–1434. doi:10.2478/s11600-014-0239-0

Chen J, Wang Z, Li M, Wei T, Chen Z (2012) Bedform characteristics during falling flood stage and morphodynamic interpretation of the middle-lower Changjiang (Yangtze) River channel, China. Geomorphology 147–148:18–26. doi:10.1016/j.geomorph.2011.06.042

Parsons DR, Best JL, Orfeo O, Hardy RJ, Kostaschuk R, Lane SN (2005) Morphology and flow fields of three-dimensional dunes, Rio Parana, Argentina: results from simultaneous multi-beam echo sounding and acoustic Doppler current profiling. J Geophys Res 110:F04S03. doi: 10.1029/2004JF000231

Part II
Physically Based Modelling

Sensitivity and Uncertainty Analysis of Precipitation-Runoff Models for the Middle Vistula Basin

Marzena Osuch

Abstract In this chapter the identifiability of a conceptual rainfall-runoff HBV model parameters in several tributaries of the Middle River Vistula is addressed. For this purpose, two approaches have been used. In the first, the application of a global sensitivity analysis by Sobol resulted in estimates of the influence of HBV model parameters on the goodness of fit of the model output. An alternative approach is based on an analysis of the posterior distribution of model parameters estimated with the help of an SCEM-UA algorithm. The results of the two approaches are similar. In both cases the parameters of the soil moisture store from the HBV model are better defined than the rest of parameters, as their influence is the highest and the relative standard deviation of parameter posterior distribution is the smallest. Small differences in the results are probably due to interactions between the parameters.

Keywords Rainfall-runoff modelling · Identifiability · Sensitivity analysis · Uncertainty analysis · Middle River Vistula

1 Introduction

In the framework of the project entitled "Stochastic Flood Forecasting System: The Middle River Vistula Case Study" carried out at Institute of Geophysics Polish Academy of Sciences a flood forecasting system has been formulated and developed. The system has a modular structure and consists of several models describing rainfall-runoff processes for the catchments (tributaries of the Middle River Vistula reach), and the flow routing within the reach. Details of that system

M. Osuch (✉)
Institute of Geophysics, Polish Academy of Sciences,
Ksiecia Janusza 64, 01-452 Warsaw, Poland
e-mail: marz@igf.edu.pl

R.J. Romanowicz and M. Osuch (eds.), *Stochastic Flood Forecasting System*,
GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-319-18854-6_5

are presented in Romanowicz and Osuch (2015, this book). This chapter deals with rainfall-runoff modelling in several tributaries of the study area (the Middle River Vistula reach between Zawichost and Warsaw Port Praski gauging stations). The analyses are carried out using the conceptual rainfall-runoff HBV model (Bergström 1995), one of the most popular rainfall-runoff models. The HBV model was calibrated and validated independently for each catchment with available hydro-meteorological data. In addition to model calibration and validation, an analysis of the identifiability of its parameters is performed to estimate how well the parameters are defined within the analysed model structure.

The identifiability of model parameters is related to parameter uniqueness (Wagener and Gupta 2005; Shin et al. 2015a, b). Non-identifiability mainly results from over-parametrization and model structure errors (Sorooshian et al. 1993; Andréassian et al. 2012), errors in model inputs and observations used to calibrate models and also may result from the shape of the objective function.

In hydrological applications the identifiability was analysed by means of a sensitivity analysis. Following that approach, the estimated sensitivity measures/indexes and related parameter ranks were used as indicators of parametric identifiability (Seibert 1997; Uhlenbrook et al. 1999; van Werkhoven et al. 2009; Booij and Krol 2010; Abebe et al. 2010; Zelelew and Alfredsen 2013; Zhan et al. 2013; Herman et al. 2013; Shin et al. 2015a, b). In the alternative approach, an identifiability study was conducted by estimation of parametric uncertainty at the model calibration stage (e.g. Vrugt et al. 2003; Romanowicz et al. 2013; Osuch et al. 2015).

In this study the identifiability of the HBV model parameters is assessed in two ways: sensitivity analysis by the Sobol method and parametric uncertainty estimation using the SCEM-UA method, both discussed in the next sections. The outcomes of these assessments are compared with other studies and the study allows the HBV model and its limitations to be understood in greater depth. The sensitivity and uncertainty analysis provides an insight into the dominant processes that drive the model behaviour. It also provides a basis for model simplification/redesign which could be used in further studies.

2 Study Area

The project study area covers the Middle River Vistula reach between Zawichost and Warsaw Port Praski gauging stations (Fig. 1). In this area, several tributaries join the Vistula. The area of these tributaries is slightly asymmetric, with a little larger proportion of the area on the left-hand side of the river. Catchments of the right tributaries occupy 46.4 % of the surface, whilst the catchments of the left tributaries are not significantly larger—they constitute 53.6 % of the surface. The largest right-hand tributaries include: Sanna, Kurówka, Wieprz, Okrzejka and Świder. The largest left-hand tributaries of the Middle River Vistula are: Kamienna, Iłżanka, Radomka, Pilica and Jeziorka. A list of the largest tributaries is shown in Table 1. As is evident from Table 1, the catchments differ in size.

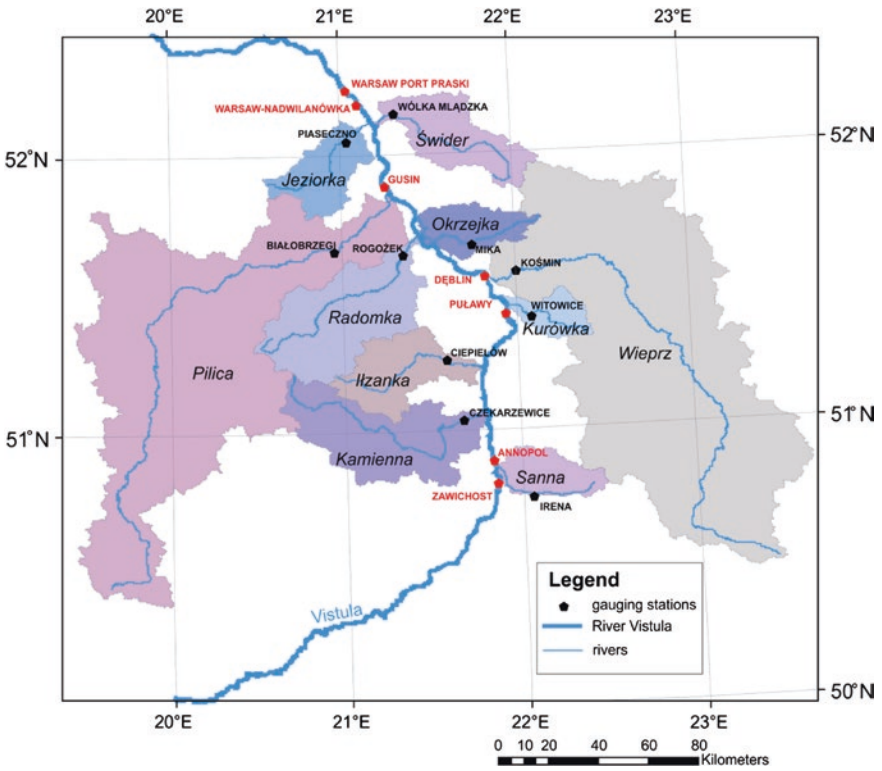

Fig. 1 Study area—tributaries of the Middle River Vistula

Table 1 A list of the largest tributaries of the Middle River Vistula and summary of catchment topographic conditions: mean, minimum, maximum and standard deviation of catchment elevation

River	Gauging station	Catchment area up to gauging station (km^2)	Mean elevation (m a.s.l.)	Min elevation (m a.s.l.)	Max elevation (m a.s.l.)	Std elevation (m)
Sanna	Irena	606.8	219.86	135	319	39.79
Kamienna	Czekarzewice	1878.4	247.75	128	610	61.73
Iłżanka	Kazanów	899.1	186.27	124	276	24.15
Kurówka	Witowice	335.4	170.59	115	230	23.65
Wieprz	Kośmin	10230.6	198.75	112	383	45.09
Radomka	Rogożek	2060.4	178.42	101	405	40.82
Okrzejka	Mika	300.2	151.98	99	205	26.70
Pilica	Białobrzegi	8664.2	213.61	94	486	55.84
Świder	Wólka Mlądzka	263.3	151.59	85	206	23.22
Jeziorka	Piaseczno 2	846.2	139.13	84	211	27.92

Additional information on land use, topography and soils was acquired from DTM constructed by the Geographic Survey of the Polish Army (DTED level 2), the CORINE Land Cover 2012 Project and the Harmonized World Soil Database (HWSD 2012).

2.1 Topography

The topography of the catchments was analysed by means of the digital terrain model DTED 2 compiled using diapositives of topographic maps at scales of 1:50000 and 1:25000 (Nita et al. 2007). The elevation data grids are given on 1 arc second, approximately 30 m. A summary of topographic conditions in all catchments is presented in Table 1. Mean catchment elevations vary from 139 for Jeziorka up to 248 m a.s.l. for the Kamienna catchment. As rainfall-runoff processes depend on differences in catchment elevation, Table 1 shows minimum and maximum elevation points for each catchment, and the standard deviation of elevation. Two catchments (Kamienna and Pilica) are characterized by variations higher than 50 m in elevation. Smaller differences in elevation are present in the other catchments. The flattest are the Kurówka and Świder catchments, with standard deviations of the elevation less than 25 m.

2.2 Land Cover

The analysis of land cover for the catchments was performed using the most recent data provided by the Warsaw Institute of Geodesy and Cartography (IGiK) in the framework of the CORINE Land Cover 2012 Project. Table 2 shows the percentages of six types of land use according to the 4th level of aggregation:

Table 2 Assessment of land cover in the tributaries according to the CORINE Land Cover 2012

River	Urban (%)	Agricultural (%)	Forest (%)	Meadows (%)	Open spaces (%)	Surface water (%)
Sanna	2.93	65.50	27.51	3.39	0.00	0.66
Kamienna	7.27	44.78	42.61	5.01	0.00	0.33
Iłżanka	3.96	54.89	31.73	9.18	0.00	0.23
Kurówka	7.95	65.29	15.60	10.69	0.00	0.47
Wieprz	5.55	61.96	21.01	10.64	0.00	0.85
Radomka	9.00	50.21	29.70	10.57	0.00	0.51
Okrzejka	2.86	54.41	32.07	9.94	0.00	0.72
Pilica	4.99	47.04	37.06	10.15	0.00	0.76
Świder	8.28	50.89	29.16	11.46	0.00	0.21
Jeziorka	11.31	58.32	15.58	12.37	1.99	0.44

urban, agricultural, forest, meadows, open spaces and surface water in the catchments. The catchments show variation of the land cover. Agricultural land cover is the most common type in all tributaries, varying from 44.78 % (Kamienna) to 65.50 % (Sanna). Forest is the second type of land cover, from 15.58 % (Jeziorka) to 42.61 % (Kamienna) of the catchment area. Urban areas were found in all catchments but with differences in percentage from 2.86 % for Okrzejka catchment to 11.31 % for Jeziorka catchment. The percentage of surface water in the catchments is relatively low, varying from 0.21 % for Świder to 0.85 % for the Wieprz catchment.

2.3 Soils

Soils and their properties belong to the most important physical catchment descriptors influencing parameters of rainfall-runoff models. Soil properties in the selected catchments were analyzed using the Harmonized World Soil Database (HWSD 2012). The HWSD database is a 30 arc second raster database with over 15,000 different soil mapping units. Each unit is linked to an attributed database providing information on soil unit composition (e.g. available water storage capacity in mm/m of soil unit AWC). The AWC classes have been estimated for all soil units, taking into account topsoil textural class and depth/volume limiting soil phases. For each catchment a weighted average AWC was estimated. The weights are related to the ratio of area occupied by each type of soil to the total catchment area. The results are presented in Fig. 2.

The lowest AWC (53 mm/m) was estimated for the Kamienna catchment. The second lowest AWC (77 mm/m) was found for the Okrzejka catchment. The Wieprz and Jeziorka catchments have average values of AWC (93 and 94 mm/m

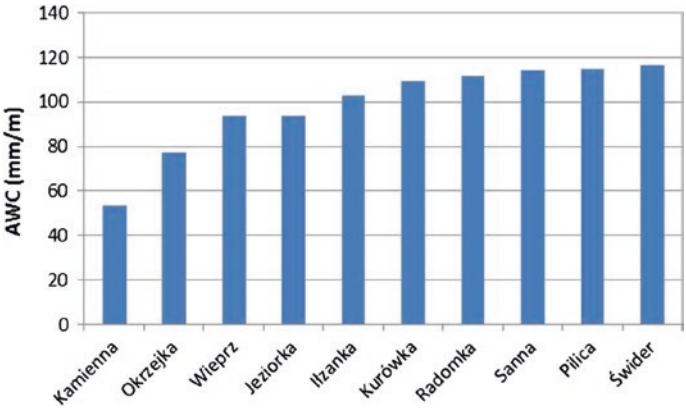

Fig. 2 Weighted averages of available water storage capacity (AWC) in mm/m for each catchment

respectively). High values of AWC (over 100 mm/m) were estimated for the Iłżanka, Kurówka, Radomka, Sanna, Pilica and Świder catchments.

3 Hydro-meteorological Data

Hydrological data in the form of daily flows from ten selected gauging stations closest to the river outlet were chosen. A list of gauging stations is presented in Table 3. Due the differences in the length of time series between stations, the period 01.11.2005–31.10.2010 was selected. Flow data for Iłżanka and Sanna are not available for that period; therefore, these two catchments were excluded from the further analysis. To compare flow conditions in the catchments, the characteristic flows derived as the minimum, maximum or mean using the annual minimum, mean and maximum daily flow values were calculated. In Poland these characteristics flows are used in hydrological practice, for example WNQ (maximum annual minimum flow), as the threshold values to distinguish hydrological drought (Tokarczyk 2013).

In Table 3 five characteristic flows: NNQ—minimum of annual minimum flows, SNQ—average of annual minimum flows, SSQ average of mean annual flows, SWQ average of maximum annual flows, WWQ—maximum of annual maximum flows) are shown. These values of characteristic flows permit catchment classification. The highest flow values of the characteristic flows are estimated for the Pilica catchment. According to this classification, the Wieprz catchment was found as second, also with high flow values compared to other catchments. Four catchments (Sanna, Kurówka, Okrzejka, and Jeziorka) have relatively low values of derived flow characteristics with estimates of WWQ smaller than 60 m^3/s.

The flow regime at a daily time step in selected gauging stations over a longer time period (01.01.1976–31.12.2010) is shown in Fig. 3. The flow regime assessed

Table 3 A comparison of hydrological conditions (characteristics flows) in eight studied catchments

No.	River	Station	NNQ (m^3/s)	SNQ (m^3/s)	SSQ (m^3/s)	SWQ (m^3/s)	WWQ (m^3/s)
1	Kamienna	Czekarzewice	2.22	3.10	7.30	36.10	74.30
2	Kurówka	Witowice	0.24	0.34	1.23	12.02	22.80
3	Wieprz	Kośmin	13.00	19.29	43.11	111.32	188.00
4	Radomka	Rogożek	1.52	2.12	7.28	30.19	71.00
5	Okrzejka	Mika	0.12	0.24	1.21	16.53	32.10
6	Pilica	Białobrzegi	14.80	16.60	3871	139.70	225.00
7	Świder	Wólka Mlądzka	0.54	0.90	4.29	32.10	57.70
8	Jeziorka	Piaseczno 2	0.10	0.59	3.02	19.78	40.60

NNQ—minimum annual minimum flow, SNQ—average annual minimum flow, SSQ average mean annual flow, SWQ average maximum annual flow, WWQ—maximum annual maximum flow

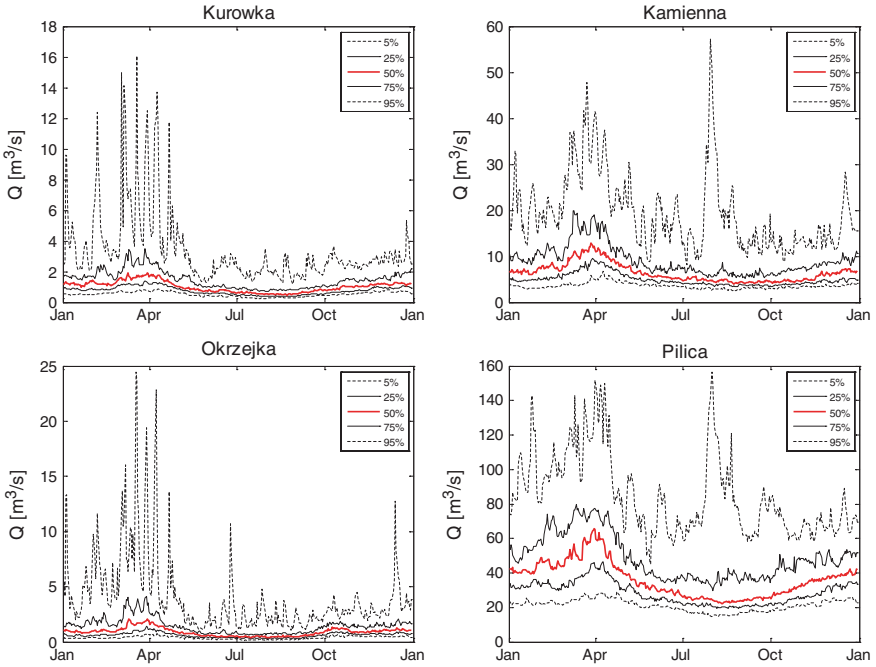

Fig. 3 Hydrological regime of selected catchments

using median and the 75th percentile shows that in all analysed stations the highest flows are observed in the spring months (March and April) as a result of snowmelt. In the case of left tributaries of the Middle River Vistula (Kamienna, Radomka, Pilica and Jeziorka), summer floods are also observed, even with higher flow values than in spring but not as often as spring floods. Low flow conditions are most common in summer and early autumn.

4 Methods

4.1 Rainfall-Runoff Models

For the rainfall-runoff modelling, a lumped version of a conceptual HBV model was applied (Bergström and Forsman 1973a, b; Bergström 1995; Lindström 1997; Lindström et al. 1997; Booij 2005; Booij and Krol 2010; Romanowicz et al. 2013; Osuch et al. 2015; Piotrowski et al. 2015). The HBV model is used as a standard tool in many countries but almost every nation/institution uses their own modified version of the model. A detailed description of the applied version of HBV model is presented by Piotrowski et al. (2015). In this version the HBV model has a structure with four conceptual storages taking into account the dominant processes

in rainfall-runoff modelling: (a) snow melting and accumulation, (b) soil moisture, (c) fast and (d) slow runoff (see Fig. 4).

The input variables to the HBV model are the daily sum of precipitation, the mean daily air temperature and the estimate of potential evapotranspiration (PET). In this study PET was calculated by the Hamon method (Hamon 1961).

The model has fourteen parameters: FC—maximum soil moisture storage (mm), β—nonlinear runoff parameter, LP—limit for potential evapotranspiration, α—nonlinear response parameter, KF—upper storage coefficient (1/day), KS—lower storage coefficient (1/day), PERC—percolation rate (mm/day), and CFLUX—maximum capillary rate (mm/day), TT—threshold temperature (°C), TTI—temperature threshold interval length (°C), CFMAX—degree-day factor (mm/°Cd), DTTM—a constant to be added to the TT parameter due to the difference between the threshold temperature for snow melt and the threshold temperature for rain/snow (°C), CFR—refreezing factor of water released from melting snow (–), WHC—water holding capacity of snow (mm/mm).

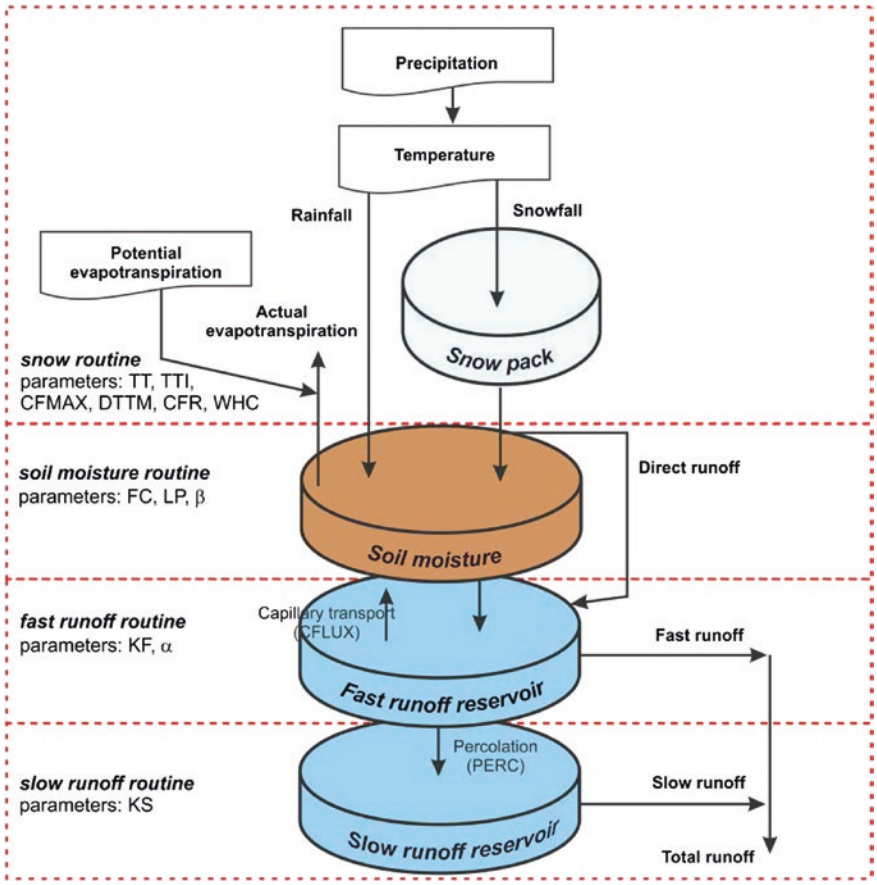

Fig. 4 The HBV model structure

Parameters: TT, TTI, CFMAX, DTTM, CFR and WHC are related to snow processes. The FC, LP and β parameters are related to the second conceptual store (soil moisture). CFLUX represents the capillary transport between the fast runoff reservoir and the soil moisture reservoir. The fast runoff is described by KF and α parameters. PERC describes the transport between the fast and slow runoff reservoirs (percolation). The slow runoff reservoir is represented by the KS parameter.

4.2 Model Calibration

Model calibration together with an estimation of parametric uncertainty was carried out using the SCEM-UA algorithm (Shuffled Complex Evolution Metropolis; Vrugt et al. 2003). This method combines a controlled random search, competitive evolution and complex shuffling together with Monte Carlo Markov Chain (MCMC) algorithm for the estimation of posterior parameter distributions. The posterior distribution is proportional to the product of the likelihood function and the prior pdf. In this study, the sum of squared errors (denoted as SSE) was used as a likelihood function. This SCEM-UA algorithm was used in order to combine the uncertainty analysis of the predictions with the parameter optimization. The choice of the likelihood function is important from the point of view of the aim of modelling. Here, the SSE was applied as the most general criterion.

The SCEM-UA algorithm can be presented in seven steps.

1. In the first step, a sample of s points (parameter set) is randomly chosen from the prior distribution and then the posterior density is computed for each point.
2. In the second step, the *s* points are sorted in order of decreasing posterior density and stored in a matrix D[1:s, 1:n + 1] where n denotes the number of parameters. The additional column n + 1 stores the posterior density function values.
3. The parallel sequences are initialized with the starting points S_1, S_2, …, S_q such that S_k is D[k, 1:n + 1] where k = 1, 2, …, q, where q denotes the number of complexes.
4. A set of s points is divided into q complexes C_1, C_2, …, C_q. Each one contains $m = s/q$ points. The points from the first complex are selected according to the formula q(j − 1) + 1 from the sorted points; the second complex contains every q(j − 1) + 2 ranked points and similarly for the next complexes, where j = 1, 2, …, m.
5. Later, each of the parallel sequences evolves following the Sequence Evolution Metropolis which is responsible for evolution. According to the SEM algorithm new candidate points in each of the parallel sequences are selected. A detailed description of the SEM algorithm is presented by Vrugt et al. (2003).
6. In the next step the complexes are (i) unpacked back into matrix D, (ii) than sorted out according to decreasing posterior density and (iii) again the s points are dived into q complexes following the same procedure
7. In the 7th step the convergence statistics by the Gelman and Rubin method are checked. If these criteria are fulfilled, the algorithm stops at that point; otherwise come back to the evolution step (step nr 4).

The application of the SCEM-UA algorithm requires the specification of two parameters: q—the number of complexes and s—the population size. The choice depends on the analysed problem. In the case of complex-shaped posterior probability density distributions characteristic for hydrological problems, the recommended values of these two parameters are $s \geq 250$ and $q \geq 10$ (Vrugt et al. 2003).

In this study the recommendation for complex problems was followed ($q = 10$, $s = 250$, maximum number of iteration is 50,000). MCMC is carried out using an informal likelihood measure, namely the sum of squared errors (denoted as SSE). The best fit is obtained for the lowest values of SSE.

In order to avoid problems with heteroscedastic and non-Gaussian error distributions, a Box-Cox power transformation (Box and Cox 1964) was applied to the simulated and observed flow data.

4.3 Sensitivity Analysis

Sensitivity study is the second method enabling an estimation of parametric identifiability. It this work, a sensitivity analysis by the Sobol method was applied (Archer et al. 1997; Saltelli et al. 2004). The method assumes that the total variance of the output from the model calculated by the Monte Carlo method can be decomposed into the sum of the variances explained by the various parameters of the model, and by interactions between them.

$$V(Y) = \sum_{i} V_i + \sum_{i<j} V_{ij} + \sum_{i<j<m} V_{ijm} + \cdots + V_{12\ldots(l-1)} \tag{1}$$

where: Y—analysed model output, $V(Y)$—the total variance of the analysed model output, i, j, m—the model parameters, l—the number of model parameters, V_i—the model output variance explained by the i-th parameter, V_{ij}—the model output variance, which is explained by the interaction between the parameters i and j, V_{ijm}—the model output variance explained by the interaction between the parameters i, j and m, $V_{12\ldots\,(l-1)}$—the model output variance explained by the interaction between parameters up to $(l-1)$ order.

The total variance of the output from the model can be decomposed into the sum of the partial variances. Decomposition is unambiguous if parameters are not correlated.

The results of the sensitivity analysis are usually presented as two sensitivity measures: the first order and total order sensitivity indexes. The first order sensitivity index (S_i) describes the influence of the parameter i on the model output.

$$S_i = \frac{V_i}{V(Y)} \tag{2}$$

The second sensitivity measure, the total sensitivity index, denotes the effect of a given parameter and the interactions of this parameter with other parameters on the model output.

The first order sensitivity indexes were used as a measure of the identifiability of model parameters. It is assumed the higher Si values are, the higher identifiability of the parameters.

In addition to the estimates of S_i values, the associated errors are also estimated, following the method presented in Homma and Saltelli (1996), with the help of the GSAT toolbox developed and presented by Cannavó (2012). The first order sensitivity index is used as a measure of parametric identifiability within the tested model structure.

5 Results

Application of the uncertainty assessment by the SCEM-UA method and also a sensitivity analysis by the Sobol method requires determination of the a priori distributions of the model parameters. In this work, the model parameters were drawn from uniform distributions within predefined upper and lower parameter boundaries (Table 4) based on studies by Seibert (1999, 2003), Booij and Krol (2010), Abebe et al. (2010), and Deckers et al. (2010). In both cases the hydrological model and its input are characterized by 18 parameters, where 14 are the appropriate HBV model parameters and the remaining 4 are the weights applied to input data from four meteorological stations (MS1, MS2, MS3, MS4). The weights for the input data were calibrated in order to take into account an uncertainty related to the estimation of the catchment rainfall. A list of meteorological stations for each analysed catchemnt is presented in Table 5.

Table 4 The HBV model parameters and their ranges used for sensitivity analysis and calibration

No.	Parameter unit	Description	Minimum	Maximum
1	FC (mm)	Field capacity, maximum soil moisture storage	0.1	1000
2	β (–)	Non-linearity parameter	0.01	7
3	LP (–)	Factor limiting potential evapotranspiration	0.1	1
4	α (–)	Non-linearity parameter	0.1	3
5	KF (1/day)	Fast runoff parameter	0.0005	0.3
6	KS (1/day)	Slow runoff parameter	0.0005	0.3
7	PERC (mm/day)	Rate of percolation	0.01	100
8	CFLUX (mm/day)	Rate of capillary rise	0	100
9	TT (°C)	Threshold temperature for snowfall	−3	3
10	TTI (°C)	Threshold temperature interval length	0	7
11	CFMAX (mm°C^{-1} d^{-1})	Degree day factor, rate of snowmelt	0	20
12	CFR (–)	Refreezing factor	0	1
13	WHC (mm mm^{-1})	Water holding capacity of snow	0	0.8
14	DTTM (°C)	Value to be added to TT to give the threshold temperature for snow melt	0	1

Table 5 A list of meteorological stations for each analyzed catchment used as input data

Catchment	MS1	MS2	MS3	MS4
Jeziorka	Skierniewice	Warszawa Okęcie	Puławy	Siedlce
Kamienna	Kielce	Bodzentyn	Łaziska	Sandomierz
Kurówka	Puławy	Lublin	Włodawa	Siedlce
Okrzejka	Jarczew	Puławy	Siedlce	Warszawa Okęcie
Pilica	Częstochowa	Łaziska	Silniczka	Skierniewice
Radomka	Puławy	Łaziska	Jarczew	Skierniewice
Świder	Jarczew	Siedlce	Warszawa Okęcie	Puławy
Wieprz	Puławy	Tomaszów Lubelski	Lublin	Włodawa

For consistency, the analyses are carried out within the same time periods as for the distributed flow routing model presented in Chap. 6 (Kochanek et al. 2015; this book). The model was calibrated using data from the period 01.11.2005–31.10.2010 and validated on observed flows from 01.11.2000 to 31.10.2001. For the calculation of the objective function only the last four years were used, assuming that after one year of simulation the effect of the initial conditions is not discernible.

5.1 Sensitivity Analysis

The Sobol first and total order sensitivity indices were calculated for eight catchments. In each case, 18 parameters were sampled separately using quasi-random Sobol sequences from a uniform distribution within the ranges presented in Table 4. A base of $b_n = 1000$ was used which, by simple calculation, results in $((2 * nr\_params) + 2) * b_n) = 38{,}000$ model evaluations for each catchment.

The outcomes of the sensitivity study by the Sobol method in the form of the percentage of the total model output variance explained for all study catchments are presented in Fig. 5. For the readability of the pie charts, only parameters with contribution higher than 1 % are labelled. In all catchments, the parameters β and FC have the highest influence on the model output. These two parameters (β and FC) explain 53 % of the variability of the HBV model output for Kamienna to 86 % for Pilica. It is worth mentioning that these two most important parameters are related to the second conceptual store (soil moisture). The contribution of the parameter β varies from 28 % for Kamienna catchment to 53 % for the Świder catchment. The maximum soil moisture in the catchment (FC) is the second parameter according to values of S_i in all catchments. This parameter explains from 16 % to 34 % of total model output variance.

The contribution of other parameters is much smaller (less than 7 %). The sum of influence of parameters other than β and FC vary from 14 % for Pilica to 46 % for the Kamienna catchment.

In addition to values of the Sobol first order sensitivity index, the absolute errors (Eq. 1) were also estimated. In general, the values of absolute errors depend

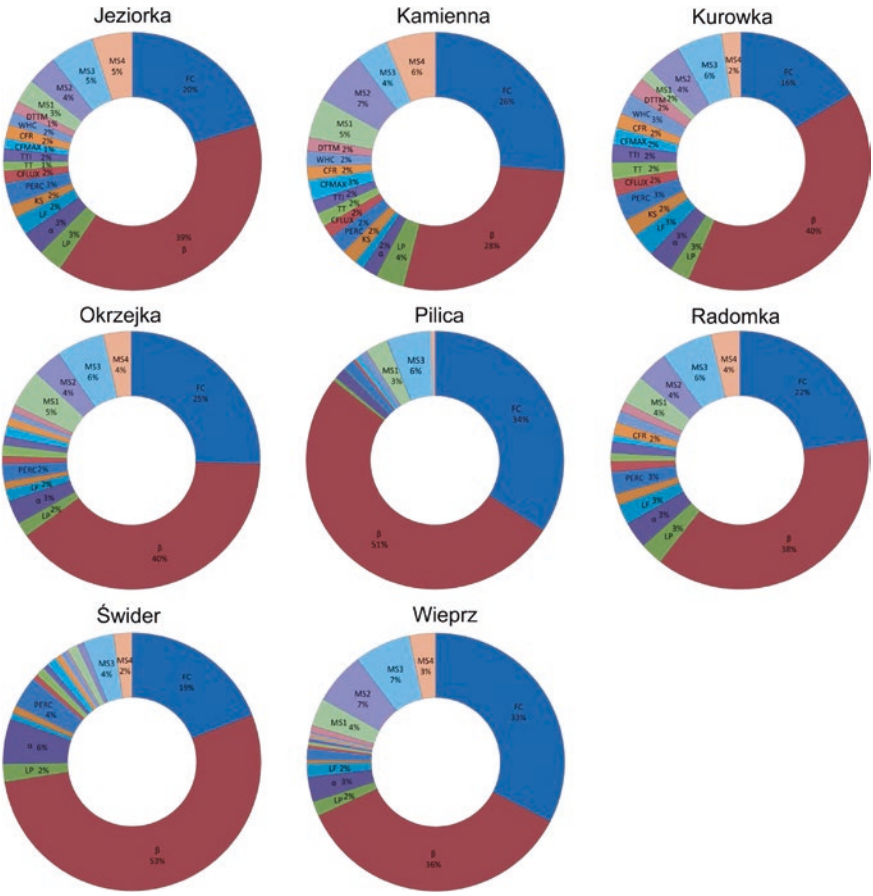

Fig. 5 The results of a sensitivity analysis by the Sobol method for eight analyzed catchments. The pie charts present percent of the total model output variance explained by model parameters. MS1, MS2, MS3 and MS4 denote weights applied to data from meteorological stations

on the analysed parameter and catchment. The highest values of absolute errors are obtained for the β and FC parameters. Figure 6 presents a comparison of estimates of the Sobol first order sensitivity index with related error for FC and β parameters for eight catchments. The values of S_{FC} are in range of 0.14–0.27, with the highest values for the Kamienna and Wieprz catchments. The results for the β parameter vary in the range from 0.27 to 0.46. The lowest values of S_{β} are obtained for catchments with the highest influence of the FC parameter.

5.2 Model Calibration by the SCEM-UA Method

As a result of the application of SCEM-UA, the posteriors of each HBV model parameter were derived. For each catchment the algorithm generated 10 parallel

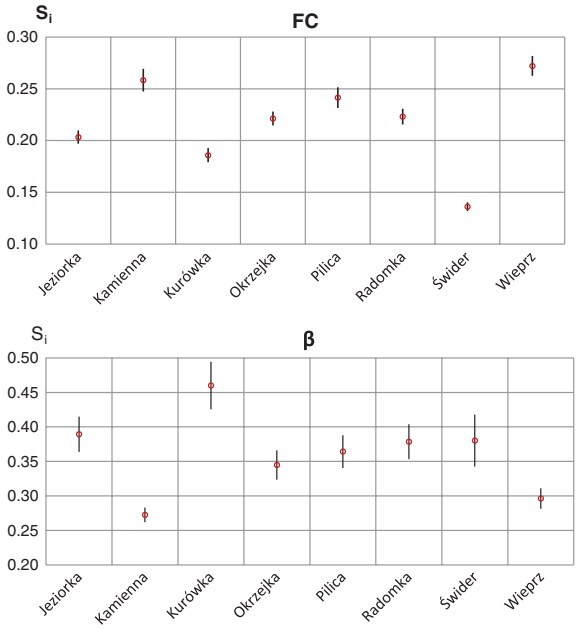

Fig. 6 A comparison of the Sobol first order sensitivity index together with its error estimated for eight studied catchments

sequences (q), each with 250 samples (s) using a population size of 25 points (m). In all analysed catchments, the SCEM-UA algorithm converged to a stationary posterior distribution.

The results of the model calibration using the SCEM-UA algorithm in the form of the best parameter set according to sum of squared errors (SSE), median and standard deviation of the posterior distribution are presented in Table 6. The optimum parameter values vary between catchments as well as median and standard deviation of the posteriors distributions.

In each case the posteriors of model parameters were examined. The shape of the posterior density depends on the catchment and on the parameter. An example of some posteriors estimated for Kurówka catchment is presented in Fig. 7. Some of the parameters, for example FC, MS1 (Puławy), MS2 (Lublin) and MS4 (Siedlce) follow a normal distribution, whilst some other evaluated parameters were excessively skewed to lower or upper values of the prior parameter range (e.g. α, TT, TTI). The application of multiplicative weights of input data gives very good results for most cases *inter alia* for the Kurówka catchment.

Hydrograph simulations based on the posterior parameter distribution show that the model was able to reproduce the timing of the observed flows with high accuracy.

The model parameter identifiability was tested by an estimation of parametric uncertainty at the model calibration stage. The evaluated error quantifies the extent of the identification accuracy. Due to the different ranges and units of the parameters, a comparison of relative standard deviation (RSD) defined as the absolute values of the coefficient of variation (the ratio of the standard deviation to

Table 6 The model calibration results

		FC (mm)	B (–)	LP (–)	KS (1/mm)	PERC (mm/d)	CFMAX (mm/°Cd)
Jeziorka	Best	553.87	1.34	0.12	3.91E−03	0.49	2.06
	Mean	348.10	2.26	0.42	4.40E−03	0.60	2.73
	Std	141.05	0.62	0.21	5.76E−04	0.10	0.59
Kamienna	Best	105.51	3.30	1.00	5.78E−03	0.91	14.81
	Mean	90.96	3.37	0.88	5.38E−03	1.50	15.78
	Std	18.11	0.31	0.09	4.22E−04	0.75	0.75
Kurowka	Best	57.43	3.95	0.49	1.07E−02	2.75	6.07
	Mean	57.23	4.12	0.58	1.17E−02	2.98	6.68
	Std	4.84	0.62	0.07	1.03E−03	0.82	2.84
Okrzejka	Best	409.57	2.07	0.40	4.45E−03	0.89	13.58
	Mean	372.82	1.91	0.37	4.57E−03	0.97	12.46
	Std	55.95	0.13	0.03	2.66E−03	0.20	1.23
Pilica	Best	83.91	2.73	0.85	1.90E−02	6.00	2.14
	Mean	94.71	2.47	0.75	1.87E−02	5.99	2.56
	Std	16.60	0.38	0.15	7.47E−04	0.01	0.54
Radomka	Best	20.60	2.34	0.33	4.08E−03	0.90	9.73
	Mean	22.22	1.98	0.30	4.43E−03	0.93	14.35
	Std	7.69	0.38	0.06	1.31E−03	0.44	4.58
Swider	Best	78.30	2.42	1.00	3.06E−02	3.43	3.99
	Mean	77.01	2.52	0.99	2.95E−02	3.69	5.88
	Std	4.65	0.15	0.01	1.91E−03	0.25	1.52
Wieprz	Best	52.69	3.73	0.66	1.07E−02	5.30	2.04
	Mean	71.10	2.88	0.61	1.25E−02	5.79	1.26
	Std	9.19	0.26	0.04	1.20E−03	0.32	0.39

Estimated values of the best parameter set according to applied objective function and two measures (mean and standard deviation) describing posterior distributions for the most important parameters

the mean) for 18 parameters in eight catchments has been calculated. The results, RSD expressed as a percentage, are shown in Figs. 8 and 9 on a logarithmic scale. The values of the RSD strongly depend on the parameter and catchment. A comparison of the RSD between parameters indicates that the highest vales of RSD are for: CFLUX, KF, CFR, and DTTM but huge differences are seen between catchments. The smallest values of RSD for almost all catchments are obtained for the β, LP, KS, FC, PERC and CFMAX parameters.

Comparing the RSD between catchments, it is seen that the results of calibration are worst for the Jeziorka catchment. This is probably due to the multimodality of posteriors. An example of the posteriors for the Jeziorka catchment is presented in Fig. 10. In the case of three parameters related to soil moisture store in the HBV model, two optima are noticed. These results indicate that the model in the present form could not be applied and should be the subject to additional tests (e.g. choice of different parameter ranges).

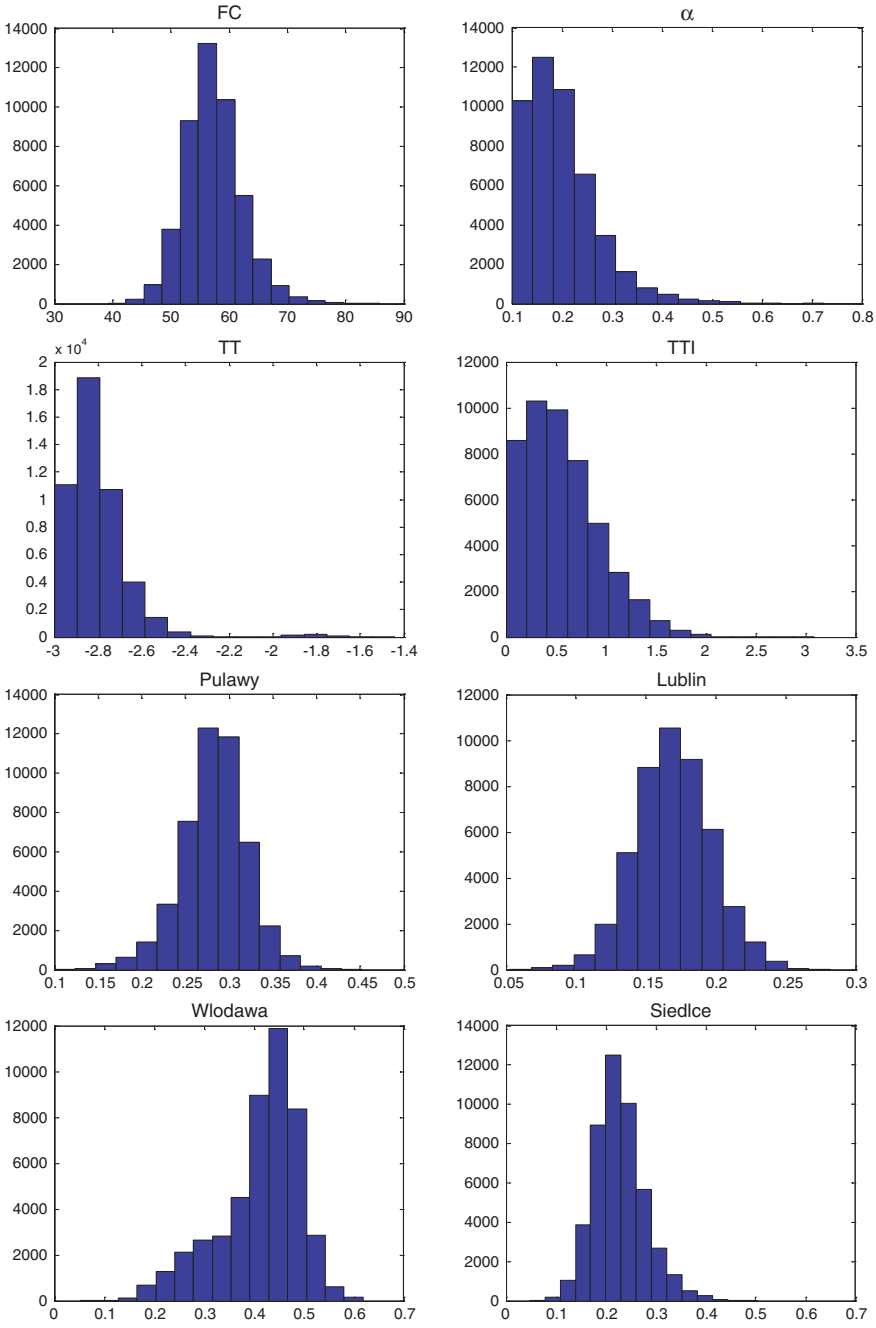

Fig. 7 Histograms of the estimated posteriors for selected parameters in Kurówka catchment

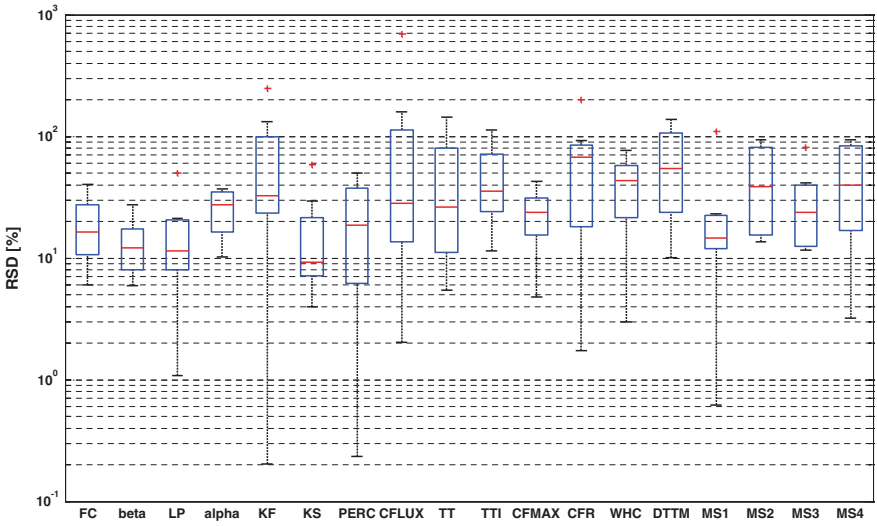

Fig. 8 A comparison of the absolute values of the coefficient of variation calculated for 18 parameters in eight catchments. Boxplots present the variability of the RSD values for the parameters. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually

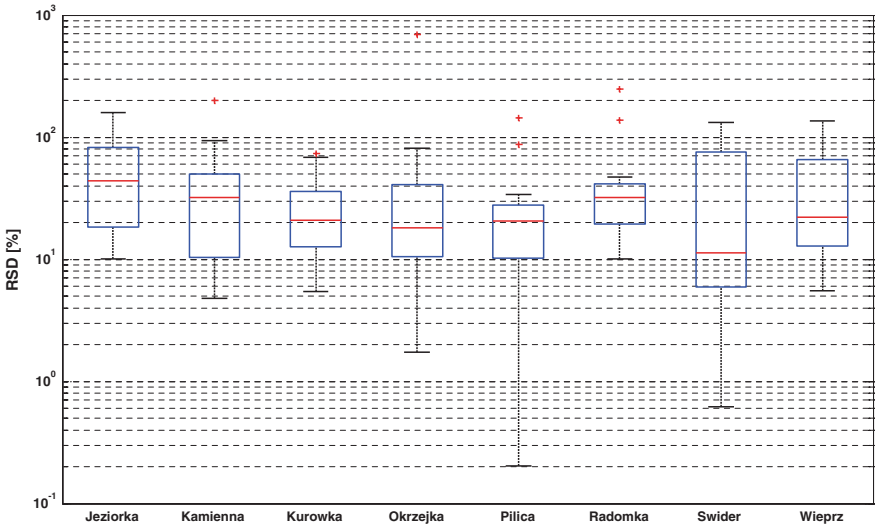

Fig. 9 A comparison of the absolute values of the coefficient of variation calculated for 18 parameters in eight catchments. Boxplots present the variability of the RSD values for the catchments. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually

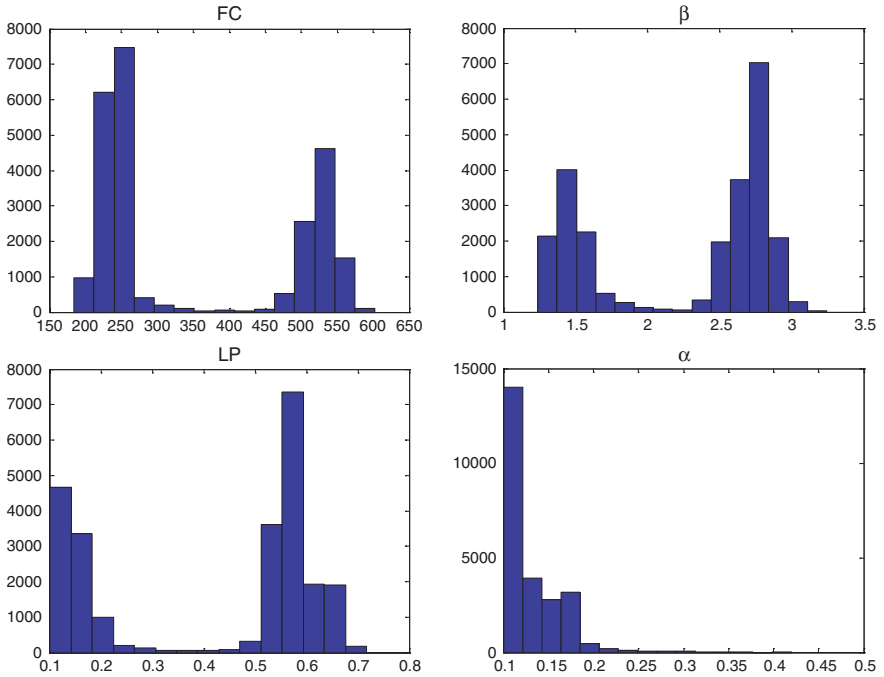

Fig. 10 Posterior histograms for the Jeziorka catchment

The smallest values of the RSD are derived for the Pilica and Kurówka catchments.

6 Summary and Conclusion

The chapter assesses the identifiability of HBV model parameters in several tributaries of the Middle River Vistula. For this purpose, two different approaches are used. In the first approach, an application of the global sensitivity analysis by Sobol resulted in the evaluation of an influence of HBV model parameters on the goodness of fit of the model output. The highest values of the Sobol first order sensitivity index S_i were estimated for β and FC parameters which are related to the second conceptual store (soil moisture) in the HBV model. These two parameters (β and FC) explain 53 % for Kamienna to 86 % for Pilica variability of the HBV model output. The contribution of other parameters is much smaller; the sum of the influence of parameters other than β and FC varies from 14 % for Pilica to 46 % for Kamienna catchment.

An alternative approach is based on an analysis of the posterior distribution of model parameters, estimated using the SCEM-UA algorithm. In that case the identifiability is assessed taking into account the shape of the posterior pdf and its

standard deviation. The outcomes from this method indicate that the shape of posterior density depends on the catchment and on the parameter. In the cases tested, three types of shape could be distinguished: (i) similar to a normal distribution, (ii) excessively skewed to lower or upper values of the prior parameter range and (iii) multimodal. The models with multimodal posteriors could not be applied in the present form and should be subjected to additional tests.

Due to the different ranges and units of the parameters, a comparison of the absolute values of the coefficient of variation relative errors (RSD) (%) for 18 parameters in eight catchments was performed. The values of the RSD strongly depend on the parameter and catchment. A comparison of the RSD between parameters indicates that the highest vales of the RSD (i.e. the worst result) are for: CFLUX, KF, DTTM, and CFR but huge differences are found between catchments. The smallest values of the RSD (i.e. the best result) for almost all catchments are evaluated for β, LP, KS, FC, PERC and CFMAX parameters.

The results of the identifiability assessment using two methods, sensitivity analysis by the Sobol method and relative error resulting from uncertainty analysis, lead to similar conclusions. In both cases the parameters of the soil moisture store from the HBV model are more easily identified as their influence is the highest and the relative standard deviation of posterior distribution is the smallest. Small differences in the results probably are due to interactions between parameters. The results of this study indicate that soil moisture store plays a dominant role in the HBV model description of the catchment response to rainfall. Further research is carried out to describe the interactions between the model parameters.

Acknowledgments This study was supported by the project "Stochastic flood forecasting system (The River Vistula reach from Zawichost to Warsaw)" carried by the Institute of Geophysics, Polish Academy of Sciences, on the order of the National Science Centre (contract No. 2011/01/B/ST10/06866). The hydro-meteorological data were provided by the Institute of Meteorology and Water Management (IMGW), Poland.

References

Abebe NA, Ogden FL, Pradhan NR (2010) Sensitivity and uncertainty analysis of the conceptual HBV rainfall-runoff model: implications for parameter estimation. J Hydrol 389:301–310

Andréassian V, Le Moine N, Perrin C, Ramos M-H, Oudin L, Mathevet T, Lerat J, Berthet L (2012) All that glitters is not gold: the case of calibrating hydrological models. Hydrol Process 26(14):2206–2210. doi:10.1002/hyp.9264

Archer G, Saltelli A, Sobol I (1997) Sensitivity measures, ANOVA-like techniques and the use of bootstrap. J Stat Comput Simul 58:99–120

Bergström S (1995) The HBV model. In: Singh VP (ed) Computer models of watershed hydrology. Water Resources Publications, Highlands Ranch, pp 443–476

Bergström S, Forsman A (1973a) Development of a conceptual deterministic rainfall–runoff model. Nord Hydrol 4:147–170

Bergström S, Forsman A (1973b) Development of a conceptual deterministic rainfall-runoff model. Nord Hydrol 4(3):147–170

Booij MJ (2005) Impact of climate change on river flooding assessed with different spatial model resolutions. J Hydrol 303:176–198

Booij MJ, Krol MS (2010) Balance between calibration objectives in a conceptual hydrological model. Hydrol Sci J 55:1017–1032

Box GEP, Cox DR (1964) An analysis of transformations. J Roy Stat Soc B 26:211–252

Cannavó F (2012) Sensitivity analysis for volcanic source modeling quality assessment and model selection. Comp Geosci 44:52–59

Deckers DLEH, Booij MJ, Rientjes THM, Krol MS (2010) Catchment variability and parameter estimation in multi-objective regionalisation of a rainfall-runoff model. Water Resour Manage 24:3961–3985

Hamon WR (1961) Estimation potential evapotranspiration. J Hydraul Div Proc Amer Soc Civil Eng 107–120

Herman JD, Kollat JB, Reed PM, Wagener T (2013) Technical note: method of Morris effectively reduces the computational demands of global sensitivity analysis for distributed watershed models. Hydrol Earth Syst Sci 17(7):2893–2903. doi:10.5194/hess-17-2893-2013

Homma T, Saltelli A (1996) Importance measures in global sensitivity analysis of nonlinear models. Reliab Eng Syst Saf 52(1):1–17

HWSD, 2012. FAO/IIASA/ISRIC/ISSCAS/JRC (2012) Harmonized world soil database (version 1.2). FAO, Rome, Italy and IIASA, Laxenburg, Austria

Kochanek K, Karamuz E, Osuch M (2015) Distributed modelling of flow in the middle reach of the River Vistula, this issue

Lindström G (1997) A simple automatic calibration routine for the HBV model. Nord Hydrol 28(3):153–168

Lindström G, Johansson B, Persson M, Gardelin M, Bergström S (1997) Development and test of the distributed HBV-96 hydrological model. J Hydrol 201(1–4):272–288

Nita J, Małolepszy Z, Chybiorz R (2007) A digital terrain model in visualization and interpretation of geological and geomorphological settings. Prz Geol 55:511–520

Osuch M, Romanowicz RJ, Booij M (2015) The influence of parametric uncertainty on the relationships between HBV model parameters and climatic characteristics. Hydrol Sci J. doi:10.1080/02626667.2014.967694

Piotrowski AP, Napiorkowski MJ, Napiorkowski JJ, Osuch M, Kundzewicz ZW (2015) Are so called modern metaheuristics successful in calibrating simple conceptual rainfall-runoff models?. Hydrol Sci J, submitted

Romanowicz RJ, Osuch M (2015) Semi-distributed flood forecasting system for the River Vistula reach, Chapter 9, this book

Romanowicz RJ, Osuch M, Grabowiecka M (2013) On the choice of calibration periods and objective functions: a practical guide to model parameter identification. Acta Geophys 61(6):1477–1503

Saltelli A, Tarantola S, Campolongo F, Ratto M (2004) Sensitivity analysis in practice: a guide to assessing scientific models. Wiley, New York

Seibert J (1997) Estimation of parameter uncertainty in the HBV model. Nord Hydrol 28(4/5):247–262

Seibert J (1999) Regionalisation of parameters for a conceptual rainfall runoff model. Agric For Meteorol 98–99:279–293

Seibert J (2003) Reliability of model predictions outside calibration conditions. Nord Hydrol 34:477–492

Shin MJ, Guillaume JHA, Croke BFW, Anthony J, Jakeman AJ (2015a) A review of foundational methods for checking the structural identifiability of models: results for rainfall-runoff. J Hydrol 520:1–16

Shin M-J, Guillaume JHA, Croke BFW, Jakeman AJ (2015b) A review of foundational methods for checking the structural identifiability of models: results for rainfall-runoff. J Hydrol 520:1–16

Sorooshian S, Duan Q, Gupta VK (1993) Calibration of rainfall-runoff models: application of global optimization to the Sacramento soil moisture accounting model. Water Resour Res 29(4):1185–1194. doi:10.1029/92WR02617

Tokarczyk T (2013) Classification of low flow and hydrological drought for a river basin. Acta Geophys 61(2):404–421

Uhlenbrook S, Seibert J, Leibundgut C, Rodhe A (1999) Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. Hydrol Sci J 44(5):779–797

van Werkhoven K, Wagener T, Reed P, Tang Y (2009) Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models. Adv Water Resour 32(8):1154–1169. doi:10.1016/j.advwatres.2009.03.002

Vrugt JA, Gupta HV, Bouten W, Sorooshian S (2003) A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resour Res 39(8):1201. doi:10.1029/2002WR001642

Wagener T, Gupta HV (2005) Model identification for hydrological forecasting under uncertainty. Stoch Environ Res Risk A 19(6):378–387. doi:10.1007/s00477-005-0006-5

Zelelew MB, Alfredsen K (2013) Sensitivity-guided evaluation of the HBV hydrological model parametrization. J Hydroinf 15(3):967–990

Zhan C, Song X, Xia J, Tong C (2013) An efficient integrated approach for global sensitivity analysis of hydrological model parameters. Environ Model Softw 41:39–52

Distributed Modelling of Flow in the Middle Reach of the River Vistula

Krzysztof Kochanek, Emilia Karamuz and Marzena Osuch

Abstract The article aims to present a step-by-step procedure of distributed modelling of river flow by contrasting two commonly used modelling tools: MIKE 11 and HEC-RAS. The middle River Vistula reach constitutes the case study for the 'debate'. Although different in form, level of complexity and price, both models lead to similar results for flood modelling, including the uncertainty of the simulated flow. The slight superiority of the MIKE 11 model stems from the automatic estimation of its parameters.

Keywords The River Vistula · Distributed modelling · MIKE 11 · HEC-RAS

1 Introduction

Floods are the most destructive natural phenomenon in Poland. The immense flood in 1997 triggered a number of central government flood-protection initiatives that included inter alia the modernisation of the monitoring system (Systems of Monitoring and Country Protection; in Polish: Systemy Monitoringu i Osłony Kraju, Lindel et al. 1997; Mierkiewicz et al. 1999; Kadłubowski 2005). This

K. Kochanek (✉) · E. Karamuz · M. Osuch
Institute of Geophysics, Polish Academy of Sciences,
Księcia Janusza 64, 01-452 Warsaw, Poland
e-mail: kochanek@igf.edu.pl

E. Karamuz
e-mail: emilia_karamuz@igf.edu.pl

M. Osuch
e-mail: marz@igf.edu.pl

R.J. Romanowicz and M. Osuch (eds.), *Stochastic Flood Forecasting System*,
GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-319-18854-6_6

system automatically transfers measurements to the operational systems responsible for flood forecasting and protection. Flood forecasts prepared by national operational systems in the Vistula basin are based on HBV rainfall-runoff models and the flow routing hydrodynamic models (Lindel et al. 1997; Kadłubowski et al. 2011). However, the flood incidents in the River Vistula in spring 2010 revealed shortcomings in the national system of flood forecasting and protection. As a result, there is still a growing demand for the development of new flood modelling and decision support tools that will enable the fulfilment of the needs of Polish flood protection services. At the same time, the rapid development of computer technology in numerical distributed modelling now allows us to supplement deterministic flood modelling results by flow forecast uncertainty, indispensable when preparing credible flood warnings. As a result, the distributed models of catchment and river valleys together with uncertainty analysis are now basic tools, apart from classical flood frequency analysis, for hydrological design and flood forecasting (Moussa and Bocquillon 1996). Dozens of commercial and non-commercial versions of modelling software packages compete with each other, mostly in the quality and ease of use of the pre- and post-processors, leaving the mathematical foundation basically the same: the 1D Saint Venant's formulae (Saint Venant 1871) accompanied by the diffusive wave approximation and/or kinematic wave approximation with the Muskingum method for schematised channel routing (Chanson 2004; Price 1974; Szymkiewicz 1991; 1993). Some difference in calculating engine may stem from different numerical approaches, solvers, etc. Of these practical versions of modelling packages, two deserve special consideration: MIKE 11 by the Danish Hydraulic Institute (DHI 2004) and HEC-RAS by the United States Corps of Engineers (Barkau 1993). Both packages are widely known and have been commonly used in hydrological projects globally and both have gained their devoted supporters. The MIKE 11 package is a complex decision support system with a highly developed graphical user interface offering a vast variety of applications in practical hydrology. On the other hand, we have the free-of-charge and similarly advanced HEC-RAS model that is commonly used by practitioners and scientists due to its availability and good performance.

The main aim of this paper is to present, step-by-step, an application of flow routing tools using the River Vistula reach between Zawichost and Warsaw as an example. At the same time, we present and compare the two most popular distributed modelling software packages in Poland, i.e. MIKE 11 and HEC-RAS. In this section of the book we show the subsequent stages of an application of flow routing tools by means of the exemplary case study rather than by focusing on the technical aspects of model application. The latter can be easily performed by following the instruction manuals of both packages. It is also important to note that we put stress on the hydrological realism within the framework of data availability and the time necessary to achieve reasonably good results. Furthermore, we will

try to find an answer to the question as to whether the use of the same (as much as possible) input data to two similar modelling packages can result in similar (ideally equal) flow predictions.

2 Step 1—Problem Description and a Study of the Modelled Reach

The precise description the purpose of the modelling is the first and, perhaps, key stage of the modelling process. Our objective is to build a model for flood forecasting; consequently, all the methodology presented in this paper aims to achieve this goal.

The practice of modelling starts with a study of the river reach to be modelled, the hydrology and hydro-morphological parameters of the river, the land cover, and water management issues. The choice of the modelling tools should depend on the problem to be solved and the availability of observations but also the skills and experience of the modellers. A proper identification of the problem and careful study of the modelled area usually speeds up the whole modelling procedure and eliminates errors that could appear in the later stages of model preparation.

The River Vistula is the largest river in Poland, with a length of 1047 km; its catchment area equals 194,424 km^2 (168,700 km^2 in Poland). Despite a long tradition of hydrological monitoring (over 100 years for some gauging stations), large parts of the Vistula catchment (tributaries) are ungauged or scarcely monitored; the gauging stations along the main river channel are located unevenly (every few or even few dozen kilometres) from −105.5 (the sources) to 941.5 km (the estuary). The Vistula River is of a semi-natural nature—there are no large dams or other hydro-technical facilities in the whole Middle Vistula reach, which is unique in Europe for such a big river. Recently, the industrial, cultural and environmental development of cities and towns along the course of the river resulted in an urgent need for a reliable flood protection policy.

Our research considered the 225-km-long river reach from the town Zawichost (287.6 km of the course) to the northern end of Warsaw (513.3 km)—see the map in Fig. 1. The hydro-morphology of the floodplain is rich in oxbow lakes and creeks. The River Vistula channel is relatively straight and, to some extent, regulated but it sometimes creates large meanders (Zieliński 1999; Sarnacka 1987). There are dikes irregularly located in the main channel. After World War II flood-protecting levees were built along almost the whole modelled reach on both banks of the river; however, the first dikes were built in the 19th century (Łajczak et al. 2006). The levees narrow the valley to 730–2000 m (Kowalska 2010) (Fig. 1).

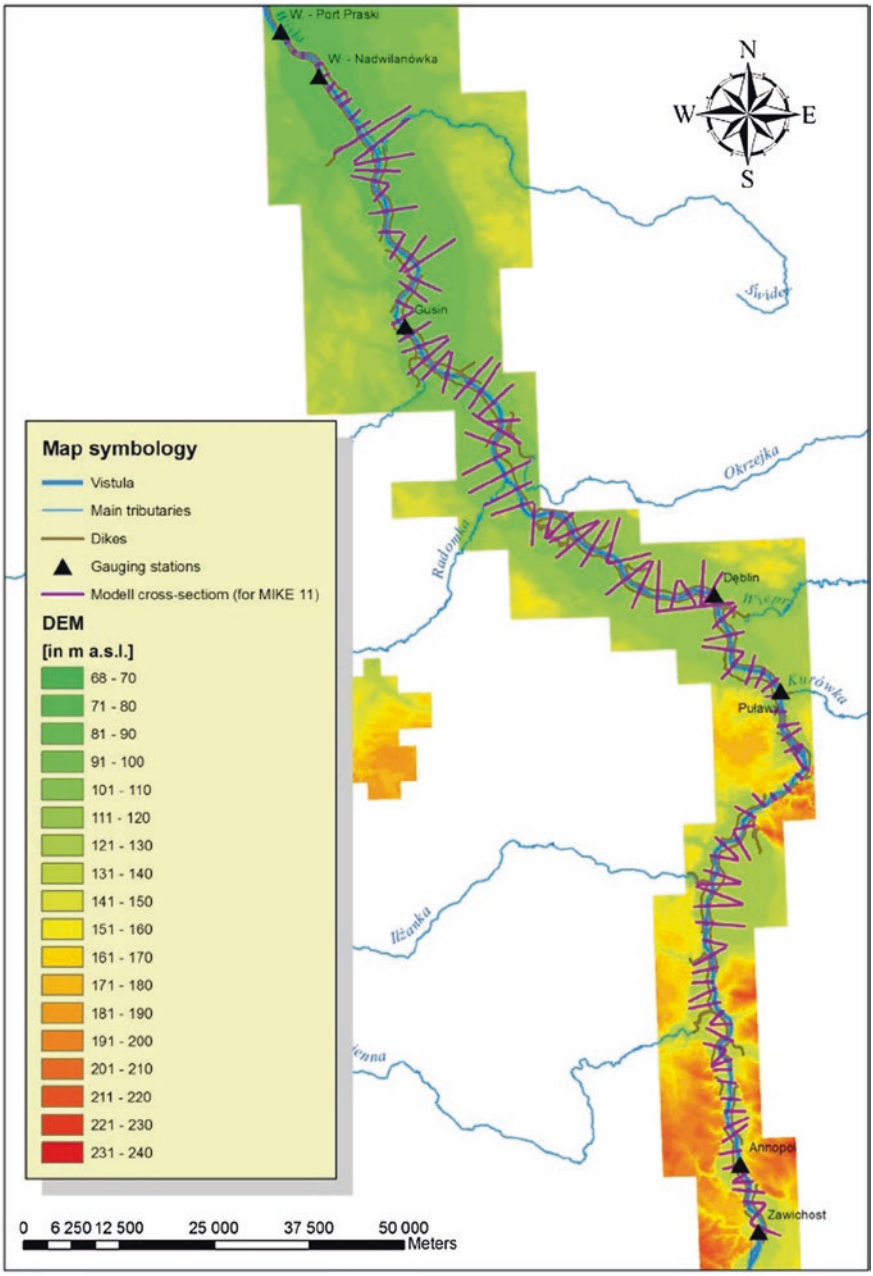

Fig. 1 Map of the modelled reach of the River Vistula

3 Step 2—Choice of a Hydrodynamic Model

The idea of simulating discharge and water levels at different sections of the river reach by flow routing techniques is not new (for example Chanson 2004). Water flow in an open channel can be, without loss of generality, approximated by the Saint Venant equations; however, certain conditions must be fulfilled (Szymkiewicz 2000; DHI 2004):

- the water is incompressible and homogeneous (i.e. negligible variation in density)
- the bottom-slope is small (the cosine of the angle it makes with the horizon may be taken as 1)
- the wave lengths are large compared to the water depth. This ensures that the flow everywhere can be regarded as parallel to the bottom, i.e. vertical accelerations can be neglected and a hydrostatic pressure variation along the vertical can be assumed.
- the flow is subcritical; supercritical flow can be also modelled by MIKE 11; the HEC-RAS algorithm performs also computations in a mixed regime.

Some simple cases of flow routing process can be solved by analytical methods (MacDonald 1996; Szymkiewicz 2000). However, in practice the Saint Venant equations are solved by numerical techniques.

In this article we show the results of flood wave transformation for the River Vistula reach between Zawichost and Warsaw by means of the MIKE 11 and HEC-RAS modelling packages.

3.1 MIKE 11

MIKE 11 HD (HydroDynamic) (DHI 2004) is one of the most widely used hydrodynamic models. It was developed at the Danish Hydraulic Institute for the simulation of water surface profiles and discharge in rivers. The solution of the equations of continuity and momentum is based on an implicit finite difference scheme developed by Abbott and Ionescu (1967). Details of the scheme are given in DHI (2004). The MIKE 11 HD model can be applied to a river of any size and gradient, i.e. from steep streams to tidal estuaries. The model is able to cater for irregular open channels and can account for varying flow resistance through different branches of the network. The flow and water level at the boundaries can be steady or can vary with time. The model has the capability to simulate flow over embankments and through weirs and bridges.

3.2 HEC-RAS

HEC-RAS is a 1-D unsteady, open channel flow routing model, created by the United States Corps of Engineers (Barkau 1993). It is equipped with a user-friendly graphical user interface which provides plots and tables helpful in the pre- and post-processing of models. The software can co-operate with popular GIS packages and, especially, can be used to extract geometrical data from a digital elevation model or to establish the range of floodplains by means of water elevation.

The calculation scheme is based on the Unsteady Network Model (Barkau 1993) developed by the United States Army Corp of Engineers. The HEC-RAS uses the Saint Venant equations solved using the Preissmann scheme. The Saint Venant equations are solved separately for the flow in the channel and for the overbank area. A thorough description of the model can be found in the HEC-RAS River Analysis Hydraulic Reference Manual (Brunner 2010).

4 Step 3—Designing the Digital Model of the Reach Geometry

Experience shows that an accurate determination of cross sections and river network geometry is crucial for the numerical modelling of flood-wave transformation. The variable character of the River Vistula is partly reflected in the variability of the river bed (see characteristic cross-sections in Fig. 2). The cross-sections for both the MIKE 11 and HEC-RAS models were created with the use of data from three sources: (1) in the strict riverbed (bank to bank) the measurements were made by the Department of Operational Oceanography Maritime Institute (DOOMI) in Gdańsk (Brzezińska et al. 2012) and (2) the river floodplain cross-sections were based on a digital terrain model (DTM) provided by the Centre for Geodetic and Cartographic Documentation, and (3) cross-sections in the city of Warsaw were based on measurements by Warsaw University of Technology (Falacinski et al. 2009).

The Department of Operational Oceanography Maritime Institute in Gdańsk performed bathymetric surveys of 100 cross-sections in the Vistula riverbed, perpendicular to the stream, by means of classical techniques using a single beam sonar and hydrographical measurement unit type hybrid (so-called hydroacoustic survey method); in the shallower parts of the river the geodetic methods were used. As a result of combining these techniques, the bottom of the river bed is represented with horizontal and vertical resolutions of less than 0.05 m. Each cross-section, therefore, is mapped by at least dozens, if not hundreds, of bathymetric points. The special character of the Warsaw city reach is commented on and described in detail by Magnuszewski (2015, this book). The bathymetric measurements in Warsaw include 116 cross-sections, of which only 12 and 64 were considered in the MIKE 11 and HEC-RAS models, respectively. The differences in the number of cross sections in MIKE 11 and HEC-RAS stem from the technical limitation of the MIKE's pre-processor. A detailed flow routing model for the city

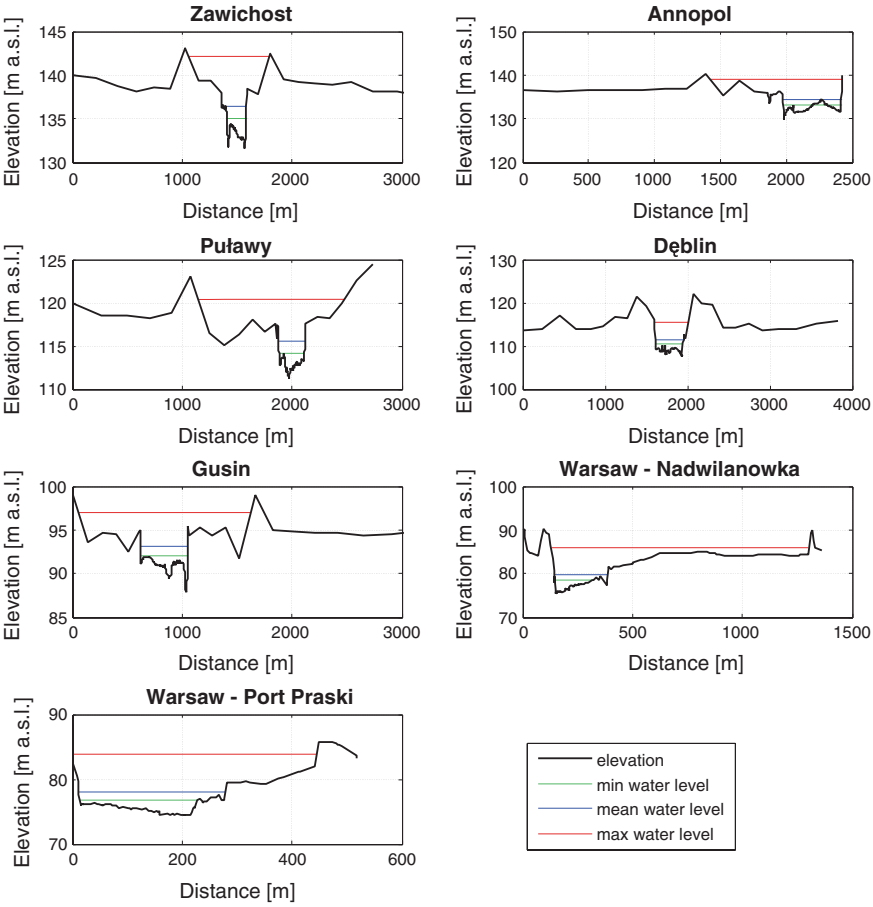

Fig. 2 The reference cross-sections

of Warsaw with all 116 cross-sections and floodplains was the subject of a separate article (Kiczko et al. 2013).

Along the studied reach, sections are located at about 2000 m apart (more frequently in Warsaw). The distances between cross-sections were based on a digital hydrological map of the River Vistula and the coordinates of the measurement points provided by the DOOMI. The shape of the modelled river reach (Fig. 1) was derived from the geometry of the cross-sections, the official digital hydrologic map of Poland (approved by the Polish Ministry of Environment and Main Water Management Office) and digital elevation model of the River Vistula valley. Over the course of the modelled length there are several tributaries, such as Kamienna, Pilica and Wieprz, but our results indicated that their impact on overall flow in the main channel is negligible and the model without their inflow performs sufficiently well. These tributaries do not compete with the main river in terms of discharge, therefore, after testing, they were omitted from the models.

5 Step 4—Boundary and Initial Conditions

Boundary and initial conditions are necessary for the derivation of a solution of the Saint Venant equations. By convention, initial conditions describe the state of the system variables (e.g. water levels and discharges) at time t = 0, whilst the boundary conditions specify the state of the variables at certain spatial locations (e.g., at the beginning and the end of the modelled reach). The MIKE 11 package offers three types of boundary conditions at the downstream- or upstream-end cross-section: (i) constant values of discharge or water level, (ii) time-varying values of discharge or water level, (iii) a functional relationship between discharge and water level (rating curve)—only downstream. Additionally, HEC-RAS also allows us to impose downstream boundary conditions; (iv) normal water level estimated from Manning equation. The lower boundary conditions affect the results in the near location; therefore, they should be used far enough downstream from the study area. The initial conditions are the values of the discharge and water level at a certain time (typically the starting point in time, t = 0) and place (each node) of the modelled system. The influence of initial conditions decreases with time. In both the MIKE 11 and HEC-RAS modelling packages the initial conditions can be established by the user or automatically by the software.

Regardless of the choice of the types of boundary conditions they should be based on high quality historical information on flows and/or water levels. The choice of boundary condition depends on the purpose of the modelling and the availability of data. In our study, the basic database consists of daily data for water levels and discharges for the years 1950–2010 measured at 7 gauging sections from Zawichost to Warsaw (Table 1).

Table 1 clearly shows that, however relatively long, the sequences of both flow measurement and the conditions do not always include the same measurement period (compare also Fig. 3 and 4).

The historical datasets for the years 1950–2010 are complete for almost all gauges (Zawichost, Annopol, Puławy, Dęblin, Gusin and Warsaw–Nadwilanówka). For the Warsaw–Port Praski profile only water level measurements were collected

Table 1 Historical datasets used for modelling

Gauging station	Gauging reference levels (m a.s.l.)	km	Discharge	Water level
Zawichost	133.38	287.6	1-11-1950–31-10-2010	1-11-1971–31-10-2010
Annopol	131.38	298.4	1-11-1950–31-10-2010	1-11-1950–31-10-2010
Puławy	113.92	374.9	1-11-1950–31-10-2003	1-11-1969–31-10-2003
Dęblin	109.15	393.7	1-11-1970–31-10-2010	1-11-1969–31-10-2010
Gusin	91.74	461.5	1-11-1974–31-10-2010	1-11-1974–31-10-2010
Warsaw–Nadwilanówka	78.79	504.1	1-11-1967–31-10-2010	1-11-1967–31-10-2010
Warsaw–Port Praski	76.08	513.3	1-11-1950–31-10-1967	1-11-1980–31-10-2010

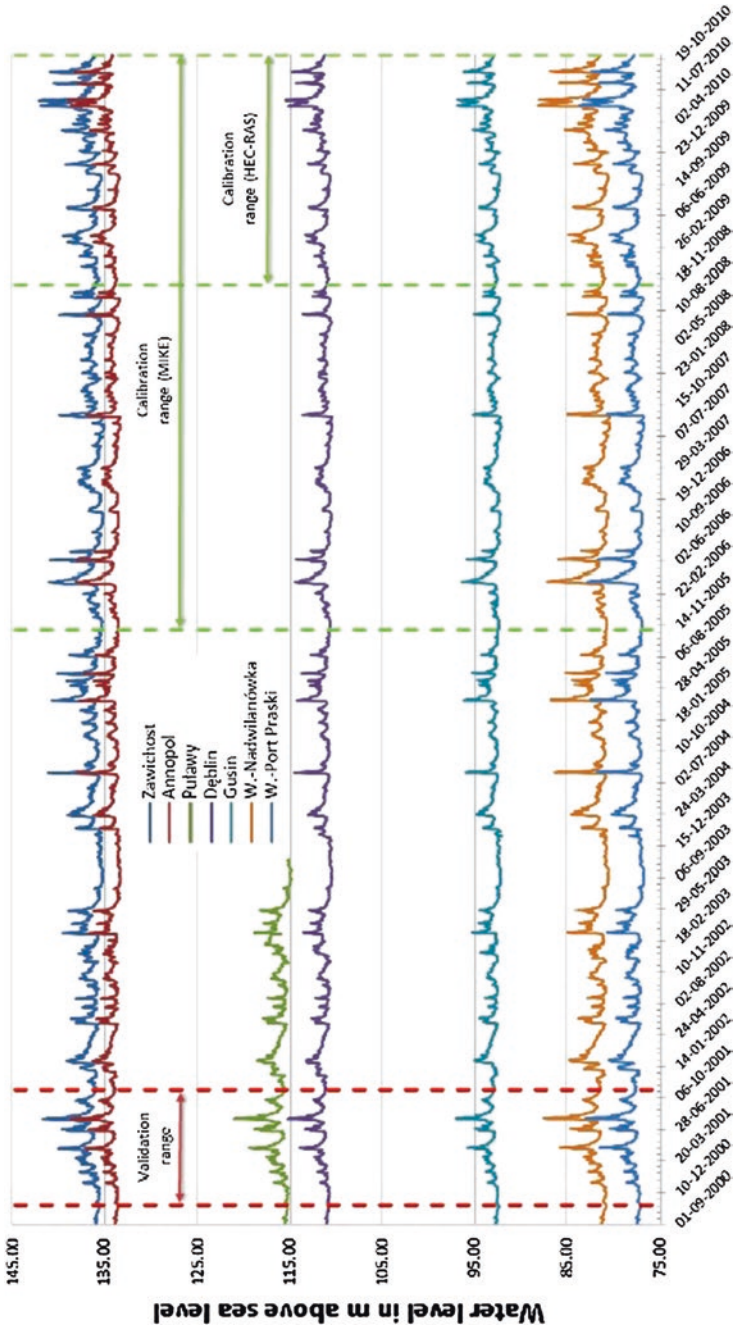

Fig. 3 Water levels (in meters above the sea level) for seven gauging stations. The dates encompass both calibration and validation periods

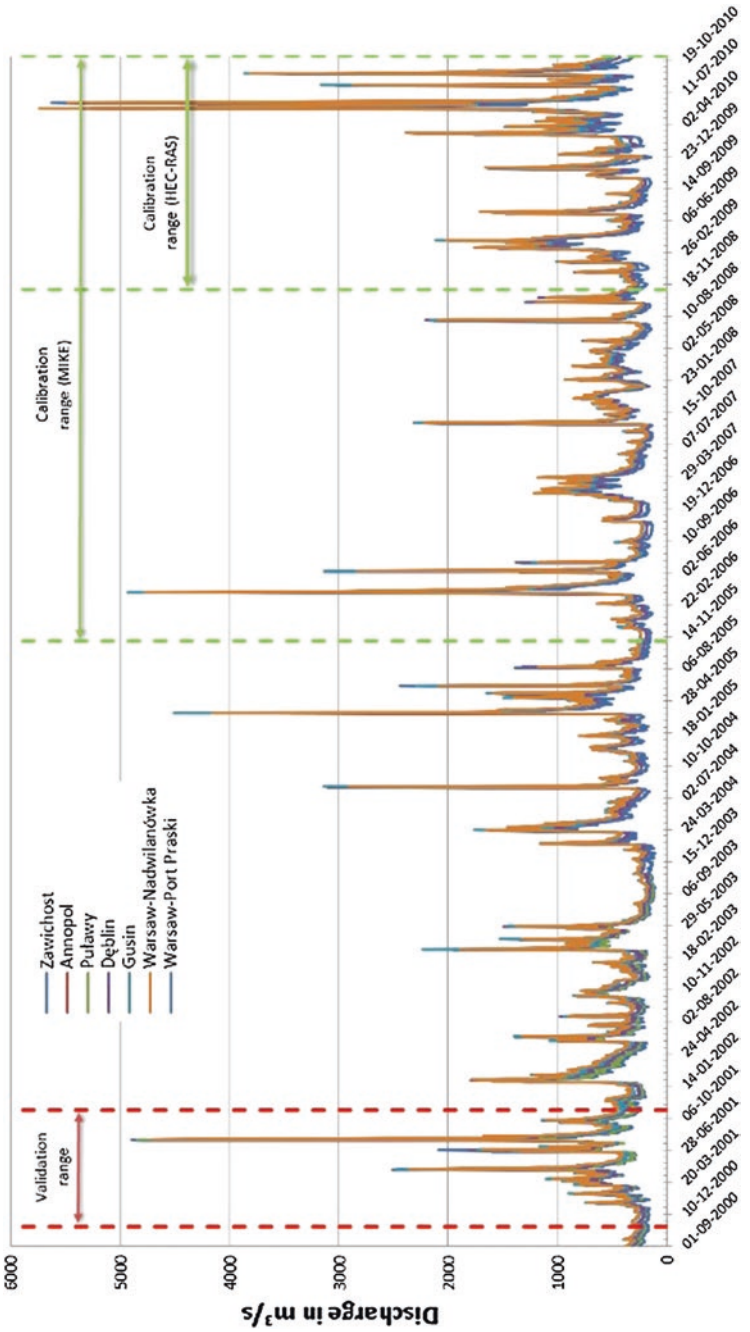

Fig. 4 Discharge for seven gauging stations. The dates encompass both calibration and validation periods

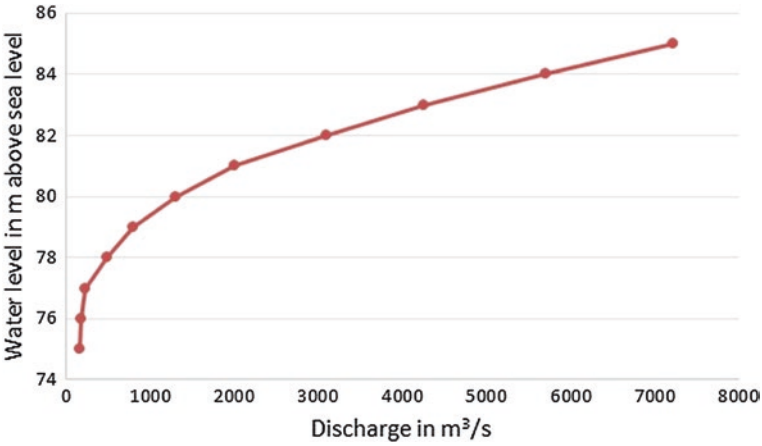

Fig. 5 The rating curve for Warsaw–Port Praski gauge as the lower boundary condition for MIKE 11

in this period. Note also that Puławy gauging station was moved from chainage 372.5 to 374.9 km in 2003. Similarly, the Warsaw–Nadwilanówka gauging station was moved from 504.2 to 504.1 km in 1994, which resulted in a raising of the reference levels by 0.716 and 2.134 m, respectively (compare Fig. 3). Nevertheless, the relatively short distance between Warsaw–Nadwilanówka and Warsaw–Port Praski allows us to transfer water flows from the profile Nadwilanówka to Port Praski in order to derive the rating curve (Fig. 5). Puławy gauging station, due to its proximity to Deblin and short datasets played only supportive role in the process of calibration and verification of the models. This discharge-water level relationship is then applied as the lower boundary condition for the MIKE 11 model, whereas for HEC-RAS the normal-depth boundary condition was used. The flow hydrograph (Zawichost gauging station) determines the upper boundary condition for both models.

6 Step 5—Estimation of Model Parameters

The estimation of model parameters consists of the minimisation of differences between observations and model simulations. We focused on the Manning's n friction coefficient as the main parameter to be identified for both the MIKE 11 and HEC-RAS models.

The number of parameters that may have influenced the output of the models is much larger, such as lateral inflows, wind strength, bridges and structures. However, bearing in mind the simplicity principle (the simpler the model, the better) and due to the shortage of information on the values of these parameters,

they were omitted. However, it is worth bearing in mind that the parameters of the numerical finite difference scheme of the modelling engine should be set at the values that guarantee a high level of stability of the model's calculation scheme. The spatial and temporal steps of finite difference approximation of the flow equations should be sufficiently small, but on the other hand, they should allow us to obtain acceptable results within a reasonable time of calculations, especially when the model is meant to be a part of a continuous modelling or flood forecasting tool. In both models the spatial step was chosen taking into account the adequate Courant number that guarantee the accuracy of the numerical scheme. According to Popescu (2014), the Abbott-Ionescu scheme works properly with Courant number (Cr) smaller than 10. The study presented in Cunge et al. (1980) shows that for the Preissman and Abbott-Ionescu schemes Courant number equal to 25 is the highest value leading to accurate results. In our case the requirement of $Cr < 10$ was equivalent to $\Delta t \approx 30$ min which was used in the study.

6.1 MIKE 11

The actual process of identification starts with a sensitivity analysis which provides an answer to the issue of model calibration strategy. The results also inform us about the dependence of model parameters between profiles, the possible range of the identified parameters and their mutual influence in particular cross-sections. From among the range of methods used to analyse the sensitivity, a few seem especially suitable within the context of flood modelling, namely: the method of Morris, the method of Sobol, visual analysis and the correlation method. All four were carried out for the MIKE 11 model of the River Vistula; however, for brevity, only the final conclusions of the Morris' method will be presented, whereas a detailed analysis with diagrams can be found in previous chapter (Osuch 2015; this issue).

The sensitivity analysis for the MIKE 11 Vistula model considers the influence of Manning's n in seven subsequent referencing gauging stations (Zawichost, Annopol, Dęblin, Gusin, Warsaw–Nadwilanówka and Warsaw–Port Praski) on the Nash-Sutcliffe coefficients estimated in seven profiles. The following assumptions have been made: number of parameters = 7, number of sensitivity-related levels = 10, total number of trajectories = 1000 and optimal number of trajectories = 15; the uniform distribution function served as the parent distribution function. The Morris' indices (μ^* and σ) serve as criteria of benchmarks of direct and overall influence of the parameter for a particular gauging station on the other referencing cross sections.

The main conclusion resulting from the sensitivity of the model response to the variation of the Manning's n is that the goodness of fit at controlled cross sections is under influence of more than one model parameter. This effect is more spectacular in forwarding cross-sections where almost all parameters have significant influence on model output.

The optimisation of one global value of the Manning's n for the whole modelled reach of the river did not bring a satisfactory result. The considerable decrease in energy line and geomorphological differences between following cross-sections, but also the data scarcity, influenced the stability of the results and the errors practically disqualified a global approach to model identification. Therefore, the river's reach was split between the gauging stations (Zawichost, Annopol, Dęblin, Gusin, Warsaw–Nadwilanówka, Warsaw–Port Praski) and roughness coefficients were estimated for each profile. In the remaining cross-sections the Manning's n coefficients were linearly interpolated between the controlling cross-sections (MIKE 11).

Furthermore, in each of the seven controlling cross-sections the Manning's n coefficients were optimized to achieve the optimal solution, i.e. to minimise the difference between the model results and measurements. As a goodness-of-fit criterion the Nash-Sutcliffe multiplied by -1 was selected. At the same time the values of the optimised coefficient were kept within the range characteristic of the Vistula channel (0.005–0.2).

The calibration dataset period covers the years 2005–2010 (five hydrological years) for MIKE 11 and 2008–2010 (two hydrological years) for HEC-RAS. The periods were selected in terms of the stability of the rating curves and river geometry. These periods are possibly closest to the period when the cross section measurements were made. The large variability in flow that occurred during these years permits a proper model calibration. However, the chosen objective function results in a better simulation of mean and high water levels and flows (Gupta et al. 2009).

The validation was performed for measurements (water levels and flows) from 01.11.2000 to 31.10.2001 (1 hydrological year). The validation period includes one of the highest flood waves of the last decade.

The optimisation of Manning's n parameter in MIKE 11 was done automatically, with regard to the observed time series. Model calibration was carried out using the Simplex Nealder-Mead algorithm (Press et al. 2002) within selected feasible parameter ranges of the Manning coefficient from 0.005 to 0.2. Model performance was evaluated by the Nash-Sutcliffe criterion (Nash and Sutcliffe 1970) set for water levels and discharges for Zawichost, Annopol, Dęblin, Gusin, Warsaw–Nadwilanówka and Warsaw–Port Praski. The estimation based either on water levels or discharges led to erratic results. The results of sensitivity analysis indicated that the model fit at particular profiles strongly depends on Manning coefficients not only from the neighbouring cross-sections; therefore, the objective function is the sum of Nash-Sutcliffe for referencing cross sections for both stage and flow (except for Warsaw–Port Praski). The Nash-Sutcliffe coefficient can vary from $-\infty$ to 1, so in our case the maximum value of the objective function can reach 11.

The local method of optimisation such as the Simplex Nealder-Mead algorithm requires multiple starting points of the algorithm in order to find the global solution. We tried 10 starting points, the best of which achieved a value of objective points of 9.03 (starting point nr 6). The results of the optimisation are presented in Table 2. The optimal Manning's n is predominantly equivalent to weedy, deep

Table 2 The results of identification of Manning's n coefficient for 10 optimisation algorithm starting points for MIKE 11

Starting point no.	Optimal Manning's n						Objective function
	Zawichost	Annopol	Dęblin	Gusin	Warsaw–Nadwilanówka	Warsaw–Port Praski	
1	0.0496	0.0487	0.0050	0.0459	0.0770	0.0500	9.0277
2	0.0494	0.0485	0.0077	0.0450	0.0783	0.0116	8.9483
3	0.0498	0.0489	0.0051	0.0457	0.0766	0.0483	9.0256
4	0.0474	0.0486	0.0116	0.0480	0.0755	0.0249	8.9337
5	0.0501	0.0473	0.0080	0.0466	0.0823	0.0151	8.9260
6	0.0499	0.0484	0.0050	0.0459	0.0761	0.0537	9.0298
7	0.0696	0.0478	0.0431	0.0497	0.0742	0.0066	8.1919
8	0.0525	0.0394	0.1987	0.0433	0.0710	0.0510	7.3167
9	0.0494	0.0499	0.0050	0.0457	0.0767	0.0405	9.0150
10	0.0494	0.0485	0.0077	0.0450	0.0783	0.0116	8.9483
Mean:	0.0517	0.0476	0.0297	0.0461	0.0766	0.0313	
Std. dev.:	0.0064	0.0029	0.0605	0.0018	0.0029	0.0192	

pools (Chow 1959, p. 110), which is close to reality, since within its modelled reach the River Vistula is a partly regulated, weedy and relatively deep channel. However, in two cases the optimal values of the Manning's n are much below or above the expected results, namely Dęblin (0.005) and Warsaw–Nadwilanówka (0.076). Being aware of this, we have to accept these unreal values because no other set Manning's n coefficients provides such a good fit between the simulations and the measurement data.

The MIKE 11 software provides the possibility of the division of cross-sections into three zones of different friction coefficients regarding water level. However, the optimisation of three Manning's coefficients for each referencing cross section (instead of one) brings only small (negligible) improvement of the value of objective function but, at the same time, significantly complicates the calibration of the model.

Table 2 and Figs. 6 and 7 show that the calibration of the model is successful for all but the last gauging stations. The discrepancy for the Warsaw–Port Praski gauging station results from the errors of lower boundary condition (rating curve). However, one has to bear in mind that the purpose of this model is to provide relatively good flow predictions for the Warsaw–Nadwilanówka gauging station. The model of flood transformation within the city of Warsaw was also prepared with the use of more detailed data, but its description is beyond the scope of this chapter. Therefore, in Table 3 and later as well as in figures the results of the Warsaw–Port Praski were omitted as non-relevant.

The robustness of the MIKE 11 model predictions within a range of discharges is confirmed by the validation results (Table 4, Figs. 10 and 11).

In the case of both the calibration and verification, one can discern a general rule: the longer the distance modelled, the higher level of error the model

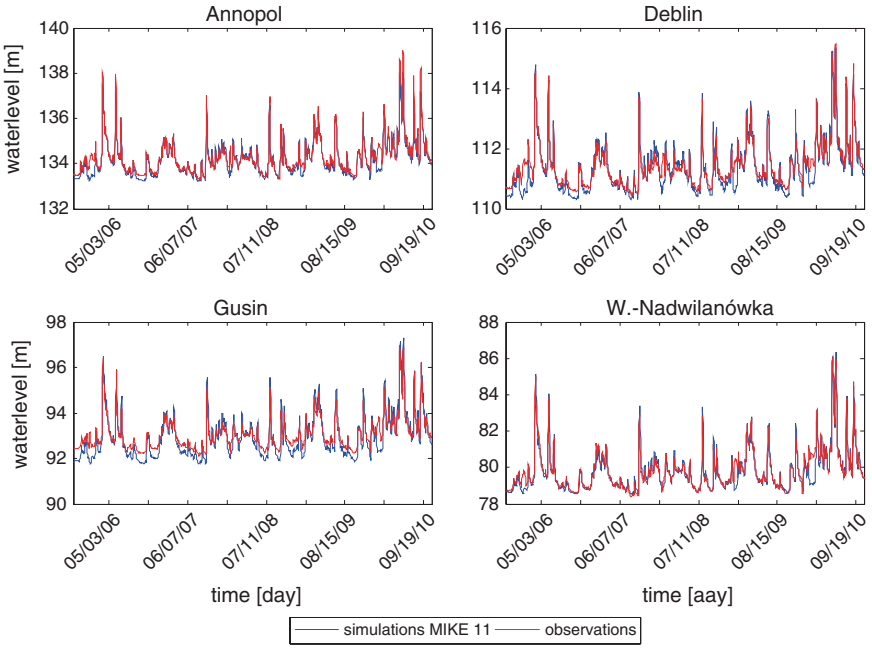

Fig. 6 The results of the MIKE 11 model calibration—water level

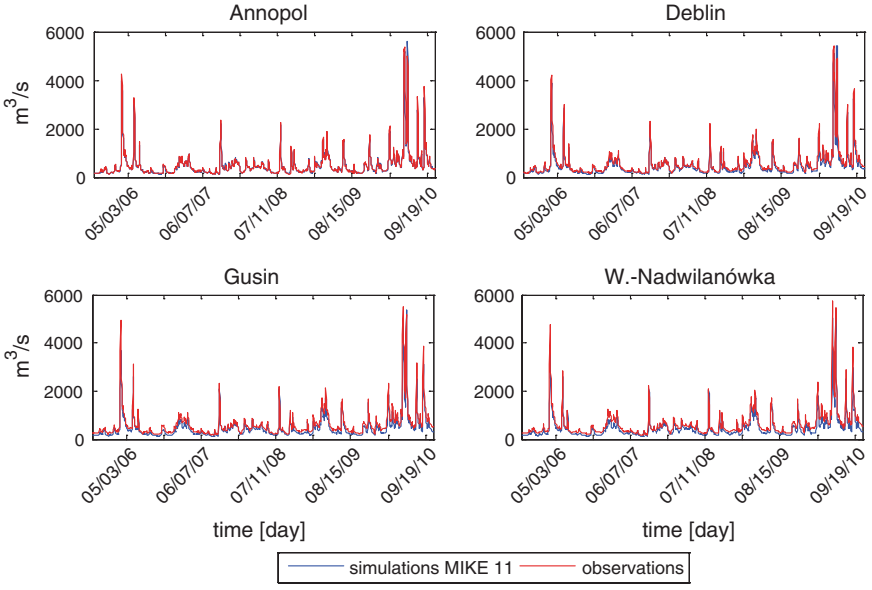

Fig. 7 The results of the MIKE 11 model calibration—discharge

Table 3 The optimal values of Nash-Sutcliffe criterion for the identification of roughness coefficients using reference gauging stations for MIKE 11

Reference gauge	Water level	Discharge
Zawichost	0.9336	0.9999
Annopol	0.9129	0.9793
Dęblin	0.8605	0.9557
Gusin	0.7513	0.8936
Warsaw–Nadwilanówka	0.8859	0.8559

Table 4 The values of Nash-Sutcliffe criterion for the validation for reference gauging stations for MIKE 11

Reference gauge	Water level	Discharge
Zawichost	0.9717	0.9998
Annopol	0.9674	0.9935
Dęblin	0.9363	0.9631
Gusin	0.8024	0.9117
Warsaw–Nadwilanówka	0.9224	0.8598

produces. This is natural, because the errors that occur in the subsequent cross sections are accumulated over the course of the river. One can easily notice a particularly good adjustment of the model to the highest discharges/water levels (even for the Warsaw gauges) which positively authorises the model's robustness for flow forecasting.

6.2 HEC-RAS

In contrast to the MIKE 11 model, the calibration of the HEC-RAS model was done by the manual iterative adjustment of calculation results to the measurement data. All the assumptions for MIKE 11 relating to the geometry of the model, the initial conditions, the parameters of the model and the numerical methods hold also for the HEC-RAS, except the lower boundary conditions which in HEC-RAS is set as the normal depth.

Although slightly different from the MIKE 11 parameters, the optimal Manning's n values (Table 5) fit into the range of sluggish reaches, weedy, deep pools (Chow 1959, p. 110) characteristic of the Vistula River in the modelled reach. Also here there are values that fall beyond the expected Manning's n coefficients, but they provide the optimal results of the model.

Similarly to the MIKE 11 case, the goodness of fit of the model to the measurement data in referencing cross-sections was expressed in the form of Nash-Sutcliffe coefficients (Table 6).

Perhaps due to the manual approach of the calibration and thus the limited number of iterations, the HEC-RAS calibration results reveal a slightly lower level of goodness-to-fit than those obtained for MIKE 11 (compare Figs. 8 and 9). One has to bear in mind, however, that automatic calibration of HEC-RAS model,

Table 5 The values of optimal Manning's n for the HEC-RAS model optimal Manning's n

	Zawichost	Annopol	Dęblin	Gusin	Warsaw–Nadwilanówka
Manning coefficient	0.0554	0.0794	0.0499	0.1012	0.0583

Table 6 The values of Nash-Sutcliffe criterion for the reference gauging stations obtained in calibration stage

Reference gauge	Water level	Discharge
Zawichost	1	0.8433
Annopol	0.9082	0.8628
Dęblin	0.8519	0.8704
Gusin	0.8825	0.8804
Warsaw–Nadwilanówka	0.9116	0.8663

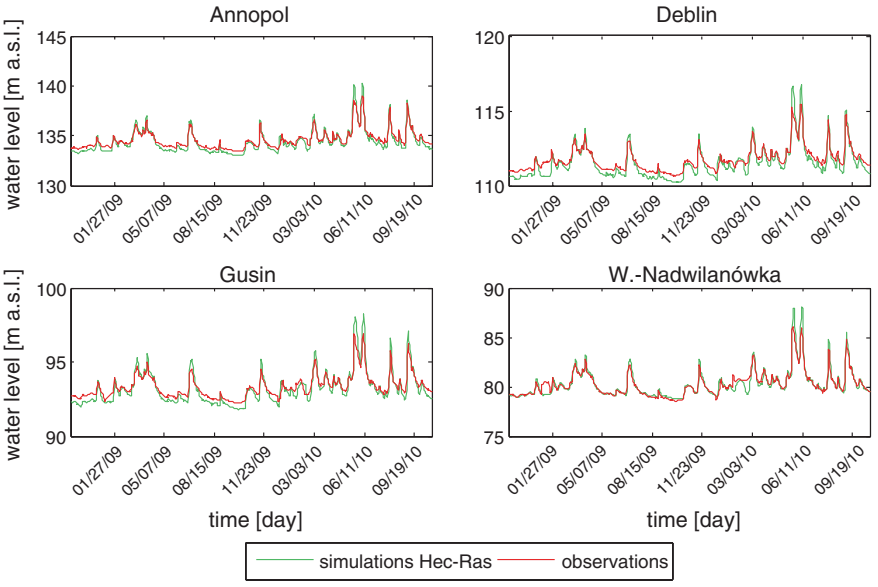

Fig. 8 The results of the HEC-RAS model calibration—water level

although possible, due to the specific form of input and output files requires technically more advanced software than for the MIKE 11 package. Nevertheless, a sufficiently good adjustment of the HEC-RAS model to the data is confirmed by the verification tests (Table 7, Figs. 10 and 11).

An analysis of the relationship between the simulated and observed values gives additional interesting insights into the validation results. Figures 12 and 13 present the H-H plots (upper rows) for MIKE 11 and HEC-RAS, respectively.

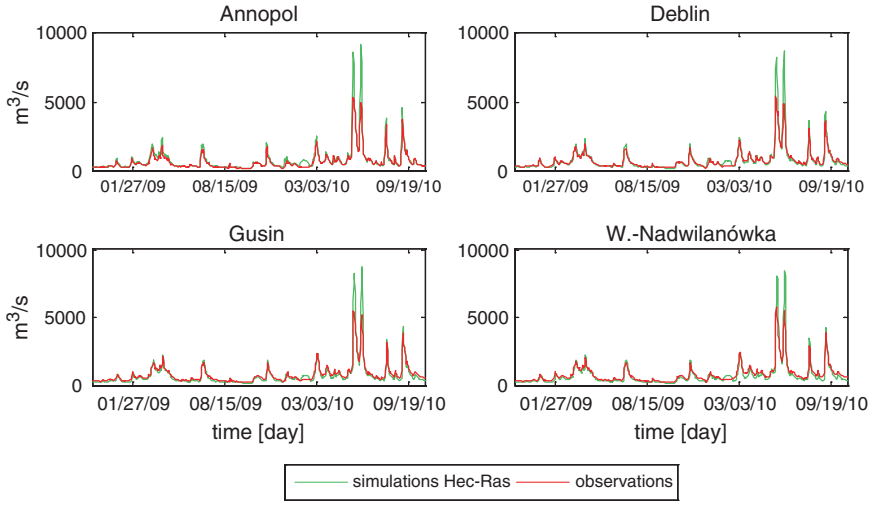

Fig. 9 The results of the HEC-RAS model calibration—discharge

Table 7 The values of Nash-Sutcliffe function for the verification for reference gauging stations

Reference gauge	Water level	Discharge
Zawichost	1	0.8609
Annopol	0.9309	0.8725
Dęblin	0.8771	0.8361
Gusin	0.8681	0.8407
Warsaw–Nadwilanówka	0.9240	0.8234

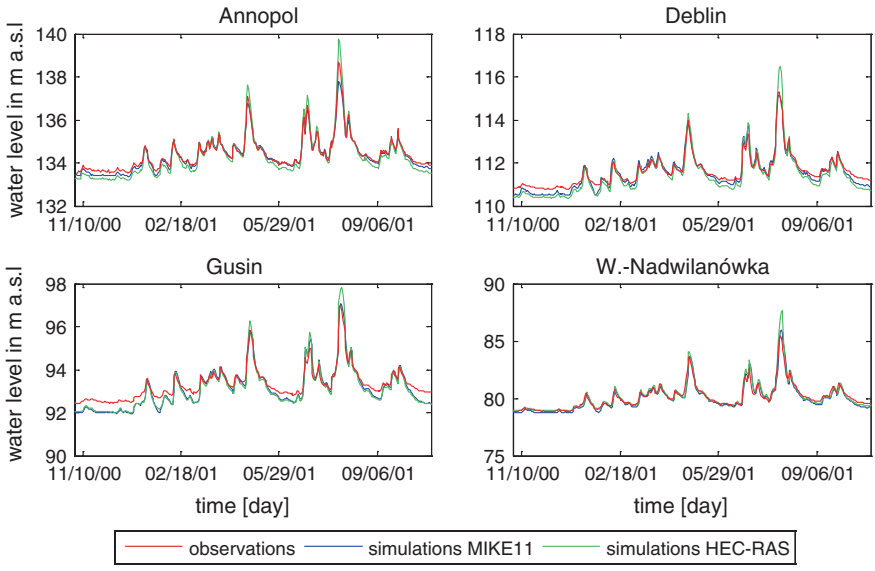

Fig. 10 The results of the MIKE 11 and HEC-RAS models validation—water level

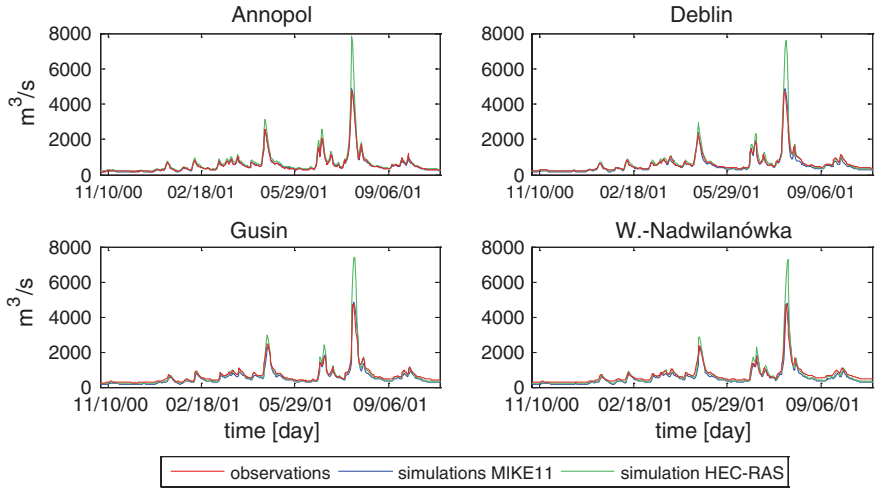

Fig. 11 The results of the MIKE 11 and HEC-RAS models validation—discharge

The MIKE 11 results show a systematic bias of the simulated results—the simulated water level is almost always higher than the observed measurements. What is interesting is that the bias is fairly constant, so it can easily be considered when analysing the water level calculations. This bias is also noticeable in HEC-RAS, but, additionally, the dissemination of points in the upper water stage records is higher than in the case of MIKE 11, which makes the de-biasing procedure more complex, if not impossible.

Figures 12 and 13 also show the Q-Q plots (lower rows) for MIKE 11 and HEC-RAS, respectively. One can easily see that the simulation by HEC-RAS provides lower-quality results than MIKE 11, especially for the controlling gauges down the river and higher flows. This is so, because the HEC-RAS model, unlike the MIKE, was calibrated manually. The results represent thus the compromise between the accuracy of the modelling and the time of calibration. However, if the automatic calibration of the HEC-RAS was possible, the results might have been as good as in the MIKE. The results of both models, technically, confirm the conclusions stemming from Fig. 11.

7 Step 6—Uncertainty Analysis

In both models the Generalised Likelihood Uncertainty Estimation (GLUE) technique (Beven and Binley 1992) was used for the estimation of parametric uncertainty. This technique is based on multiple Monte Carlo runs of a deterministic model with parameters chosen randomly from a specified an a priori distribution. The posterior distribution of the model predictions is obtained using a likelihood

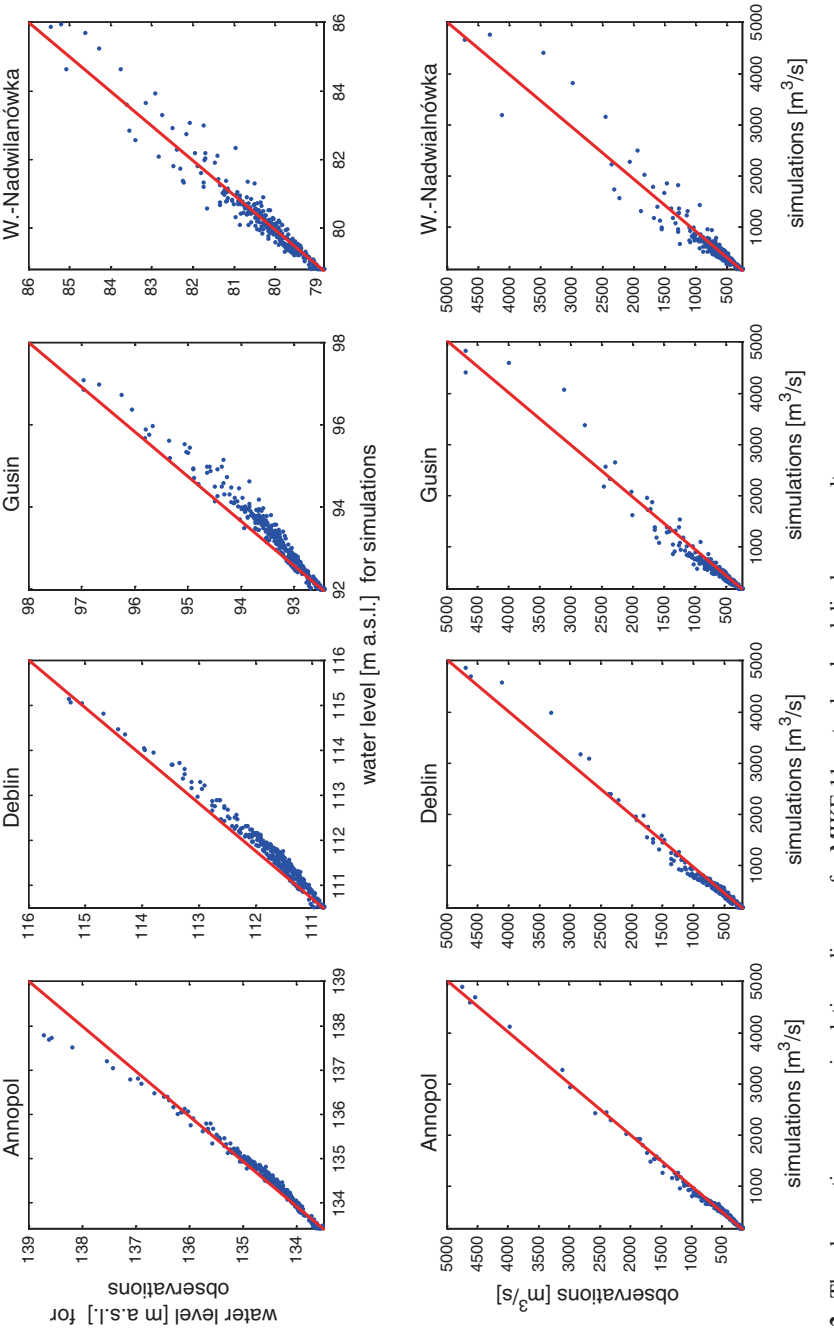

Fig. 12 The observations versus simulations diagram for MIKE 11 water level and discharge results

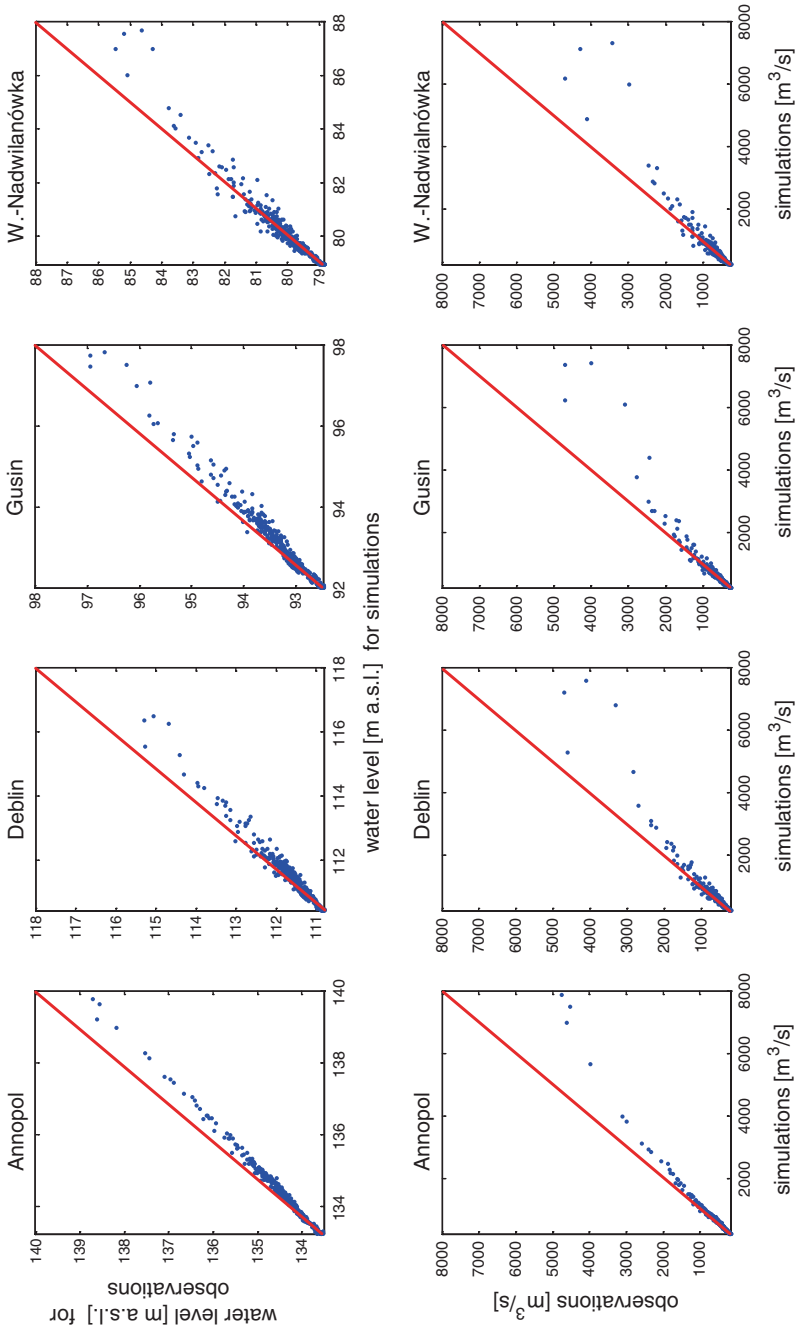

Fig. 13 The observations versus simulations diagram for HEC-RAS water level and discharge results

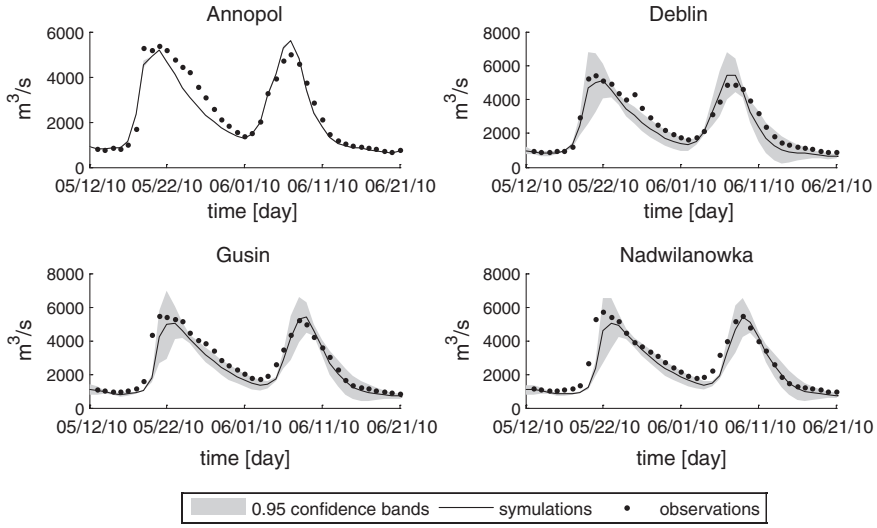

Fig. 14 The calibration uncertainty of discharge by MIKE 11 model

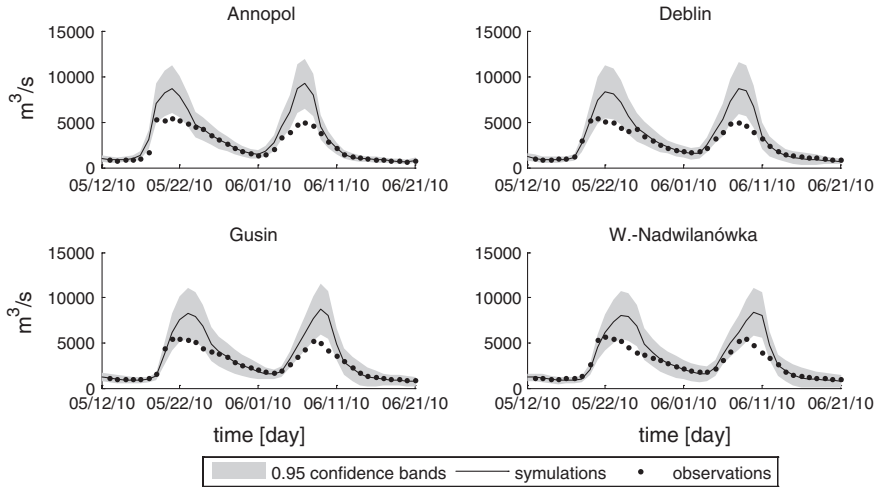

Fig. 15 The calibration uncertainty of discharge by HEC-RAS model

measure conditioned on observations. The choice of the likelihood measure should reflect the purpose of the study. Here the Nash-Sutcliff coefficient of determination was used as the likelihood measure. The results of the uncertainty analysis for the biggest recorded peak flow, where the uncertainty is most clearly visible, are presented in Fig. 14 for MIKE 11 and Fig. 15 for HEC-RAS.

Due to the fact that the number of Monte Carlo simulations of uncertainties and the range of the sampled Manning n's were different, the comparison the uncertainty results of both models is difficult. MIKE 11 was sampled 5000 times from uniform distribution [0.005 0.2] whilst HEC-RAS only 500 times from [0.005 0.5].

8 Step 7—Analysis of Results and Recommendations for the End Users

The aim of the chapter was to describe of the procedure of development of distributed flow routing model for middle River Vistula using two modelling packages: MIKE 11 and HEC-RAS.

The two modelling software packages, however different in form and capacity, eventually lead to quite similar results in the values of the objective values at calibration and validation stages. Better solutions for MIKE 11 result from the fact that the calibration of its parameters was carried out by means of automatic iterative optimisation—very difficult to repeat in case of HEC-RAS model. On the other hand, the time invested in preparation of the suitable optimisation software, testing and optimisation per se is comparable to the manual parameters identification in HEC-RAS. So before building a new distributed river model, one should not only make the choice between the available modelling tools, but also make a trade-off between challenges of calibration and expected (and accepted) results. This is particularly important when the model is supposed to serve as the flood-prediction tool and/or is going to be used in continuous modelling. The Nash-Sutcliffe coefficient is sensitive to the averaging; large flow values and low flows are less involved (Gupta et al. 2009; Osuch et al. 2015). Therefore despite very good values of the Nash-Sutcliffe coefficient, the results of validation show that both models reveal an almost perfect adjustment to the moderate water stage and discharge and are visibly worse in the peak flows (especially the HEC-RAS). The calibration of such a wide range of discharges in such a large river as the Vistula (e.g. 175–5740 m^3/s for Warsaw-Nadwilanówka within the period of calibration) is particularly difficult within the context of data scarcity and the necessary model generalisation. All objective functions tend to satisfy either the peak flows or lower flows. In consequence, a trivial conclusion may be drawn that no model can be regarded as a universal tool and that each hydrological model ought to be comprehensively designed to satisfy particular needs, as e.g. either flood modelling or low-flow-modelling, but never both at the same time. In other words: a model whose purpose is flood prediction (as in our case) should be calibrated to satisfy 'high water' at the expense of low-flows.

As to the uncertainty of the modelling results, it is important to note that, in general, the uncertainty of discharge is rather high in both models (slightly better for MIKE 11). This is unfortunate, because discharge constitutes the basic information for hydrological design. Better results, with regard to their uncertainty, for water levels give hope also for better estimations of discharge, if trustworthy rating curves are available.

In summary, both MIKE 11 and HEC-RAS give similar results, if the same input to the models is delivered. The dissimilarities stem from the technical differences and abilities of both software packages. The automatic estimation of model parameters leads to better results than the manual one.

Acknowledgment This work was financed by the project "Stochastic flood forecasting system (The River Vistula reach from Zawichost to Warsaw)" carried out by the IGF PAN for the National Science Centre (contract no. 2011/01/B/ST10/06866). The flow data were provided by the Institute of Meteorology and Water Management (IMGW), Poland.

References

Abbott MB, Ionescu F (1967) On the numerical computation of nearly-horizontal flows. J Hydraul Res 5:97–117

Barkau RL (1993) UNET one-dimensional flow through a full network of open channels user's manual version 2.1, publication CPD-66. US Army Corps of Engineers Davis CA, Hydrologic Engineering Center, Washington, DC

Beven K, Binley A (1992) The future of distributed models: model calibration and uncertainty prediction. Hydrol Process 6:279–298. doi:10.1002/hyp.3360060305

Brunner GW (2010) HEC-RAS, river analysis system hydraulic reference manual. US Army Corps of Engineers Hydrologic Engineering Center (HEC), 609 Second Street Davis, Ca 95616-4687

Brzezińska A, Cichowska D, Gajewski P, Koszałka J, Kubacka M, Nowak J, Olszewski T, Rudowski S, Wnuk K, Zelewska I (2012) Przeprowadzenie pomiaru korytowego 100 przekrojów poprzecznych mokrych Wisły na odcinku Zawichost-Warszawa (km 494), Raport z pomiarów terenowych, Zakład Oceanografii Operacyjnej, Instytut Morski w Gdańsku

Chanson H (2004) Environmental hydraulics of open channel flows. Butrerworth—Heinemonn, Oxford, p 488

Chow VT (1959) Open-channel hydraulics. McGraw-Hill, New York, p 680

Cunge JA, Holly FM, Verwey A (1980) Practical aspects of computational river hydraulica. Pitman Advanced Publishing Program, Boston

Danish Hydraulic Institute (DHI) (2004) MIKE 11 GIS reference and user manual. Horsholm, Denmark

Falacinski P, Kuzniar P, Wilk E, Danilewicz M, Klimiuk W (2009) Changes in the vertical water level and morphology of the bottom of the river channel in Warsaw reach of the Vistula River—materials in the light of archives and measurements from year 2008. Technical report, Warsaw University of Technology, Faculty of Environmental Engineering

Gupta HV, Kling H, Yilmaz KK, Martinez GF (2009) Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. J Hydrol 377(2009):80–91

Kadłubowski A (2005) Szkoła Tematyczna MANHAZ, Zarządzanie Zagrożeniami dla Zdrowia i Środowiska, 26–30 Sept 2005

Kadłubowski A, Mierkiewicz M, Budzyńska H (2011) Operational rainfall/snowmelt-runoff model for upper Narew River. In: Modelling of hydrological processes in the Narew catchment, geoplanet: earth and planetary sciences. Springer, Berlin

Kiczko A, Romanowicz RJ, Osuch M, Karamuz E (2013) Maximising the usefulness of flood risk assessment for the River Vistula in Warsaw. Nat Hazards Earth Syst Sci 13:3443–3455. doi:10.5194/nhess-13-3443-2013

Kowalska A (2010) Wpływ obwałowań na zróżnicowanie roślinności równiny zalewowej w dolinie środkowej Wisły, Nizina Mazowiecka. (Embankments influence on diversity of floodplain vegetation in the Middle Vistula River Valley, Mazovian Plain). Prace Komisji Krajobrazu Kulturowego Nr 13

Lindel S, Ericsson L, Mierkiewicz M, Kadłubowski A (1997) Integrated hydrological monitoring and forecasting system for Vistula River basin. Final report, IMGW (Poland), SMHI (Sweden)

Łajczak A, Plit J, Soja R, Starkel L, Warowna J (2006) Changes of the Vistula River channel and floodplain in the last 200 years. Geographia Polonica 79(2):65–87

MacDonald I (1996) Analysis and computation of steady open channel flow thesis. Department of Mathematics, University of Reading, UK

Magnuszewski A (2015) Influence of channel processes and vegetation in the Vistula River valley in Warsaw on flood safety, this book

Mierkiewicz M, Kadłubowski A, Sasim M (1999) System of hydrological protecting in the middle and lower Vistula River basin. Wiadomości IMGW XXXII(XLII):4 (in Polish)

Moussa R, Bocquillon C (1996) Criteria for the choice of flood routing methods in natural channels. J Hydrol 186(1–4):1–30

Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I—a discussion of principles. J Hydrol 10(3):282–290. doi:10.1016/0022-1694(70)90255-6

Osuch M (2015) Sensitivity analysis of the flow routing model—multimethod approach, this book, Chap. 7

Osuch M, Romanowicz RJ, Booij M (2015) The influence of parametric uncertainty on the relationships between HBV model parameters and climatic characteristics. Hydrol Sci J. doi:10.1080/02626667.2014.967694

Popescu I (2014) Computational hydraulics: numerical methods and modelling. IWA Publishing, London

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2002) Numerical recipes in C++. Cambridge University Press, Cambridge

Price RIC (1974) Comparison of four numerical methods for flood routing. ASCE J Hydrani Div 100(HY7):879–899

Saint-Venant B (1871) Théorie du mouvement non permanent des eaux, avec application aux crues des rivières et à l'introduction des marées dans leurs lits. Comptes Rendus des séances de l'Académie des Sci 73(1871):237–240

Sarnacka Z (1987) Evolution of the Vistula valley between the outlets of Radomka and Świder in the Late Glacial and Holocene. In: Starkel L (ed) Evolution of the Vistula River valley during the last 15,000 years part II Prace Geograficzne. Special issue nr 4

Szymkiewicz R (1991) Finite-element method for the solution of the Saint Venant equations in an open channel network. J Hydrol 122:275–287

Szymkiewicz R (1993) Solution of the inverse problem of the Saint Venant equations. J Hydrol 147:105–120

Szymkiewicz R (2000) Modelowanie matematyczne przepływów w rzekach i kanałach. Wydawnictwo PWN, Liczba stron 265

Zieliński J (1999) Wisła Środkowa po przeprowadzeniu prac regulacyjnych (Middle Vistula after regulation). In: Kucharczyk M (ed) Problemy ochrony i renaturalizacji dolin dużych rzek Europy (Problems of protection and re-naturalisation of big European rivers). UMCS Publishing House, Lublin, pp 195–200

Sensitivity Analysis of the Flow Routing Model for the Middle River Vistula-Multi-method Approach

Marzena Osuch

Abstract In this study a sensitivity analysis of a one-dimensional flow routing model MIKE 11 is applied to the Middle River Vistula reach located between Zawichost and Warsaw Port Praski gauging stations. Three methods of sensitivity analysis are compared, namely, Morris, Sobol and a correlation-based approach allowing for the choice of appropriate model calibration strategy. The influence of the model parameters (Manning roughness coefficients) on the goodness of fit of the water levels at seven locations has been estimated. The results show that the roughness coefficients of the analysed cross-sections and the next downstream reaches have a significant influence on model output (Nash Sutcliffe values). Only the results for the gauging station in Warsaw Port Praski differ, due to the influence of the boundary condition (rating curve in Warsaw). A comparison of the outcomes from the three methods shows that the results from the beginning of the study reach are more similar to each other than for the cross-sections located downstream, as a consequence of uncertainty propagation from the upstream cross-sections and the influence of the downstream boundary condition. Generally, the results indicate that an appropriate model calibration strategy should not consist of a series of model optimizations for independent sub-reaches but rather it should be integrated over all reaches and simultaneously take into account the model fit at each controlled cross-section.

Keywords Sensitivity analysis · Model calibration · MIKE 11 · Middle River Vistula

M. Osuch (✉)
Institute of Geophysics, Polish Academy of Sciences,
Ksiecia Janusza 64, 01-452 Warsaw, Poland
e-mail: marz@igf.edu.pl

R.J. Romanowicz and M. Osuch (eds.), *Stochastic Flood Forecasting System*,
GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-319-18854-6_7

1 Introduction

In hydrological practice, mathematical models are developed and applied to descriptions of environmental systems. Flow routing in rivers is described by hydrodynamic models of flow in an open channel (Chow 1959; Mahmood et al. 1975; Szymkiewicz 2000; Madsen et al. 2003; Gasiorowski et al. 2015). These models are based on the basic principles of physics (mass, momentum and energy conservation). In the case of floods, and especially when a flood wave exceeds the embankments, the flow is complex over a wide range of space scales and is unsteady in time. Therefore, it should be modelled as a three-dimensional phenomenon, but in practice, due to the lack of data and inadequate computational power, the use of 3D models is limited (Hunter et al. 2007). In many cases, the flow in a compound channel is described as two-dimensional (Bates et al. 1998; Feldhaus et al. 1992; Gee et al. 1990). In practice, especially when the study is aimed at a prediction of water levels at a particular cross-section without significant concern to represent floodplain flow and storage accurately, 1D hydraulic models are used (Bates and Roo 2000). The use of 1D models to describe the flow through the area of complex geometry is often considered as too large a simplification of the flow field (Knight and Shiono 1996), but it is just an additional source of error which is small compared to other errors, resulting from the uncertainty of the boundary conditions, the cross-sectional geometry, and also the roughness coefficients (Ali and Goodwin 2002; Bates 2007).

The flow routing models are characterized by a number of parameters that should be estimated taking into account the observations. Only a few have a significant influence on the simulations and associated model output and should be calibrated. It is generally accepted that a sensitivity analysis (SA) assessing the influence of model parameters/factors on the model output should be conducted as a part of the calibration efforts and model implementation (Romanowicz and Macdonald 2005; He et al. 2011). The SA methods are quantified in multiple ways but in general they are divided into two categories: local and global (Saltelli et al. 2004; Song et al. 2015). Local sensitivity methods give an estimate of analysed factors on the model output around a specific point of interest. A global sensitivity analysis reflects the importance of factors (model parameters) through the multivariate space of a model response (Herman et al. 2013). This category includes the following methods: screening (the most commonly used being the Morris method), regression, variance-based, meta-modelling regionalized sensitivity analysis and entropy-based. In practice, several methods of SA are applied. A detailed review of concepts, methods and hydrological applications of SA are presented, e.g., in Song et al. (2015).

It should be noted that in only a few studies has a sensitivity analysis been applied to flow routing models (Hall et al. 2005; Pappenberger et al. 2006, 2008; Cloke et al. 2008; Kiczko et al. 2008; Liu and Sun 2010).

Pappenberger et al. (2008) presented an assessment of the influence of the following factors: Manning channel roughness, Manning roughness floodplain, weighting parameter of the numerical scheme and boundary conditions, on the results of flood inundation models using five SA methods (Sobol, Kullback-Leiber entropy, Morris, regionalised sensitivity analysis and regression). The

results indicated that different methods produce different assessments of parameter importance. The outcomes of the analysis with an assumption of Nash-Sutcliffe coefficients as a model output indicate that the channel roughness coefficient is the most important parameter. The influence of the other factors is much smaller.

Kiczko et al. (2008) evaluated the influence of the channel and floodplain roughness coefficients, lateral inflow coefficient, delay of lateral inflows and boundary conditions on model output. The analysis carried out for the Upper Narew indicated that channel roughness coefficients together with lateral inflow are the factors with highest impact on model output.

Hall et al. (2009) reviewed a set of SA methods taking into account their applicability to hydraulic models. Three methods of modelling a simple pipe bend using an advanced Shallow Water Equation solver were used as a case study. The results indicated that the incautious use of SA methods could provide misleading conclusions, especially in the case of nonlinear models. The application of local SA methods may result in an incomplete picture of model response over the range of variability of model inputs.

An analysis of the influence of parameters of the 1D hydrodynamic model is also a main theme of this work, which has been carried out in the framework of the project entitled "Stochastic flood forecasting system: the Middle River Vistula case study" (Romanowicz and Osuch, Chap. 9, this book). As the title suggests, it aims to formulate and develop a flood forecasting system for the studied river reach. The proposed system has a modular structure and includes different types of models describing flow routing inter alia 1D hydrodynamic model MIKE 11. It is generally assumed that the model should be as simple as possible but at the same time should provide accurate results of simulations for mean and high water levels. The choice of model is also constrained by the available hydrological data and meteorological forecasts. For this reason, lateral inflows are excluded from the flow routing model. In addition, a comparison of flows in the River Vistula and their tributaries confirms this simplification of the model. The chosen 1D hydrodynamic model should be calibrated on the basis of observed water levels and flows estimated through rating curves. The first trial calibration using one parameter value for the whole reach or an independent calibration for individual sub-reaches did not give satisfactory results. For this reason, a sensitivity analysis is conducted to help in the choice and development of an efficient and effective flow routing model calibration strategy. In this chapter the following research objectives are identified:

- an assessment of the influence of the factors (model parameters) on accuracy of the MIKE 11 model output described by means of Nash Sutcliffe coefficients calculated for water levels at seven controlled cross-sections located along the Middle River Vistula (Nash and Sutcliffe 1970)
- an application of three methods of sensitivity analysis: Morris, Sobol and a correlation-based approach to flow routing model
- a comparison of the estimates of sensitivity measures describing the direct influence of model parameters and an overall impact together with an assessment of nonlinearity effects and interactions between parameters obtained from the three methods applied

Fig. 1 Study area—the Middle River Vistula reach between Zawichost and Warsaw Port Praski gauging stations

The analyses are performed for the middle reach of the River Vistula located between Zawichost and Warsaw Port Praski gauging stations (Fig. 1).

2 Study Area

In this chapter, the Middle River Vistula reach between Zawichost and Warsaw Port Praski is presented. Hydrological measurements are carried out at 7 gauging stations along that river reach: Zawichost, Annopol, Puławy, Dęblin, Gusin, Nadwilanówka and Warsaw Port Praski. As shown in Fig. 1, there are several tributaries but their influence on the main channel flow is negligible and they were excluded from the analysis.

3 Flow Routing Model

A one dimensional unsteady flow routing model (MIKE 11) was applied in this study. The river geometry was described by 112 cross-sections along the river reach. A detailed report on the cross-sections and river geometry is presented in Kochanek et al. (Chap. 6, this book). Flow hydrographs in Zawichost and the rating curve at Warsaw Port Praski were applied as model boundary conditions.

In the model the variability of the roughness coefficients was simplified and only one Manning roughness coefficient was used to represent the cross-section, neglecting differences between the main channel and floodplains. To justify this simplification it is worth mentioning that in all cases analysed the maximum water level was below the levee crown. The model requires the determination of Manning coefficients for all the analysed cross sections. In this work it was assumed that the MIKE 11 model is controlled by seven Manning roughness coefficients at the following monitored cross-sections: Zawichost, Annopol, Puławy, Dęblin, Gusin, Nadwilanówka and Warsaw Port Praski. In the cross-sections situated along the reaches between the gauging stations, the Manning coefficients were automatically linearly interpolated by the MIKE 11 software.

As a measure of model output the Nash-Sutcliffe coefficient calculated on the basis of simulated and observed water levels in the seven controlled cross sections was applied. The analysis was conducted using water level observations from the period 01 November 2000 to 31 October 2010.

4 Sensitivity Analysis

A study of the influence of model parameters on the model outputs was carried out using three methods: Morris, Sobol, and the correlation tests. A description of these methods is presented in the following sections.

4.1 The Morris Method

To evaluate the influence of model parameters on modelled output, a sensitivity analysis using the Morris method was applied (Morris 1991). In this method the sensitivity is estimated on the basis of a number of local changes at different points in the possible range of model parameters, called the elementary effects. The global results are obtained by regional exploration of the input space (a chosen number of trajectories).

Each parameter x_i is perturbed along a grid of size Δ_i in the chosen parameter space (between selected limits). It is assumed that for a model with p parameters each trajectory consist of a sequence of p perturbations. Each trajectory allows us to estimate one elementary effect for each parameter. The elementary effect for parameter i is calculated according to the following equation:

$$EE_i = \frac{f(x_1, \ldots, x_i + \Delta_i, \ldots, x_p) - f(x)}{\Delta_i} \tag{1}$$

where f(x) denotes the prior point in the trajectory. The analysis applied to a single trajectory is strongly dependent on the location of the initial point in parameter space. Therefore, in the Morris method, the analyses are repeated for a chosen number of trajectories sampled according to some recommended methods. In the approach originally proposed by Morris, the trajectories are randomly sampled over the parameter space. More accurate estimates of global sensitivity analysis are obtained following Campolongo et al. (2007, 2011) and Ruano et al. (2012) where trajectories that maximize the coverage of parameter space are selected.

The results of this method are presented in the form of two measures: μ^* and σ. The first measure, μ_i^*, represents the overall influence of a factor/parameter i on the output (in this case MIKE 11 model parameter). It was introduced by Campolongo et al. (2007) and is computed from the mean of the absolute values of the elementary effects which allows the problem of effects of an opposite sign to be solved.

$$\mu_i^* = \frac{\sum_{i=1}^{r} |EE_i|}{r} \tag{2}$$

The second measure, σ_i, estimates the ensemble of the factor's higher order effects (i.e. nonlinearity or interaction with other parameters) for parameter i.

$$\sigma_i = \frac{\sum_{i=1}^{r} (EE_i - \mu_i)^2}{r} \tag{3}$$

These two measures allow for a classification of the MIKE 11 model parameters into two groups, taking into account their effect on the analysed model output (here, the Nash-Sutcliffe coefficients), and permit a choice of an appropriate MIKE 11 model calibration strategy.

4.2 The Sobol Method

The Sobol method, based on an analysis of variance output from the model (Archer et al. 1997), is the second approach applied here. In this method, it is assumed that the total variance of the output from the model, calculated using the Monte Carlo method, can be decomposed into the sum of the variance of each input factor explained by the fractional contribution (model parameters in this case), and by interactions between them.

$$V(Y) = \sum_{i} V_i + \sum_{i<j} V_{ij} + \sum_{i<j<m} V_{ijm} + \cdots + V_{1,2\ldots(l-1)} \tag{4}$$

where Y is the analysed model output (in this chapter the Nash-Sutcliffe coefficients calculated on basis of simulated and observed water levels at seven gauging stations), $V(Y)$ is the total variance of the analysed model output, i, j, m are model parameters, l the number of model parameters, V_i the model output variance explained by influence of the ith parameter, V_{ij} the model output variance explained by influence of the interactions between parameters i and j, V_{ijm} the model output variance explained by the influence of the interactions between parameters i, j and m, and $V_{1,2\ldots(l-1)}$ is the model output variance explained by influence of the interactions between parameters up to $l - 1$ order. The decomposition is unique if the input factors are independent of each other.

As for the Morris method, the Nash-Sutcliffe coefficient calculated on the basis of simulated and observed water levels was applied as a measure of model output.

The outcomes of a sensitivity study by the Sobol method are usually presented as the first order, higher order and total order sensitivity measures. The first order sensitivity index (S_i) describes the influence of the parameter i on the model output.

$$S_i = \frac{V_i}{V(Y)} \tag{5}$$

Higher order sensitivity measures describe the influence of interactions between model parameters on model output. For example, the second order sensitivity measure between the parameters i and j can be written as:

$$S_{ij} = \frac{V_{ij}}{V(Y)} \tag{6}$$

The model sensitivity to the interactions among subsets of parameters, the so-called higher order effects, is investigated using the Sobol total sensitivity measure: S_{Ti}. This measure represents all interactions which involve parameter i.

For example, the total sensitivity measure for the Manning roughness coefficient for the cross-section in Zawichost ($n_{Zawichost}$) can be written as the sum of the (i) first order sensitivity measure for $n_{Zawichost}$, (ii) six second order measures assessing interactions between: $n_{Zawichost}$ and $n_{Annopol}$, $n_{Zawichost}$ and $n_{Puławy}$, $n_{Zawichost}$ and $n_{Dęblin}$, $n_{Zawichost}$ and n_{Gusin}, $n_{Zawichost}$ and $n_{Nadwilanówka}$, $n_{Zawichost}$ and n_{Warsaw}, and (iii) the corresponding third, fourth, fifth and sixth order sensitivity indices.

The total sensitivity measure was approximated using numerical integration in a Monte Carlo framework by the so-called 'freezing unessential variables' method (Sobol 2001; Saltelli 2002; Saltelli et al. 2008). According to Homma and Saltelli (1996) application of the approximation method gives accurate and informative results at reasonable computational cost.

4.3 The Correlation Based Method

Two correlation coefficients, namely Pearson and Spearman, were used as measures of the influence of model parameters on the analysed outputs in the third study method. As in the two previous methods, an application of the correlation-based method consists of three steps: (a) Monte Carlo parameter sampling, (b) model simulations, and (c) the derivation of sensitivity measures.

The Pearson correlation coefficient (r_P) is a measure of the linear dependence between two variables, whilst the Spearman correlation coefficient (r_S) is an indicator of a monotonic dependence. In both cases, the values of r_P and r_S range from -1 to 1. A value of 0 indicates no correlation whilst -1 or 1 indicate, respectively, a perfect negative or positive correlation. There is a difference in the behaviour of these two measures in their sensitivity to outliers and to the shape of the function describing dependence (curvature), namely, r_P is sensitive and r_S is not. As a result, a comparison of these two correlation measures allows for a quantification of the nonlinearity effect. In both cases, together with the correlation values, a corresponding *p*-value statistic for determination of the significance of correlation coefficient is estimated. It has been assumed that a *p*-value smaller than 0.05 indicates a statistically significant correlation and also an important influence on the model output.

5 Discussion and Interpretation of Results

In this section, the results of the assessment of an influence of model parameters on model outputs using three methods (Morris, Sobol, and the correlation-based approach) are presented. In all cases we analysed a one-dimensional flow routing model for the Middle River Vistula between gauging stations Zawichost and Warsaw Port Praski.

5.1 The Morris Method

An application of the Morris method requires the determination of a set of parameters: the number of levels *p* (usually within the range of [0 10]), and the number

of trajectories R which describes the sample size and is therefore connected to computational cost (Zhan et al. 2013). In this chapter it was assumed that all the MIKE 11 parameters (the Manning roughness coefficients at controlled cross-sections in: Zawichost, Annopol, Puławy, Dęblin, Gusin, Nadwilanówka and Warsaw Port Praski) are sampled from a uniform distribution within the range [0.005 0.2], the number of levels p equals 10 and the number of trajectories $R = 15$. As a result of these assumptions, the sample size (N) is equal to 120.

The results of a sensitivity analysis by the Morris method, the estimated two sensitivity measures (μ* and σ) are presented in Fig. 2 in the form of screening plots which have these two measures as the x- and y-axes. These types of graphs make the interpretation of the results and classification of the model parameters easier. According to the classification being introduced here, the parameters with values of μ* (i) higher than 1 are 'very important', (ii) within the range [0.1 1) are 'important', (iii) within the range [0.01 0.1) are 'unimportant' and (iv) below 0.001 are 'irrelevant' for the analysed model output. Taking into account two sensitivity measures (μ* and σ), the model parameters could be classed into three groups (i) parameters with negligible effect—low values of μ* and σ, (ii) parameters with linear and additive effects for high values of μ* and low values of σ, (iii) parameters with nonlinear or interaction effects for high values of μ* and σ.

Seven panels in Fig. 2 show the outcomes for seven analysed model outputs, NS coefficients, calculated for simulated and observed water levels at the following cross-sections: Zawichost, Annopol, Puławy, Dęblin, Gusin, Nadwilanówka and Warsaw Port Praski. The results of the assessment for the Zawichost gauging station are presented in the upper left panel. It is seen that for the appropriate calibration of water levels in Zawichost, the parameter $n_{Zawichost}$ has the highest influence. The influence of the Manning roughness coefficient from the Annopol gauging station could be described as very important. The impact of the other parameters ($n_{Puławy}$) on the Nash-Sutcliffe values in Zawichost is unimportant or irrelevant ($n_{Dęblin}$, n_{Gusin}, $n_{Nadwilanówka}$ and n_{Warsaw}). The estimates of σ indicated that the nonlinearity effect is less pronounced for $n_{Zawichost}$ than for $n_{Annopol}$.

The sensitivity measures estimated for NS in Annopol are presented in the upper right panel of Fig. 2. It is seen that the highest influence on model output comes from $n_{Annopol}$. The impact of $n_{Puławy}$ was classified as important. The estimates of μ* and σ for others parameters are smaller than 0.1, indicating an unimportant or irrelevant influence on the analysed model output.

The NS coefficient estimated for simulated and observed water levels at the Puławy gauging station shows a strong influence from $n_{Puławy}$, $n_{Dęblin}$ parameters. Two parameters ($n_{Annopol}$ and n_{Gusin}) were classified as important and three others ($n_{Zawichost}$, n_{Warsaw} and $n_{Nadwil.}$) as unimportant.

The results of the assessment for the Dęblin gauging station indicated that four parameters ($n_{Dęblin}$, $n_{Annopol}$, n_{Gusin} and $n_{Puławy}$) could be regarded as very important. The influence of the other three parameters is assessed as unimportant for the MIKE 11 fit at Dęblin.

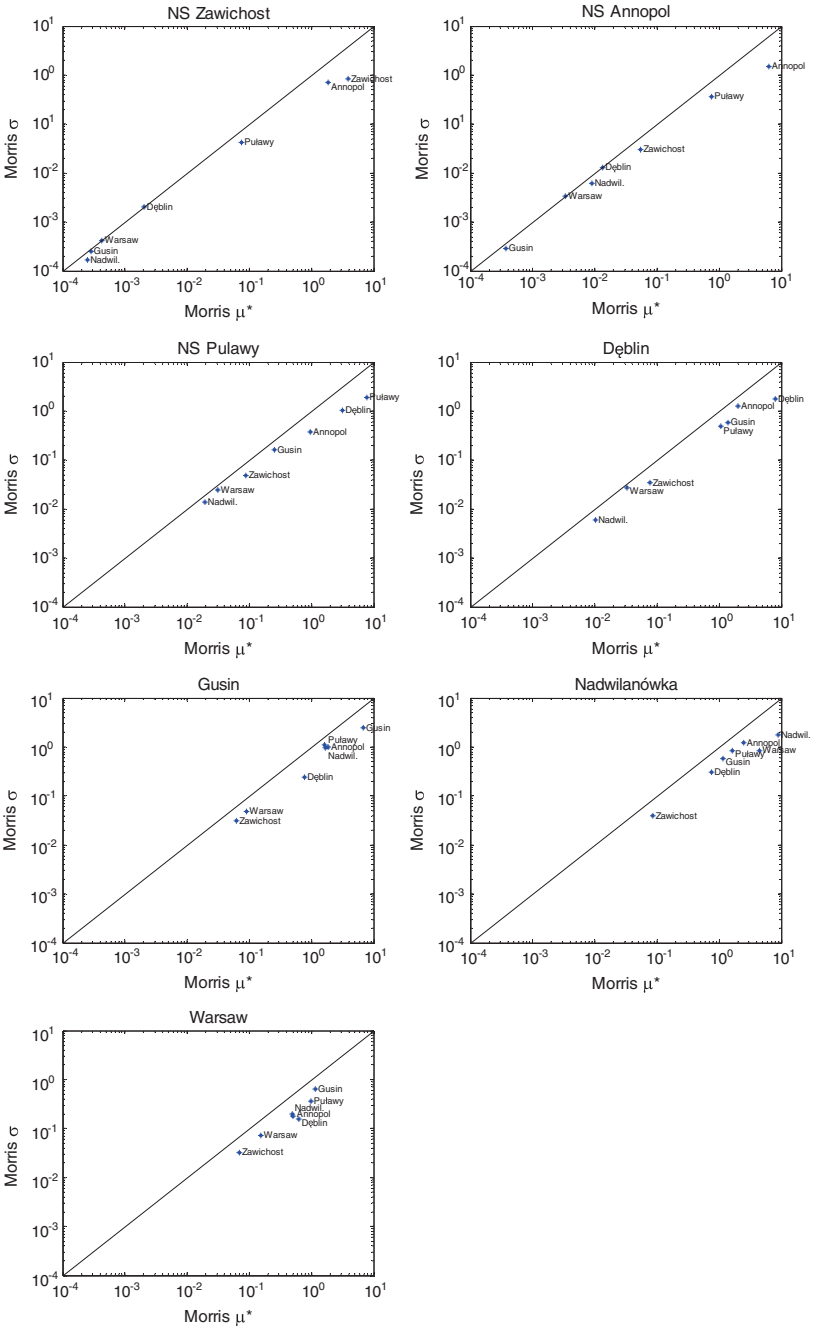

Fig. 2 Results of sensitivity analysis by Morris method for NS estimated at Zawichost, Annopol, Puławy, Dęblin, Gusin, Nadwilanówka and Warsaw Port Praski gauging stations

In the case of Gusin, four parameters are very important for proper model calibration. The influence of parameter $n_{Dęblin}$ is important, whilst $n_{Zawichost}$ and n_{Warsaw} are unimportant for the calculated NS values.

The results for the Nadwilanówka gauging station are similar to those for Gusin except for the influence of n_{Warsaw} which was classified as very important.

The estimates of μ^* and σ for Warsaw Port Praski indicated that n_{Gusin} could be classified as very important, five parameters ($n_{Puławy}$, $n_{Nadwil.}$, $N_{Dęblin}$, $n_{Annopol}$ and n_{Warsaw}) as important and $n_{Zawichost}$ as unimportant for the model fit in Warsaw Port Praski.

Summarizing, the results of a sensitivity analysis by the Morris method indicate that in almost all cases the parameter that has the greatest impact is the Manning roughness coefficient from the analyzed cross-section. For example, in the case of NS calculated for simulated and observed water levels in the Annopol gauging station the parameter that has the highest influence is the Manning channel roughness coefficient in Annopol. The Manning roughness coefficient from the nearest downstream cross-section is the second parameter, with an important influence on the analyzed output in most cases. The Warsaw Port Praski cross-section shows different results, probably resulting from the influence of the downstream boundary condition.

5.2 The Sobol Method

The Sobol first and total order sensitivity indices were calculated. Seven MIKE 11 model parameters were sampled separately, using quasi-random Sobol sequences from uniform distribution within the ranges [0.005 0.2]. A base of $bn = 200$ was used, resulting in 3200 model evaluations.

The outcomes of the sensitivity study by the Sobol method are usually presented as the first order and total order sensitivity measures. Figure 3 presents the estimated values of the first order sensitivity index for seven analysed model outputs. The results of the sensitivity assessment for the Nash-Sutcliffe values calculated using simulated and observed water levels in Zawichost are presented in the upper left panel. The variance of NS at Zawichost gauging station is explained by two parameters: $n_{Zawichost}$ and $n_{Annopol}$. In the case of two model outputs, in the Annopol and Dęblin cross-sections, the quality of fit depends only on the Manning roughness coefficients from these cross-sections. The Sobol first order sensitivity indices for Puławy depend on two parameters: $n_{Puławy}$ and $n_{Dęblin}$. For the next three analysed outputs, at Gusin, Nadwilanówka and Warsaw Port Praski, the estimated variance of NS is explained by three or more parameters. In the case of the Nadwilanówka cross-section, the MIKE 11 model goodness of fit depends on all parameters with the highest influence of $n_{Nadwilanówka}$. The estimates of S_i for Warsaw Port Praski indicate an influence of almost all model parameters except $n_{Zawichost}$. Only in this case, the Manning roughness coefficient does not have the

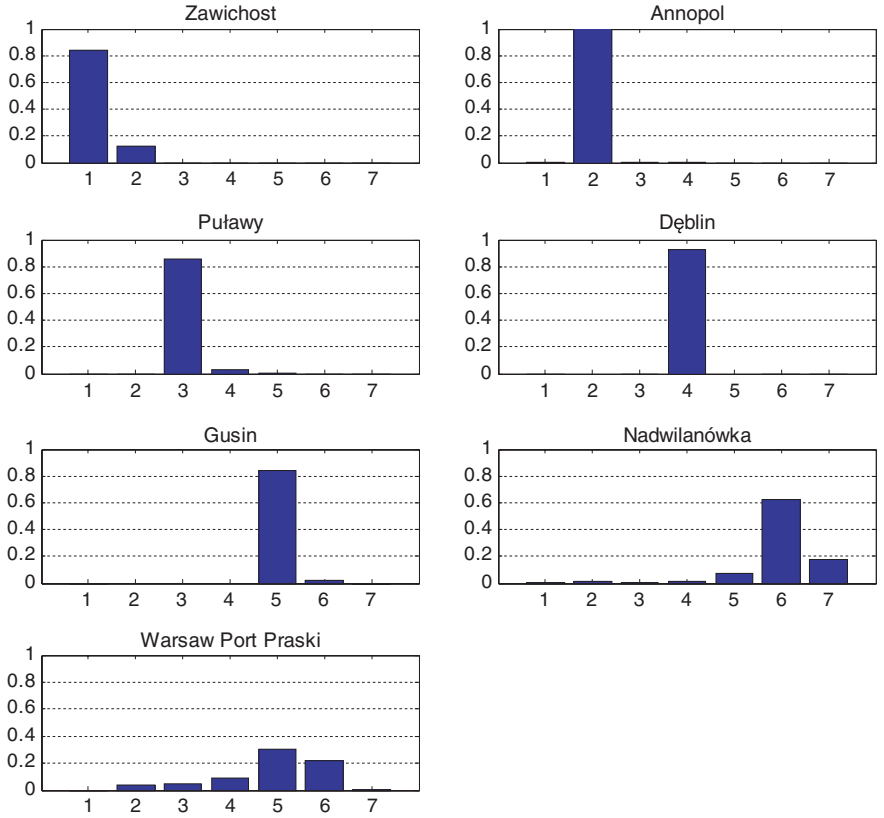

Fig. 3 The results of sensitivity analysis by the Sobol method. The panels present the estimated first order sensitivity indices for the seven analysed model outputs. The x-axis denotes the MIKE 11 model parameter values (manning roughness coefficients at the cross-sections: *1*—Zawichost, *2*—Annopol, *3*—Puławy, *4*—Dęblin, *5*—Gusin, *6*—Nadwilanówka, *7*—Warsaw Port Praski)

highest influence on the model output. Probably, it is a result of an influence of a downstream boundary condition.

In the method of Sobol, the overall impact of the model parameters, including nonlinearity effect and interactions between parameters, was assessed using the total order sensitivity index. The results are shown in Fig. 4. They are very similar to the estimates of first order sensitivity indices, thus indicating a negligible influence of nonlinearity and no interactions between model parameters.

5.3 The Correlation-Based Approach

Pearson and Spearman correlation coefficients were calculated for the quasi-randomly sampled MIKE 11 model parameters and seven model outputs. The same sample (N = 3200) was used as for the Sobol method.

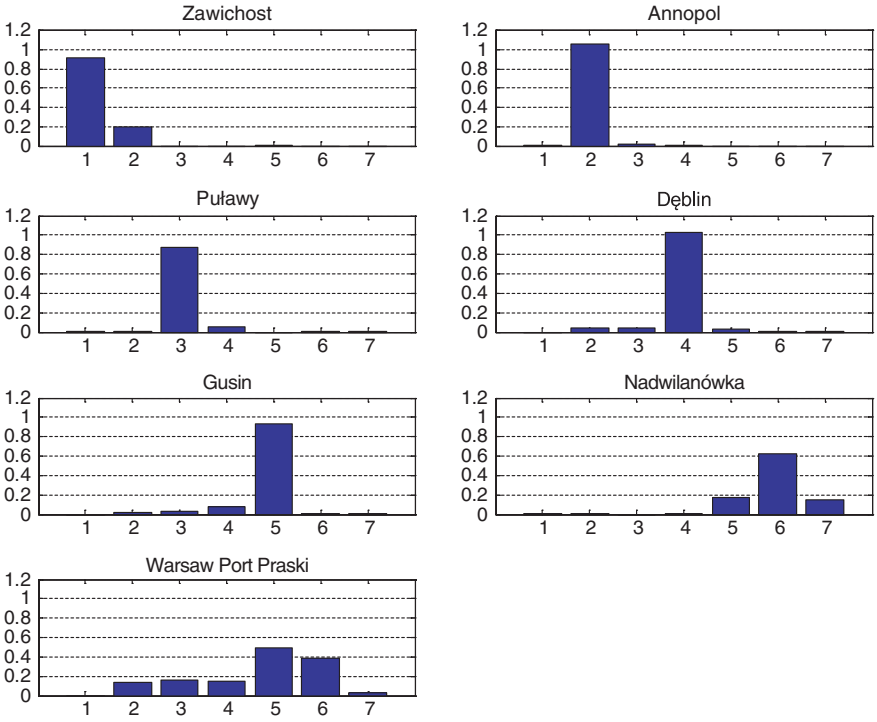

Fig. 4 The results of a sensitivity analysis by the Sobol method; assessment of the total order sensitivity indices of the model parameters on the NS values calculated on basis simulated and observed water levels. The x-axis denotes the MIKE 11 model parameter values (manning roughness coefficients at the cross-sections: *1*—Zawichost, *2*—Annopol, *3*—Puławy, *4*—Dęblin, *5*—Gusin, *6*—Nadwilanówka, *7*—Warsaw Port Praski)

The estimated Pearson correlation coefficients together with their statistical significance at the 0.05 level are presented in Table 1. The estimated r_P values vary between the analysed outputs and model parameters. The highest values are obtained for the parameters from cross-sections where the NS is calculated. This behaviour is not visible in the case of the Warsaw Port Praski gauging station. The number of parameters with statistically significant r_P varies from one for Annopol to six for Warsaw Port Praski. On the basis of the estimated r_P values the entire reach could be divided into two parts: (a) an upper part from Zawichost to Dęblin, and (b) a lower part from Gusin to Warsaw. In the upper part the model outputs are correlated mainly with the parameters from the analysed cross-sections whilst in the lower part almost all parameters have a significant influence on model output.

The results for the influence of model parameters on model outputs using the Spearman correlation coefficient are presented in Table 2. The outcomes are similar to those for the Pearson correlation coefficients, indicating a small influence of nonlinearity between model parameters and analysed outputs. In almost all cases, the highest correlation is estimated for the Manning roughness coefficient from the cross-section where the model output is calculated. In addition to the correlation

Table 1 The results of a sensitivity study by a correlation-based method (Pearson coefficient of correlation)

		n Zwichost	n Annopol	n Puławy	n Dęblin	n Gusin	n Nadwil.	n Warsaw
NS Zawichost	corr	**−0.8549**	**−0.3257**	0.0024	−0.0179	−0.0142	0.0213	0.0123
	pval	0.0000	0.0000	0.8932	0.3133	0.4232	0.2292	0.4885
NS Annopol	corr	−0.0188	**−0.7766**	−0.0011	−0.0154	−0.0269	0.0182	−0.0101
	pval	0.2893	0.0000	0.9513	0.3852	0.1287	0.3044	0.5683
NS Puławy	corr	0.0003	**−0.0415**	**−0.9213**	**−0.2027**	−0.0074	0.0017	0.0124
	pval	0.9842	0.0190	0.0000	0.0000	0.6762	0.9239	0.4827
NS Dęblin	corr	**−0.0392**	−0.0191	−0.0280	**−0.7927**	−0.0068	0.0051	0.0022
	pval	0.0268	0.2815	0.1136	0.0000	0.7003	0.7738	0.9030
NS Gusin	corr	−0.0302	**−0.0881**	**−0.0474**	**−0.0868**	**0.1618**	**0.0925**	−0.0122
	pval	0.0885	0.0000	0.0075	0.0000	0.0000	0.0000	0.4896
NS Nadwil.	corr	−0.0190	−0.0165	**−0.0512**	**0.0440**	**0.0897**	**−0.7462**	**−0.4248**
	pval	0.2843	0.3526	0.0038	0.0130	0.0000	0.0000	0.0000
NS Warsaw	corr	0.0128	**−0.2784**	**−0.2721**	**−0.2209**	**−0.4850**	**−0.3452**	**−0.0452**
	pval	0.4687	0.0000	0.0000	0.0000	0.0000	0.0000	0.0107

Bold values indicate r_P statically significant at 0.05 level

Table 2 The results of a sensitivity study by a correlation method (Spearman coefficient of correlation)

		n Zwichost	n Annopol	n Puławy	n Dęblin	n Gusin	n Nadwil.	n Warsaw
NS Zawichost	corr	**−0.8528**	**−0.3147**	0.0032	−0.0223	−0.0115	0.0313	0.0083
	pval	0.0000	0.0000	0.8563	0.2082	0.5150	0.0768	0.6384
NS Annopol	corr	−0.0146	**−0.7667**	−0.0047	−0.0133	−0.0255	0.0179	−0.0107
	pval	0.4108	0.0000	0.7926	0.4520	0.1493	0.3122	0.5447
NS Puławy	corr	0.0037	**−0.0508**	**−0.9223**	**−0.1838**	−0.0108	−0.0005	0.0159
	pval	0.8355	0.0041	0.0000	0.0000	0.5406	0.9778	0.3684
NS Dęblin	corr	−0.0313	**−0.0457**	**−0.0485**	**−0.7861**	−0.0103	−0.0019	0.0154
	pval	0.0774	0.0098	0.0061	0.0000	0.5590	0.9127	0.3849
NS Gusin	corr	−0.0067	**−0.0999**	**−0.1038**	**−0.1106**	**−0.1835**	−0.0002	−0.0129
	pval	0.7063	0.0000	0.0000	0.0000	0.0000	0.9894	0.4658
NS Nadwil.	corr	−0.0108	−0.0124	**−0.0493**	−0.0064	**0.0416**	**−0.8092**	**−0.4376**
	pval	0.5430	0.4846	0.0054	0.7190	0.0187	0.0000	0.0000
NS Warsaw	corr	0.0093	**−0.3164**	**−0.3201**	**−0.2765**	**−0.5200**	**−0.4190**	**−0.0524**
	pval	0.6011	0.0000	0.0000	0.0000	0.0000	0.0000	0.0031

Bold values indicate r_S statically significant at 0.05 level

coefficients their significance at the 0.05 level is also tested. The number of parameters with a statistically significant Spearman correlation coefficient varies between the analysed model outputs, from one in the case of Annopol to six for the Warsaw Port Praski cross-section.

6 Comparison of the Results

This section presents an assessment of the influence of the MIKE 11 model parameters on the Nash Sutcliffe values estimated on the basis of simulated and observed water levels at seven cross-sections: Zawichost, Annopol, Puławy, Dęblin, Gusin, Nadwilanówka and Warsaw Port Praski. Three methods of sensitivity analysis have been applied: the Morris, Sobol, and correlation-based approach. As a result, a number of sensitivity indices/measures were derived. Generally, they could be classified into two groups: (i) measures assessing the direct influence of individual parameters and (ii) measures of the overall influence of individual parameters together with their interactions and nonlinearity effects. In the following two sections, SA measures classified above are compared.

6.1 A Comparison of a Direct Influence of MIKE 11 Model Parameters on the NS Criterion

In this study three measures of the direct influence of model parameters on the model output were tested. They are: (i) μ^* in the Morris method, (ii) S_i—first order sensitivity index from the Sobol method and (iii) r_p—the Person correlation coefficient. The results are presented in Fig. 5 in seven panels, separately for each analysed model output. The double y-axes present estimated values of: Si, μ^* and absolute value of r_P, respectively. All the methods applied give similar results (ranks of parameters) regarding the very important and important parameters. The highest differences in the parameter ranks, but small in the values of sensitivity measures, are found for unimportant or irrelevant parameters. A comparison of the results obtained by three methods along the river reach shows that the results at the beginning of the reach are more similar to each other than those at the cross-sections located downstream as a consequence of uncertainty propagation from the upstream located cross sections.

6.2 A Comparison of the Overall Influence of Individual MIKE 11 Model Parameters

In addition to the estimation of a direct influence of model parameters on model output, the model parameter interactions and model non-linearity have also been assessed. Therefore, three measures have been calculated: S_{Ti}—the Sobol total order sensitivity index, σ from the Morris method and r_S—the Spearman correlation coefficient. In the case of the latter, only the non-linearity effect is taken into account. The results for each analysed model output (NS coefficient at seven gauging station-related cross-sections) are presented in Fig. 6. Generally, the outcomes from

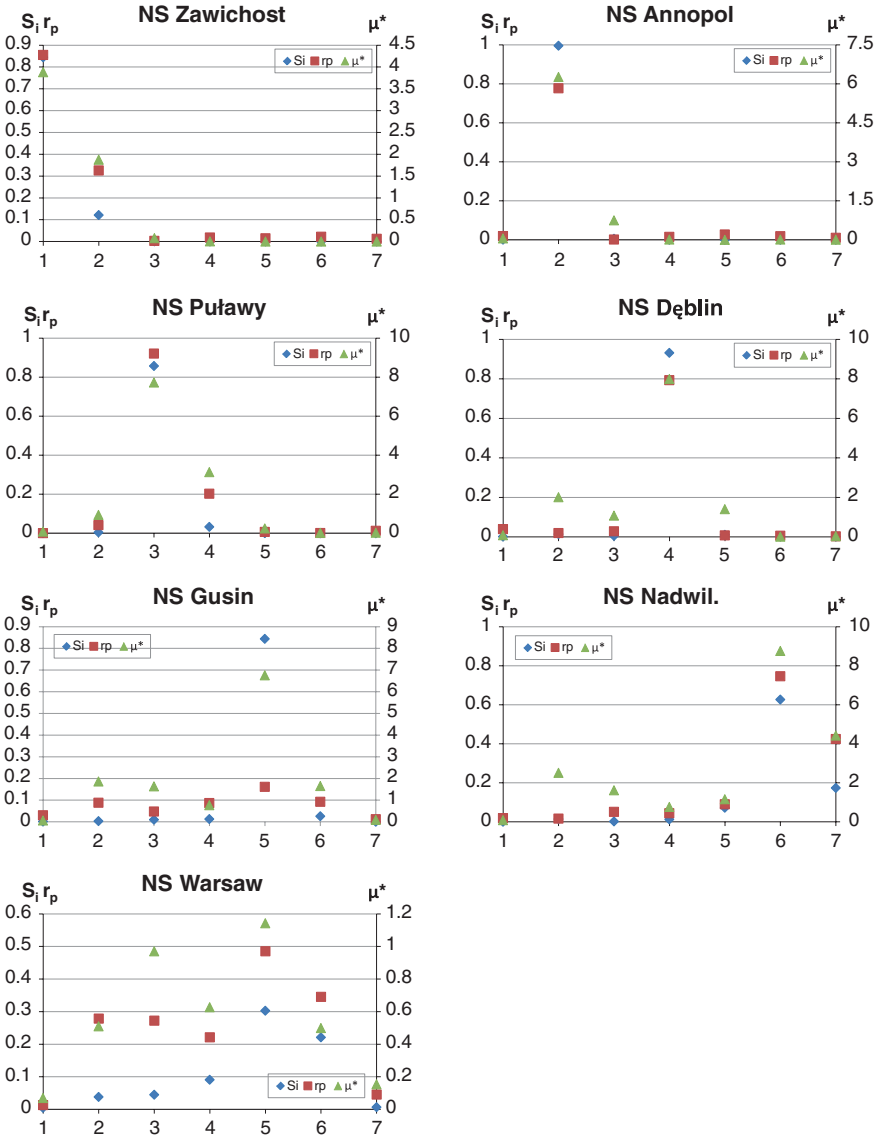

Fig. 5 A comparison of the sensitivity analysis results using three measures: S_i, $|r_p|$ and μ^*. The x-axis denotes the MIKE 11 model parameters (manning roughness coefficients at the cross-sections: *1*—Zawichost, *2*—Annopol, *3*—Puławy, *4*—Dęblin, *5*—Gusin, *6*—Nadwilanówka, *7*—Warsaw Port Praski)

the three methods applied are similar for the parameters with a high influence on the model output. Significant differences in the parameter ranking and in the estimated measures were found for the rest of the parameters. The results for the two most upstream located cross-sections are similar, contrary to the estimates for the

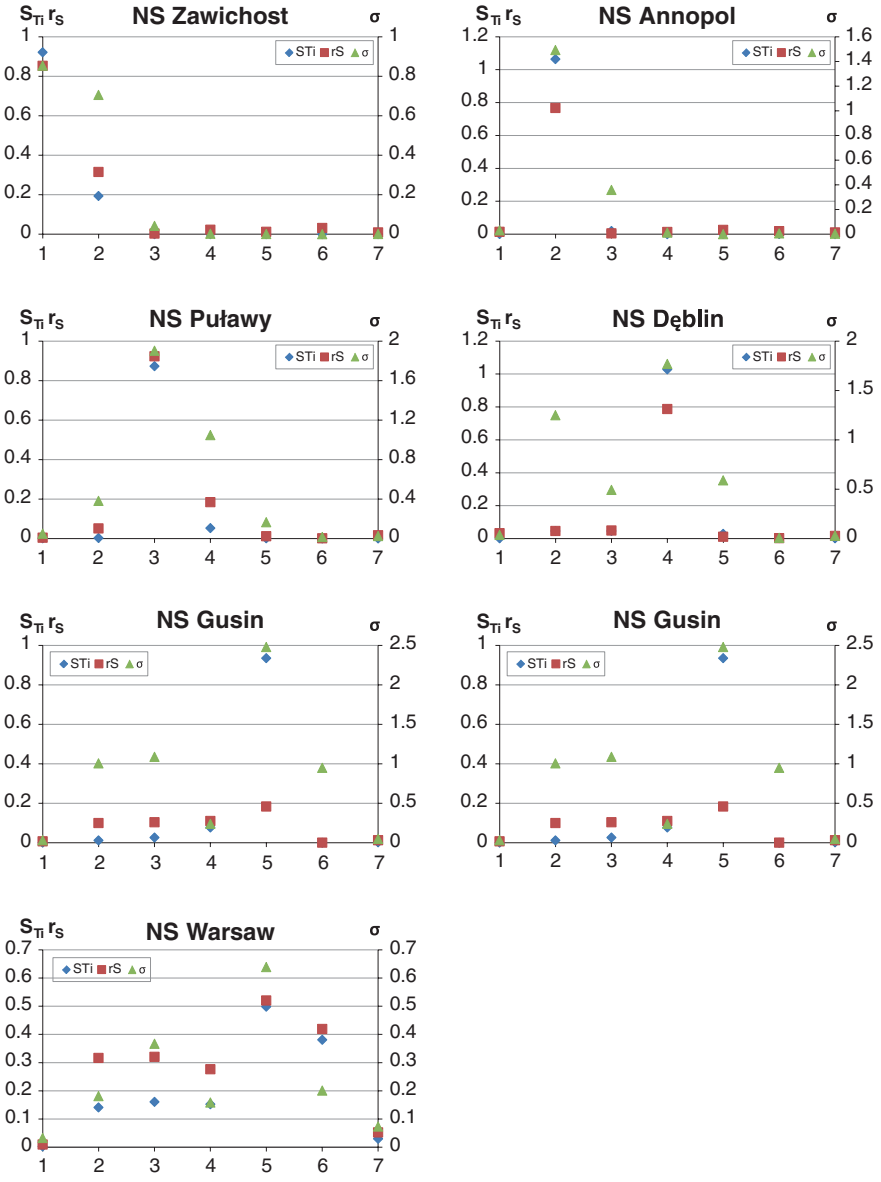

Fig. 6 A comparison of the sensitivity analysis results using three measures: S_{Ti}, r_S and σ. The x-axis denotes the MIKE 11 model parameters (manning roughness coefficients at the following cross-sections: *1*—Zawichost, *2*—Annopol, *3*—Puławy, *4*—Dęblin, *5*—Gusin, *6*—Nadwilanówka, *7*—Warsaw Port Praski)

downstream located sections. These differences in the results at the cross-sections in the lower end are due to uncertainty propagation from the upstream located cross-sections and also an influence of downstream boundary condition.

7 Conclusions

This study presents the results of sensitivity analyses by the three methods: Sobol, Morris and correlation-based approach applied to the flow routing model MIKE 11 for the reach of the River Vistula between Zawichost and Warsaw Port Praski. The applied model is characterized by Manning roughness coefficients for each of 114 cross-sections. It is assumed that only seven roughness coefficients from the controlled cross-sections: Zawichost, Annopol, Puławy, Dęblin, Gusin, Nadwilanówka and Warsaw Port Praski are randomly drawn in the SA procedures. The Manning values for other cross-sections are linearly interpolated by MIKE 11 software. The Nash-Sutcliffe efficiency criterion estimated on the basis of simulated and observed water levels in each of the control sections was selected as the model output.

We compared the direct and total influence of MIKE 11 model parameters on the output using three SA methods. The results show that the roughness coefficients of the analysed cross-sections and the next downstream reaches have a significant influence on model output (NS values). Only the results for the gauging station in Warsaw Port Praski do not confirm this finding, probably due to the influence of the downstream boundary condition (the rating curve at Warsaw Port Praski cross-section). The influence of the other parameters is negligible for the first two stations or less important in the case of Puławy and Dęblin gauging stations. The results for the stations located downstream are characterized by a large influence of almost all parameters.

Comparison of the results of the sensitivity measures between the methods used show significant compliance. The differences were found only for parameters with relatively little impact on the values of the coefficients of NS. A comparison of the results of the three methods along the river reach shows that the results from the beginning of the reach are more similar to each other than to those for the cross-sections located downstream as a consequence of uncertainty propagation from the upstream located cross-sections.

Generally speaking, the outcomes from the study indicate that appropriate model calibration strategy should not consist of a series model optimization for independent sub-reaches but rather it should be performed as multiobjective optimization or single objective optimization that integrate the goodness of fit criteria for all control cross-sections.

Acknowledgments This work was supported by the project “Stochastic flood forecasting system (The River Vistula reach from Zawichost to Warsaw)” carried by the Institute of Geophysics, Polish Academy of Sciences, on the order of the National Science Centre (contract No. 2011/01/B/ST10/06866). The water level data were provided by the Institute of Meteorology and Water Management (IMGW), Poland.

References

Ali S, Goodwin P (2002) The predictive ability of 1D models to simulate floodplain processes. In: Falconer RA, Lin B, Harris EL, Wilson CAME (eds) Proceedings of the fifth international conference on hydroinformatics. Model development and data management. IWA Publishing, London, pp 247–252

Archer G, Saltelli A, Sobol I (1997) Sensitivity measures, ANOVA-like techniques and the use of bootstrap. J Stat Comput Simul 58:99–120

Bates PD (2007) Evaluation of inundation models, limits and capabilities of models. Technical report. www.floodsite.net

Bates PD, Roo AD (2000) A simple raster-based model for flood inundation simulation. J Hydrol 236:54–77

Bates PD, Stewart MD, Siggers GB, Smith CN, Hervouet JM, Sellin RHJ (1998) Internal and external validation of a two-dimensional finite element model for river flood simulation. Proc Inst Civ Eng, Water Marit Energy 130:127–141

Campolongo F, Cariboni J, Saltelli A (2007) An effective screening design for sensitivity analysis of large models. Environ Model Softw 22:1509–1518. doi:10.1016/j.envsoft.2006.10.004

Campolongo F, Saltelli A, Cariboni J (2011) From screening to quantitative sensitivity analysis. A unified approach. Comput Phys Commun 182:978–988

Chow V (1959) Open-channel hydraulics. McGraw-Hill Book Company, New York

Cloke HL, Pappenberger F, Renaud J-P (2008) Multi-method global sensitivity analysis (MMGSA) for modelling floodplain hydrological processes. Hydrol Process 22:1660–1674. doi:10.1002/hyp.6734

Feldhaus R, Hottges J, Brockhaus T, Rouve G (1992) Finite element simulation of flow and pollution transport applied to a part of the river Rhine. In: Hydraulic and environmental modelling: estuarine and river waters. Ashgate publishing, Aldershot, pp 323–334

Gasiorowski D, Napiorkowski JJ, Szymkiewicz R (2015) One dimensional modeling of flows in open channels. In: Rowiński PM, Radecki-Pawlik A (eds) Rivers—physical, fluvial and environmental processes. GeoPlanet Series. Springer, Berlin

Gee D, Anderson M, Baird L (1990) Large scale floodplain modelling. Earth Surf Proc Land 15:512–523

Hall JW, Tarantola S, Bates PD, Horritt MS (2005) Distributed sensitivity analysis of flood inundation model calibration. J Hydraul Eng 131(2):117–126

Hall J, Boyce S, Wang Y, Dawson R, Tarantola S, Saltelli A (2009) Sensitivity analysis for hydraulic models. J Hydraul Eng 135(11):959–969

He MX, Hogue TS, Franz KJ, Margulis SA, Vrugt JA (2011) Characterizing parameter sensitivity and uncertainty for a snow model across hydroclimatic regimes. Adv Water Resour 34:114–127. doi:10.1016/j.advwatres.2010.10.002

Herman JD, Kollat JB, Reed PM, Wagener T (2013) Technical note: method of Morris effectively reduces the computational demands of global sensitivity analysis for distributed watershed models. Hydrol Earth Syst Sci 17(7):2893–2903. doi:10.5194/hess-17-2893-2013

Homma T, Saltelli A (1996) Importance measures in global sensitivity analysis of nonlinear models. Reliab Eng Syst Safe 52:1–17

Hunter NM, Bates PD, Horrit MS, Wilson MD (2007) Simple spatially-distributed models for predicting flood inundation: a review. Geomorphology 90:208–225

Kiczko A, Romanowicz RJ, Napiórkowski JJ (2008) Sensitivity and uncertainty analysis applied to water management problem: Upper Narew case study. Publs Inst Geophys Pol Acad Sci E-9(405):75–87

Kochanek K, Karamuz E, Osuch M (2015) Distributed modelling of flow in the middle reach of the River Vistula. This book, chapter 6

Knight D, Shiono K (1996) River channel and floodplain hydraulics. In: Floodplain processes. Wiley, Chichester, pp 139–182

Liu Y, Sun F (2010) Sensitivity analysis and automatic calibration of a rainfall-runoff model using multi-objectives. Ecol Inf 5:304–310

Madsen H, Rosbjerg D, Damgard J, Hansen FS (2003) Data assimilation in the MIKE 11 flood forecasting system using Kalman filtering. In: Water re-sources systems-hydrological risk, management and development, proceedings of symposium HS02b held during IUGG2003 at Sapporo, vol 281, July 2003, IAHS Publ., pp 75–81

Mahmood K, Yevjevich V, Miller W (1975) Unsteady flow in open channels. Water Resources Publications, Colorado

Morris MD (1991) Factorial sampling plans for preliminary computational experiments. Technometrics 33(2):161–174

Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models 1: discussion of principles. J Hydrol 10:282–290

Pappenberger F, Matgen P, Beven KJ, Henry JB, Pfister L, de Fraipont P (2006) Influence of uncertain boundary conditions and model structure on flood inundation predictions. Adv Water Resour 29:1430–1449

Pappenberger F, Beven KJ, Ratto M, Matgen P (2008) Multi-method global sensitivity analysis of flood inundation models. Adv Water Resour 31:1–14. doi:10.1016/j.advwatres.2007.04.009

Romanowicz RJ, Macdonald R (2005) Modelling uncertainty and variability in environmental systems. Acta Geophysica Polonica 53:401–417

Romanowicz RJ, Osuch M (2015) Stochastic semi-distributed flood forecasting system for the Middle Vistula reach, This book, chapter 9

Ruano MV, Ribes J, Seco J, Ferrer J (2012) An improved sampling strategy based on trajectory design for application of the Morris method to systems with many input factors. Environ Model Softw 37:103–109

Saltelli A (2002) Making best use of model evaluations to compute sensitivity indices. Comput Phys Commun 145:280–297

Saltelli A, Tarantola S, Campolongo F, Ratto M (2004) Sensitivity analysis in practice: a guide to assessing scientific models. Wiley, New York

Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S (2008) Global sensitivity analysis: the primer. Wiley Online Library, New York

Sobol I (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math Comput Simul 55:271–280

Song X, Zhang J, Zhan C, Xuan Y, Ye M, Xu C (2015) Global sensitivity analysis in hydrological modeling: review of concepts, methods, theoretical framework, and applications. J Hydrol 523:739–757

Szymkiewicz R (2000) Modelowanie matematyczne przepływów w kanałach otwartych. PWN, Warszawa

Zhan C, Song Z, Xia J, Tong C (2013) An efficient integrated approach for global sensitivity analysis of hydrological model parameters. Environ Model Softw 41:39–52

Influence on Flood Safety of Channel Processes and Vegetation in the River Vistula Valley in Warsaw

Artur Magnuszewski

Abstract The influence of floodplain plant communities on flood conveyance has been estimated using a two-dimensional hydrodynamic model CCHE2D to model flow in the contemporary River Vistula channel in Warsaw. In the study the channel reach between 501 and 521 km limited by flood protective dikes has been analyzed. The influence of vegetation on flow resistance has been estimated by applying two scenarios of flood wave passage. The first scenario "0" describes the flow with existing dense vegetation and scenario "1" simulates the situation in which tall trees and dense shrubs are cleared at the right river bank between 509 and 519 km. Simulations have been run for the discharges $Q_{1\ \%} = 6430\ m^3\ s^{-1}$ and $Q_{0.1\ \%} = 8250\ m^3\ s^{-1}$. Obtained results show that the largest, 0.56 m decrease of the water surface after the trees cutting occurs at 510 km and discharge $Q_{0.1\ \%}$. Application of 2D hydrodynamic model for different scenarios of the floodplain maintenance helps to predict the flood conveyance in highly urbanized and vulnerable area of the city of Warsaw.

1 Introduction

One of the standard techniques of the flood protection is the construction of dikes which limit the area of inundated floodplain. This is a relatively inexpensive method of reducing the flood risk, but has also a number of negative

A. Magnuszewski (✉)
Hydrology Department, Faculty of Geography and Regional Studies,
University of Warsaw, Krakowskie Przedmieście 30, 00-927 Warsaw, Poland
e-mail: asmagnus@uw.edu.pl

R.J. Romanowicz and M. Osuch (eds.), *Stochastic Flood Forecasting System*,
GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-319-18854-6_8

effects. Smith and Winkley (1996) have shown that in the Mississippi river the narrowing of the corridor for flood passage has increased the channel erosion and raised the need for river training works. Regulation of the river in sake of the dikes protection led to an increase in river flow velocities, causing deepening of the channel and growth of a new floodplain. Such a process is visible also, for example, at River Elbe which has been regulated at the beginning of 19-th century (Bart et al. 2009). A similar process is observed at the regulated reach of the lower River Vistula where new floodplain was built below the Włocławek dam (Babiński 2004).

The new floodplain created between the regulated channel banks and the dike foot slopes is a place of intensive growth of plants. The vegetation in the flood corridor can have positive influence by slowing down flood wave propagation in the river valley and enhancing sedimentation at the floodplain, but also negative influence by increasing the water level during the flood which affects the risk of inundation of the area protected by dikes (Baptist 2004). In the flood management policy we have to answer the question which of these processes is more important at a given location. On the one hand, it is important to decrease the dynamics of flood but also to reduce the risk of flood caused by higher water stages. This problem is visible in case of the River Vistula valley in Warsaw. Dense vegetation which has developed on a new floodplain provides important ecological services but is disputable when we think about the flood safety.

First attempts on evaluating the influence of riparian vegetation on river hydraulics have been undertaken by Cowan (1956) and Chow (1959). They have been using different Manning roughness coefficients representing appropriate types of the land covers. Also physical hydraulic models have been used for studying the velocity field and momentum transfer between the main channel and floodplains covered by vegetation (Pasche and Rouve 1985). A new approach is the application of hydrodynamic models, which makes possible simulation of the flood passage through the given reach of the river valley assuming different vegetation cover scenarios.

At the floodplain there is a belt of accelerated flow caused by the momentum transfer from the main channel; the width of this belt is equal to half of the main channel width. At a longer distance from the channel bank, the water velocities are lower and vegetation cover in this area affects the flood conveyance to a lesser degree (Knight and Shiono 1996).

Leu et al. (2008) have performed numerical experiment to show the optimum vegetation clear cut at the floodplain. Their results have confirmed the fact that the most significant improvement of hydraulic properties is achieved by clearing the floodplain from the trees and shrubs in a parallel belt to the river bank.

The aim of this study is to show the influence of vegetation developed on a newly formed floodplains on flood conveyance properties. Very special conditions of the Warsaw's reach of the River Vistula make a problem of flood conveyance one of the important issues related to the management of the flood risk. Two scenarios are studied using 2D hydrodynamic model CCHE2D. The scenario "0" describes flow with existing dense vegetation and scenario "1"

simulates the situation in which tall trees and dense shrubs are cleared at the right bank between 509 and 519 km. Simulations have been run for the discharges $Q_{1\ \%} = 6430\ m^3\ s^{-1}$ and $Q_{0.1\ \%} = 8250\ m^3\ s^{-1}$.

2 Description of the Study Area

The River Vistula valley in the city of Warsaw has been formed at the end of Pleistocene and has got its recent shape through the Holocene period. The Warta glaciation has formed glacial uplands of Równina Warszawska and Równina Wołomińska, while younger stages of Wkra and Mława glaciations are responsible for shaping the main valley course with the outlet in the proglacial lake occupied today by the lowland of Kotlina Warszawska (Mojski 2005). Between the glacial uplands there have been formed river terraces. At the end of the Vistulian glaciation the higher Pleistocene terraces, called Otwock terrace and Falenica terrace, have been formed. At the beginning of Holocene the over-flood Praga terrace has been formed. The higher flood terrace (Wawer terrace) developed in the early Holocene. It stretches along both banks of the river, forming the widest floodplain. When forming this terrace, Vistula had the pattern of a meandering river. The traces of meander bends are visible as curved undercuts at the edge of overflood (Praga) terrace, for example in the areas of Wawer-Gocław and Powsin. The remnants of the meandering are preserved also as oxbow lakes in Wilanów and Czerniaków districts.

Until almost the end of the 19th century the Warsaw stretch of River Vistula remained in its natural state. Meandering pattern of the river channel has been changing into braiding pattern which creates islands, channel bars and side arms of the river. Initially, only short reaches of the flood protective dikes existed at the left bank (511.5–513.8 km) and the right bank (515.5–518.5 km).

The catastrophic flood of July 1884 gave an impulse for undertaking the regulation of the river and extension of the flood protective dikes. In the years 1923–1931 further regulation work was carried out, with the purpose of protecting the shores and the newly built dikes after the flood of 1924. Immediately after the World War II the river channel of Vistula in Warsaw became the debris dumping ground, so that the cross-sections of the high discharges narrowed down by some 50 % in comparison with the middle Vistula, forming the so-called Warsaw "corset" over the segment of 507–517 km. Such a situation is due to geological structure (the extension of the Praga terrace) and narrow flood protection dikes spacing fitting the bridges active cross-sections. Narrow Warsaw downtown reach of the River Vistula together with river training structures is shown in Fig. 1. Spacing between flood protective dikes in this area is shown in Fig. 2.

The water levels of the River Vistula in Warsaw are observed at the gauge Port Praski, which has been in operation for nearly 200 years. This reach has a drainage area of 84,857 km^2, and is located at the 513 km of the river course, on the right bank of the River Vistula at the inlet canal of the inland harbor. Characteristic

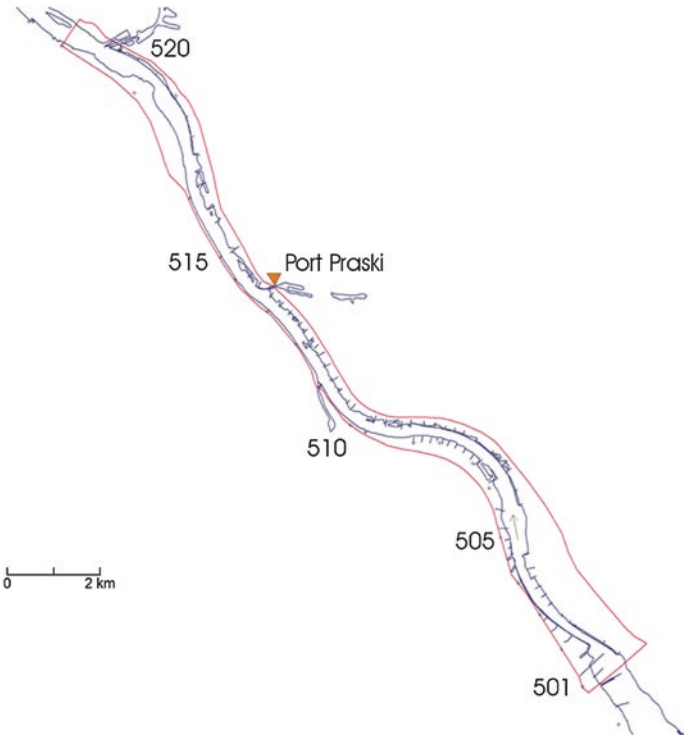

Fig. 1 Plan of the Vistula River with river training structures and boundary of the computational mesh for a CCHE2D model

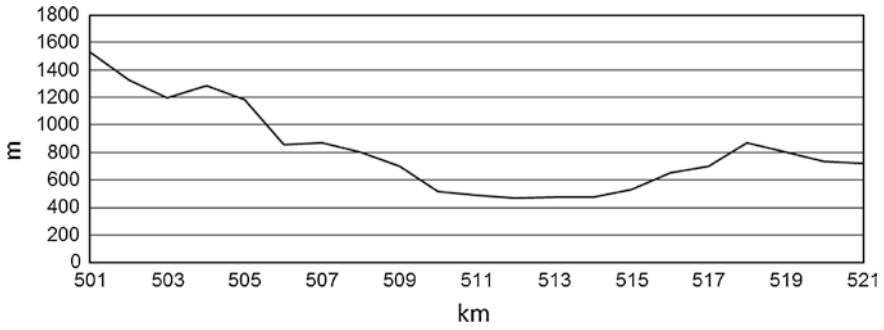

Fig. 2 The distance between the flood protective dikes and boulevards in Warsaw in a longitudinal profile

discharges are: mean low flow $Q_L = 194\ m^3\ s^{-1}$, average flow $Q_M = 561\ m^3\ s^{-1}$, $Q_{1\ \%} = 6430\ m^3\ s^{-1}$, $Q_{0.1\ \%} = 8250\ m^3\ s^{-1}$. Values of the characteristic discharges have been taken according to report by Institute of Meteorology and Water Management IMGW (2006).

In history, winter floods were more frequent and were often caused by the ice jams, although the highest discharges were caused by the intensive rainfalls in summer months. Vistula in its middle reach has a transitional character; it means that the floods created in the upper part of the catchment in the Carpathian Mountains propagate down the lowland flood corridor constrained by the flood protective dikes.

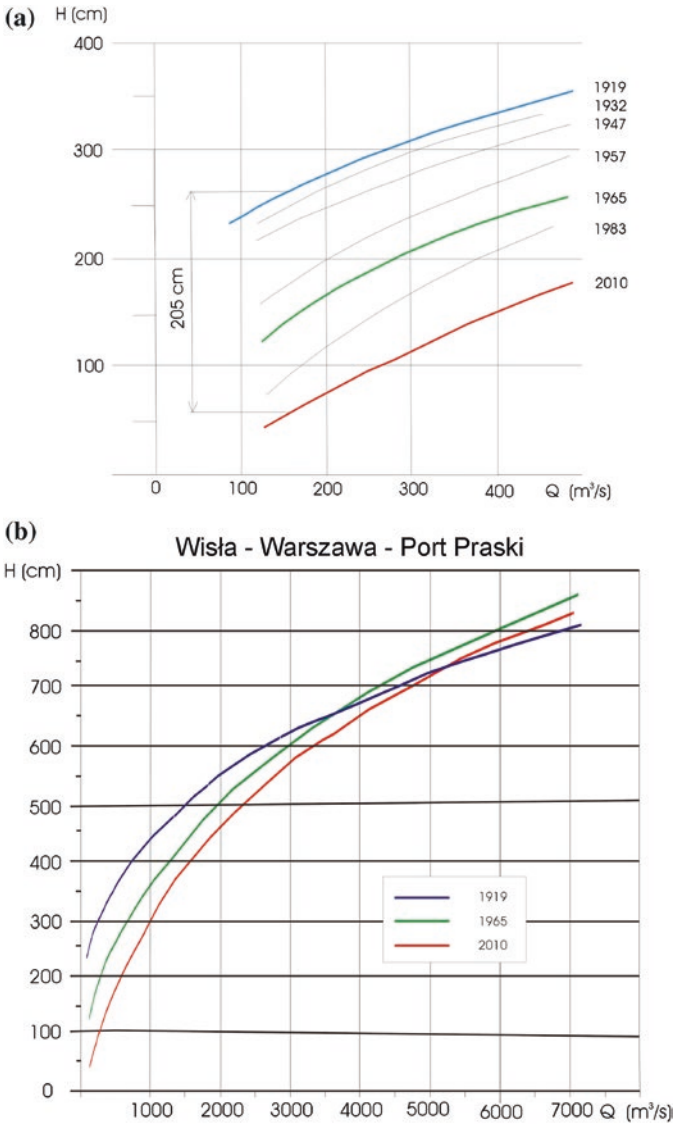

Fig. 3 Rating curves at Warsaw Port-Praski gauge (513.3 km). **a** low flow range; **b** flood range. Prepared as a compilation of results published by Kornacki (1960), Mikulski et al. (1969), Fal and Dąbrowski (2001), Kuźniar and Magnuszewski (2010), Ozga-Zieliński et al. (2010) and hydrodynamic modeling with the use of CCHE2D model

According to the design rules, the dikes should have elevation providing safe passage of the flood wave that has a recurrence of 1000 years ($Q_{0.1\ \%}$). In Warsaw flood protection dikes are spaced by 400–600 m, whilst outside of the city they are 1000–1700 m apart and are lower since they shield agricultural areas against the inundation caused by flood $Q_{1\ \%}$.

The narrowing of the river valley creates the conditions favorable for a channel bottom erosion. The lowering of the river bottom has been observed at the beginning of 20-th century when first rating curves started to be available. Starting from the first rating curve established in 1919 up to 2010, the bottom of the river has dropped by 2.05 m (Fig. 3a).

Additional processes responsible for the river bottom erosion are the dredging works, performed for the sake of obtaining proper depths for the navigation. The first dredging works for the navigation purpose have been documented in 1923.

Erosion has accelerated after the war when large volumes of sand were excavated from the river channel for the process of reconstruction of rapidly growing city. In the year 1950, about 500,000 m^3 of sand has been excavated from the River Vistula channel in Warsaw. This volume is equal to the amount of the bedload transported by the river from the upper reach of the catchment (Skibiński 1963). In the years 1952–59, the volume of dredged sand has increased to 850,000 m^3 annually, reaching almost 2,000,000 m^3 in 1978.

Fig. 4 Aerial photo of the Vistula River in Warsaw showing "corset" and new floodplain with the dense vegetation in the 513–514 km reach (photo by Marek Ostrowski)

Comparison of the curves (Fig. 3b) shows that in a zone of the highest stages the rating curve of 2010 is lower than the curve of 1960 but still higher than the initial curve of 1919. This pattern indicates that the initial flood conveyance has declined. That can be explained by the influence of the river training works from the 1960, as well as the subsequent channel erosion. The first reaction of the channel to the river training works was the development of a new floodplain between the groins (Fig. 4). Erosion processes in the channel are reflected by lower position of the newest rating curve.

3 Method Description

To study the influence of vegetation on flood conveyance in the Vistula channel in Warsaw, the hydrodynamic 2D model called CCHE2D, developed at the National Center for Computational Hydroscience and Engineering (NCCHE) at the University of Mississippi, USA, was applied. The model simulates free surface flow, and is based on the depth-averaged Navier-Stokes equations. Turbulent shear stresses were estimated using Boussinesq approximation, with the introduction of turbulent eddy viscosity. The set of equations is solved implicitly by the control volume method. Model description and examples of application are given by Altinakar et al. (2005).

In the CCHE2D model, numerical calculations are carried out at the nodes of an irregular rectangular mesh. There is a separate subroutine for mesh design and interpolation of channel geometry from DEM as well as selection of Manning roughness values. The geometric information was obtained from the DEM of 20 m resolution. DEM contains 78,600 x, y, z coordinates.

The data for DEM have been combined from a river soundings and irregular elevation points obtained from digital stereoscopy representing the floodplain and higher terraces. The Manning roughness coefficient n was assigned to land-use classes according to the literature values: 0.03 for river channel, 0.04 for open area with a low vegetation, 0.06 for settlement area, and 0.11 for forest and tall vegetation. The spatial distribution of the values was obtained from recent topographic maps.

Computing mesh has been designed for a reach of the channel between 501 and 521 km limited by the flood protective dikes. Mesh with the resolution of $i = 50$ by $j = 400$ lines has been obtained. Spacing along i lines is in the range of 20–30 m, whilst along j lines it is 40–50 m. The upper and lower boundary conditions selected for calculations are given at Table 1.

In the procedure of model verification, a steady flow has been set as the upper boundary condition and the discharge corresponding to the 1997 year flood, $Q_{1997} = 4170\ m^3\ s^{-1}$, for which a longitudinal profile of the water level during the flood peak flow has been measured. Field measurements of the water level elevation have been done by OPGK-Lublin geodetic survey company. Model results versus water level measurements show good fit 501–505 km and 515–521 km (Fig. 5).

Table 1 Values of the upper and lower boundary conditions of the CCHE2D model of the River Vistula reach between 501 and 521 km

Case	Upper boundary condition Q (m^3 s^{-1})	Lower boundary condition z (m a.s.l.)
1997	4170	80.35
1 %	6430	82.32
0.1 %	8250	83.30

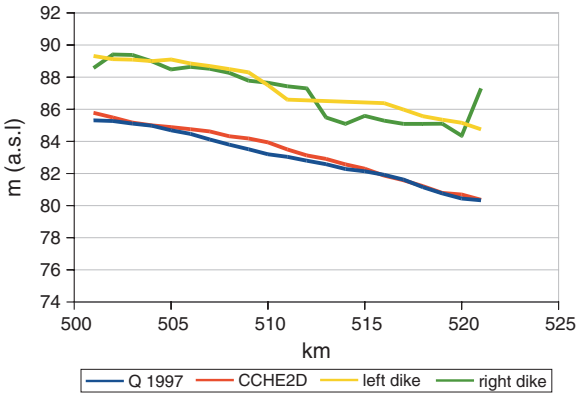

Fig. 5 Model verification results—comparison of the water surface elevation from measurements during the 1997 flood (Q 1997) at the Vistula River in Warsaw and calculated by CCHE2D model

The average difference in the water elevation between observed water levels and model calculation is 0.26 m. The difference can be explained by the fact that the water level measurements have been done during the passage of flood with unsteady flow characteristic, whilst model represented steady flow conditions. In 1D models it is a common practice to calibrate the model by the change of Manning roughness coefficients; in this study, those values have not been altered for tuning the model.

Additionally, the model was validated at Q = 3059 m^3 s^{-1}, by comparison of calculated velocity to real current meter measurements performed by the Institute of Meteorology and Water Management (IMGW) during the flood of 10 June 2006. The value of maximum vertical velocity measured at the profile of Świętokrzyski Bridge by a current-meter was 2.15 m s^{-1}, while that calculated by the CCHE2D model was 2.36 m s^{-1}.

The influence of trees and shrubs cut at the right bank of a new floodplain in the reach of 509–519 km has been simulated by changing the Manning roughness coefficient from 0.120 to 0.045. This corresponds to change (clear cut) of the dense floodplain forest (scenario 0) to meadows and park area (scenario 1). Figure 6a shows the distribution of Manning roughness coefficient in the natural state and Fig. 6b a scenario of the clear cut at 509–519 km at the right bank.

The area of the dense forest in the natural state (scenario 0) was 233.6 ha, while the area of the forest removed by the clear-cut (scenario 1) was 89.1 ha. This makes the lowering of the forest vegetation area by 38.1 %. The valley of the

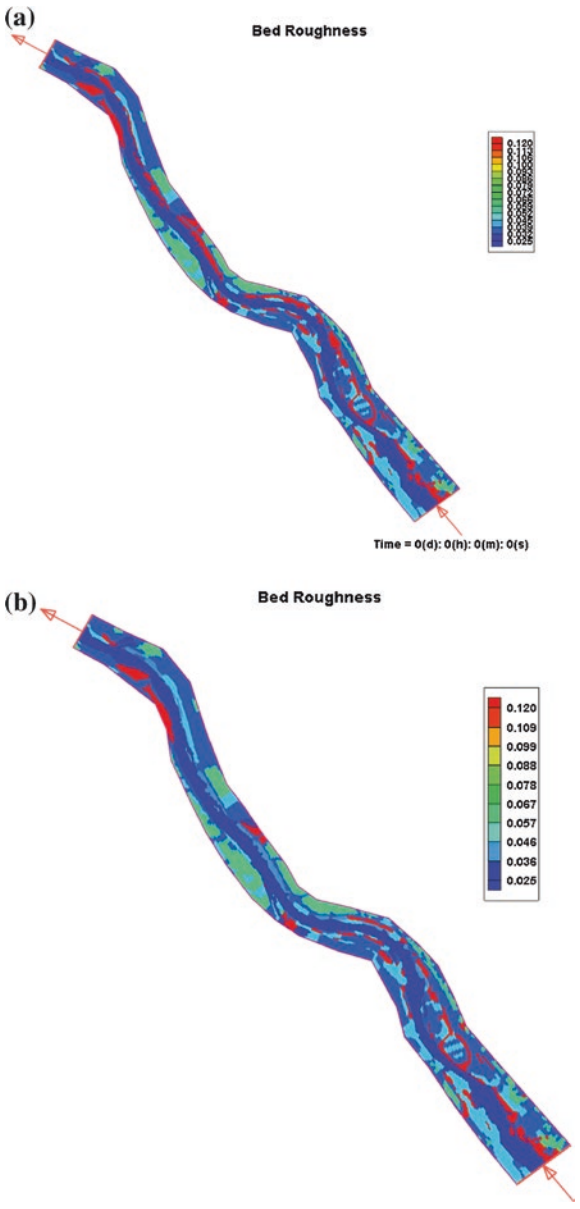

Fig. 6 **a**. Distribution of the Manning roughness coefficient in complete vegetation cover (scenario 0). **b**. Distribution of the Manning roughness coefficient with a clear cut of the tall plants on the right bank between 509 and 519 km (scenario 1)

River Vistula limited by the flood protective dikes between the towns of Płock and Dęblin belongs to the Natura 2000 "Special Protection Area".

Alluvial soils had developed on the youngest flood terrace, which are some of the richest soils in Warsaw. Most of the alluvial areas within Warsaw are covered by vegetation that resembles the natural plant communities. Except of the

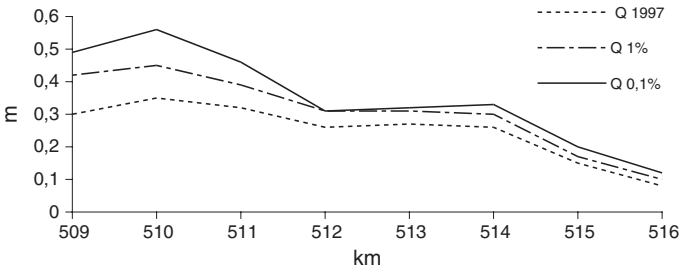

Fig. 7 Lowering of the water surface elevation at discharges of Q_{1997}, $Q_{1\ \%}$, $Q_{0.1\ \%}$ in a scenario 1 of tall plants clear-cut at the right bank floodplain between 509 and 519 km

river training structures and re-enforced sections and the areas within the vicinity of the sewers outlets, a pronounced vegetation zonation occurs on the river banks throughout the city. Species from the Class *Isoëto-Nanojuncetea* and *Bidentetea tripatiti* are found closest to the river channel. *Salix* scrub and *Salix-Populus* forests, which are better preserved near the northern and southern boundaries of the city, occur further away from the river (Sudnik-Wójcikowska and Galera 2011). The forest complex belongs to *Salici-Populetum* which according to Nature 2000 has the symbol 91E0—alluvial forests.

Calculations for two defined scenarios have been done assuming discharge of Q_{1997}, $Q_{1\ \%}$ and $Q_{0.1\ \%}$ with a steady flow conditions. The lowering of the water surface elevation due to a tall vegetation clear-cut is shown at Fig. 7.

4 Discussion of Results

Application of the 2D hydrodynamic model to study the influence of vegetation on flood conveyance makes it possible to design proper maintenance works, and select the optimal solution. This class of the models reflects the process of momentum transfer between the main channel and floodplain with a vegetation cover.

The obtained results indicate that by performing the clear cut of the tall vegetation at the right bank of the River Vistula floodplain in the city of Warsaw it is possible to lower the $Q_{0.1\ \%}$ water surface elevation below the crest of the flood protective dikes (Fig. 8). This vulnerable area is located at the right bank between 513–514 km and is difficult to protect against the flood, due to the fact that on top of the dike there is a major street with transit traffic.

The maximum decrease of the water level due to tall plants clear cut is observed at the entrance to the Warsaw "corset"—near 510 km. This is an important area which has to be well protected from flooding because in the case of dike breaching a vast, densely populated area of Warsaw could be inundated.

The specific conditions of the River Vistula channel in the city of Warsaw require special care to provide enough space for flood conveyance. Erosion of

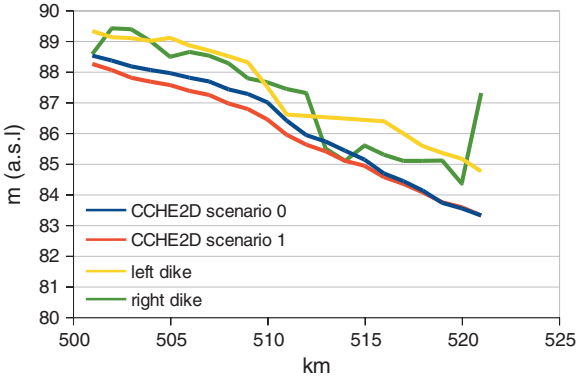

Fig. 8 Water surface elevation at discharge $Q_{0,1\ \%}$ in a scenario 0 of natural forest communities at the right bank floodplain, 509–519 km, and scenario 1 representing clear-cut of the high plants

the river channel induced by the limited spacing between flood protective dikes and river training works improves flood conveyance but enhances also the growth of a new floodplain with an environment favorable for the development of dense vegetation.

The presented results are important for understanding and evaluating the role of dense vegetation cover for decreasing the flood risk in the downtown of Warsaw. The alluvial forest in the city has an ecological importance for Nature 2000 network but in the context of flood conveyance in the Warsaw "corset" it should be re-evaluated.

Acknowledgments This paper presents the results obtained from the project: Spatial Differentiation of Information Society Vulnerability to Defined Environmental Hazards in Poland, supported by the Polish National Science Centre (DEC-2011/03/B/HS4/04933, 2012–2015). Warsaw Regional Water Authority (RZGW) has kindly provided bathymetric and DEM data. I am grateful to professor Mustafa Altinakar and Dr Yaoxin Zhang (National Center for Computational Hydroscience and Engineering, School of Engineering, The University of Mississippi) for permission to use the CCHE2D model.

References

Altinakar MS, Czernuszenko W, Rowiński P, Wang SY (ed) (2005) Computational modeling for the development of sustainable water resources systems in poland. Publ Inst Geophys Pol Acad Sci E-5(387):239–260

Baptist MJ (2004) A commentary on advances in modelling the effects of vegetation on flow, sediment transport and morphology. In: Petts G, Kennedy B (ed) Emerging concepts for integrating human and environmental water needs in river basin management. Report of a workshop sponsored by the USACE as part of the Water Operations and Technical Support Programme, 16-18 Sept, The University of Birmingham, United Kingdom, pp 72–76

Babiński Z (2004) Antropogeniczne przekształcenia dna doliny Dolnej Wisły i jej renaturyzacja w świetle prac hydrotechnicznych. Dokumentacja Geograficzna 31:9–10

Bart S, Schumberg S, Deutsch M (2009) Revising time series of the Elbe river discharge for flood frequency determination at gauge Dresden. Nat Hazards Earth Syst Sci 9:1805–1814

Chow VT (1959) Open channel hydraulics. McGraw-Hill, New York

Cowan WL (1956) Estimating hydraulic roughness coefficients. Agricul Eng 37(7):473–475
Fal B, Dąbrowski P (2001) Dwieście lat obserwacji i pomiarów hydrologicznych Wisły w Warszawie: Przepływy Wisły w Warszawie. Gospodarka Wodna 12:503–510
IMGW (2006) Wpływ roślinności porastającej międzywale rzeki Wisły w rejonie Warszawy na odcinku od km 504,1 do km 513,3 na zagrożenia powodziowe. Instytut Meteorologii i Gospodarki Wodnej, Warszawa (unpublished report)
Knight DW, Shiono K (1996) River channel and floodplain hydraulics. In: Anderson MG, Walling DE, Bates PD (eds) Floodplain processes. Wiley, Chichester
Kornacki Z (1960) Przyczyny obniżania się dna Wisły w Warszawie. Gospodarka Wodna 7:305–307
Kuźniar P, Magnuszewski A (2010) Przepływ wód wielkich Wisły w Warszawie—re-konstrukcja powodzi historycznych. In: Magnuszewski A (ed) Hydrologia w ochronie i kształtowaniu środowiska. Monografie Komitetu Inżynierii Środowiska PAN, nr 69, pp 109–118
Leu JM, Chan HC, Yafei Jia, Zhiguo He, Wang Sam SY (2008) Cutting management of riparian vegetation by using hydrodynamic model simulations. Adv Water Resour 31:1299–1308
Mikulski Z, Skibiński J, Żbikowski A (1969) Wpływ rumowiska na pracę ujęć poddennych wodociągów miejskich m. st. Warszawy przed i po spiętrzeniu rzeki Wisły. Prace i Materiały Techniczno-Ekonomiczne Rady Naukowej przy Prezydium Stołecznej Rady Narodowej. Sekcja 4, nr 37, s 44
Mojski JE (2005) Ziemie polskie w czwartorzędzie. PIG, Warszawa, pp 1–404
Ozga-Zieliński B, Szkutnicki J, Kadłubowski A, Chudy Ł (2010) Wisła w Warszawie—wybrane problemy hydrologiczne. Gospodarka Wodna 12:490–493
Pasche E, Rouve G (1985) Overbank flow with vegetatively roughened floodplains. J Hydraul Eng 111(9):1262–1278
Skibiński J (1963) Wleczenie rumowiska dennego przez Wisłę w rejonie Warszawy, Wiadomości Służby Hydrologicznej i Meteorologicznej, z 53
Smith LM, Winkley BR (1996) The response of the Lower Mississippi River to river to river engineering. Eng Geol 45:433–455
Sudnik-Wójcikowska B, Galera H (2011) Warsaw. In: Plants and habitats of European Cities. Springer, New York

Part III
Lumped Parameter Approximation to Distributed Modelling of River Flow, Distributed Model Emulators and On-line Data Assimilation

Stochastic Semi-distributed Flood Forecasting System for the Middle Vistula Reach

R.J. Romanowicz and M. Osuch

Abstract The aim of the study is development of an integrated flood forecasting system for the middle reach of the River Vistula between Zawichost and Warsaw Port Praski. It has a modular structure, consisting of flood wave transformation models for the main river reach. The system predictions are based on water level observations, as the most robust information on the system state. We describe the system structure and we show that three-day ahead predictions of water levels at Warsaw Port Praski reach of the river can be obtained using a simple, lumped parameter Stochastic Transfer Function (STF) model. Apart from a single lumped model also serially connected Single Input Single Output (SISO) and Multiple Inputs Single Output (MISO) STF models describing river sub-reaches situated between the gauging stations have been derived. In this way, the system can have either lumped or semi-distributed structures. Due to the stochastic nature of the methods applied the confidence limits of the predictions are also estimated. The ability to tackle the nonlinearity of the flood wave transformation process by means of the Hammerstein-Wiener modelling approach is studied.

Keywords Water levels predictions · Stochastic transfer function · Flood forecasting system · River vistula · Box-Cox transformation · Hammerstein-Wiener model

R.J. Romanowicz (✉) · M. Osuch
Institute of Geophysics Polish Academy of Sciences,
Ks. Janusza 64, 01-452 Warsaw, Poland
e-mail: Romanowicz@igf.edu.pl

M. Osuch
e-mail: marz@igf.edu.pl

R.J. Romanowicz and M. Osuch (eds.), *Stochastic Flood Forecasting System*,
GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-319-18854-6_9

1 Introduction

Hydrological models used in the national operating systems for the purpose of short-term flow forecasting are based on empirical relationships between water level observations at neighbouring gauging stations along the river, simplified open-channel flow routing models and hydrodynamic models (e.g. MIKE 11, IHMS HBV-HD).

The recent big floods in Poland, in 1997 and 2010, promoted a discussion on the issue of flood forecasting. The 2010 year flood was one of the largest since the 19th century in Poland. The second flood-wave followed the first only 10 days later, with high water levels along the middle and lower River Vistula reaches. Very large flood damages incurred showed that that the current flood warning and defence system is not efficient. In addition, it does not take into account uncertainty of hydrological forecasts and does not provide appropriate tools to help the decision makers to manage the risk of flooding.

The flood alert system plays an important role in decreasing the flood-risk. After the disastrous flood in 1997, the existing system of Hydrological Forecasting for the whole of Poland was upgraded and modernized as a component of the Emergency Flood Recovery Project, financed by the World Bank. The latter has the form of a System of Monitoring, Forecasting and Warning (SMOK), transformed into the System of Hydrology (SH), developed in the Institute of Meteorology and Water Management (IMGW). The operational hydrological models used for short-term flood forecasting include simplified models of flood wave transformation in the river channel and hydrodynamic models, such as MIKE 11 and IHMS HBV-HD (Kadłubowski 2005). The 2001 and 2010 floods of the Vistula showed that un-gauged tributaries played an important role in flood wave transformation along the river. This indicates that the meteorological forecast might be of major importance in predicting the evolution of a flood wave along the river.

In contrast to the purely deterministic forecasts obtained in the forecasting system currently operating on the territory of Poland, we propose a stochastic forecasting system, which gives the most probable water level predictions together with their 95 % confidence intervals. Knowledge of reliability of water level estimates is very important in determining the risk of flooding required by the flood warning system (Romanowicz et al. 2006a) and is a necessary condition of a feasibility of a flood forecasting system in Europe (Thielen et al. 2007).

Distributed hydrodynamic models are used to describe flood wave transformation in an open channel (Chow 1959; Mahmood et al. 1975; Abbott 1979; Cunge et al. 1980; Madsen et al. 2003; Szymkiewicz 2010), models based on neural networks (Park et al. 2005) and non-linear time series models (Porporato and Ridolfi 2001). The most popular among hydrodynamic models are: ISIS, Mike11, Lisflood, HEC-RAS and Sobek (e.g. Huizinga et al. 2005). However, these models require long computing times that increase when high accuracy of predictions is required. Usually this is not a problem when single analysis of the modelled

system is being performed, but may be troublesome when multiple runs of the model are required, e.g. during optimization of the model parameters or when estimating model prediction uncertainty using Monte Carlo based methods. In some practical applications, for example, flood forecasting in real-time, simplified methods provide equally good results with much lower computing costs (Romanowicz et al. 2006a). In this work, an alternative modelling approach that applies a discrete Stochastic Transfer Function (STF) model is followed, identified only on the basis of the observed time series of water levels. The modelled river reach is divided into segments according to the location of gauging stations along the river. For each segment, a lumped parameter STF model is identified, with parameters which may have physical interpretation, and therefore they are called grey box models. The methods applied in this work for the middle reach of the River Vistula are an application and extension of methods described in Romanowicz et al. (2006a, b, 2010, 2014).

The stochastic flood forecasting system presented in this work has a modular structure and consists of in-series connected STF models, that describe transformation of a flood wave between the gauging stations. In view of the experience of one of the authors, in the models, water levels instead of flows are used as driving variables (Romanowicz et al. 2006a). As water levels rather than flows are consistently measured, this approach allows errors associated with nonlinear transformation of water levels into flows (rating curve) to be avoided. The approach adopted allows for the estimation of parametric uncertainty of the STF models and the uncertainty of complete system predictions to be assessed.

The river reach between Zawichost and Warsaw Port Praski, used as a case study, is presented in Sect. 2. In Sect. 3 the research methods are presented at a glance and estimated STF models for individual parts of the river between gauging stations are given. Section 4 gives a short summary of parametric uncertainty analysis of the STF models and the propagation of uncertainty in the integrated system. Section 5 contains a summary.

2 The River Vistula Reach Between Zawichost and Warsaw Port Praski

Stochastic flood simulation system has been designed for the middle River Vistula reach between Zawichost and Warsaw Port Praski gauging stations (Fig. 1). Along that river reach, hydrological monitoring is carried out at 7 gauging stations: Zawichost, Annopol, Pulawy, Deblin, Gusin, Nadwilanówka and Warsaw Port Praski. As shown in Fig. 1 there are ten tributaries within the studied reach. Because of the substantial differences in size and flows, only three largest tributaries were taken into account: Kamienna, Wieprz and Pilica. The description of the rivers is given by (Osuch 2015, this book).

In order to extend the lead time of the forecast, the system was extended by the additional STF modules based on the available observations up-stream of

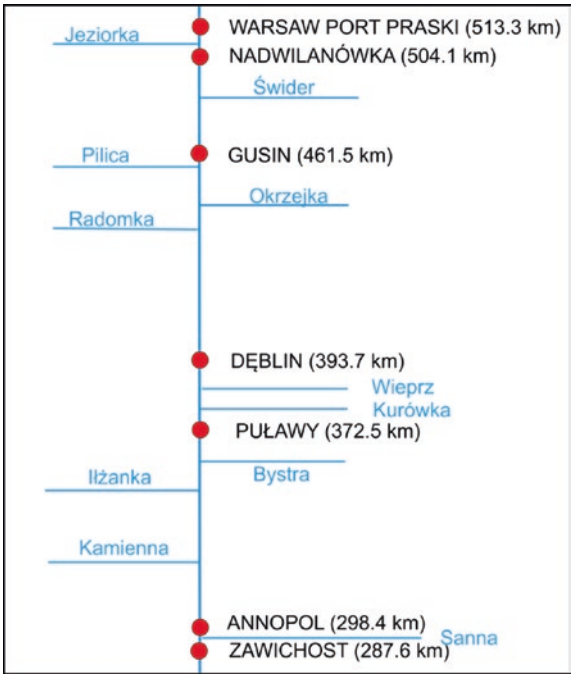

Fig. 1 Diagram of the analysed reach of the River Vistula between Zawichost and Warsaw Port Praski

Zawichost at gauging stations of Kraków (69.2 km) and Szczucin (194.1 km). The geomorphology of the river reach is typical of a lowland river, with a wide channel situated on moving sandy bed. Therefore, the channel is prone to changes of its course during high flows and, in particular, during the ice-melt induced floods.

3 Methods

3.1 Modelling the Transformation of Flood Wave Using the STF Approach Based on Water Levels

The Stochastic Transfer Function (STF) model was applied to describe the transformation of flood wave between Zawichost and Warsaw Port Praski gauging stations. It is a lumped grey-box modelling approach based on observations that follows the Data Based Mechanistic (DBM) philosophy (Young 2001). In that approach, a model structure is identified based on available observations using information-efficient statistical tools. Among the statistically feasible models only those that can be explained in a physically meaningful way are chosen. Depending on the modelled system complexity, transfer function model can take many forms. In this work, the transformation flood wave between Zawichost and

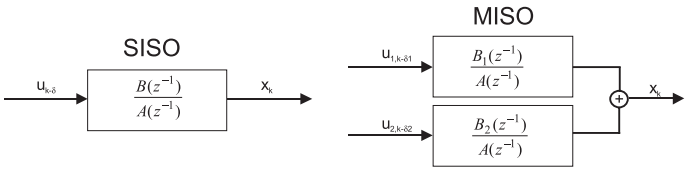

Fig. 2 The schematic representation of two types of the STF models, SISO (*left panel*) and MISO (*right panel*)

Warsaw Port Praski was described using two types of STF models, with a single input (Single Input Single Output) SISO and with two inputs (Multiple Input Single Output) MISO models, as presented in Fig. 2.

The discrete-time SISO type STF model can be presented as (Young 1984):

$$\begin{aligned} x_k &= \frac{B(z^{-1})}{A(z^{-1})} u_{k-\delta} \\ y_k &= x_k + \xi_k \end{aligned} \tag{1}$$

where $u_{k-\delta_1}$ denotes STF model input at time period k (upstream water level), x_k is the underlying 'true' water level, y_k is the noisy observation of this variable, δ is a pure, advective time delay of $\delta \Delta t$ time units, $A(z^{-1})$ and $B(z^{-1})$ are polynomials of the transfer function. These polynomials are in the form:

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}; \quad \mathrm{B}(z^{-1}) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_m z^{-m} \tag{2}$$

in which z^{-p} is the backward shift operator, i.e. $z^{-p} y_k = y_{k-p}$, and $A(z^{-1})$ is assumed to have real roots (eigenvalues) that lie within the unit circle of the complex z plane. Generally, the additive observation noise ξ_k in (Eq. 1) can be both heteroscedastic (i.e. its variance changes over time) and autocorrelated. In this study the noise was assumed to be white.

It is assumed that the noise term includes all the model and input uncertainty at the output of the system, i.e., measurement noise, unmeasured inputs, and uncertainties associated with the model structure. The orders of the polynomials, n and m, are identified from the data during the identification process. The model structure identification and estimation of parameters are performed using the Simplified Refined Instrumental Variable (SRIV) method from Captain toolbox (Taylor et al. 2007). In the approach we follow, it is assumed that input is deterministic. Therefore, in order to take into account the input uncertainty we have to apply the rules related to the transformation of the input noise through a linear system (Romanowicz and Young 2003). In order to take into account the noise heteroscedacity, the transformation of model variables is required (Box and Cox 1964). The result of model structure identification for the SISO models is given in the form of three values [n m δ] denoting the length of polynomials A, B, and pure advective delay, respectively.

Model calibration is performed using the sum of square errors of one step ahead predictions, the so-called R_T^2 criterion. It is defined as a ratio of model output variance explained by the model. It takes values between 0 and 1 (0–100 %). The higher the R_T^2 value the better the fit of the model results to observations. Following the DBM ideology (Young 2001), the STF models applied in our study should have physical interpretation. Therefore, the analysis of model parameters provides an insight into the time constant and the advective delay of the flow transformation process (Young 2002). The time constant of the first order approximation of the model dynamics has a form:

$$T = \frac{-\Delta t}{\log(a_1)} \tag{3}$$

where Δt denotes model discretisation time and a_1 is a coefficient of the polynomial (Eq. 2).

In order to take into account the influence of tributaries, the STF models with multiple inputs can be applied, the so-called Multi Input Single Output (MISO) models. In the case of a model with two inputs, a model equation has the following form:

$$x_k = \frac{B_1(z^{-1})}{A_1(z^{-1})} u_{1,k-\delta_1} + \frac{B_2(z^{-1})}{A_2(z^{-1})} u_{2,k-\delta_2} \tag{4}$$

where $u_{1,k-\delta_1}$ denotes water levels measured at the gauging station situated on main river channel with an advective delay δ_1 and $u_{2,k-\delta_2}$ denotes the tributary input (here water level) with an advective delay δ_2.

Identification of the model structure and estimation of its parameters is performed in a similar way as in the case of the SISO models. The results of model structure identification for the MISO models with two input variables are given in the form of five values [n m_1 m_2 δ_1 δ_2] denoting the length of polynomials A, B_1, B_2 and pure advective delays, respectively.

3.2 The Lumped-Parameter Linear SISO Model for the Zawichost-Warsaw Port Praski Reach

In the first instance the entire reach between Zawichost and Warsaw Port Praski was modelled by a single SISO STF. Daily observed water levels at the Zawichost gauging station served as input variable and the water levels at Warsaw gauging station were the simulated output. The best estimated STF model had a form [1 2 3], with 3-day advective delay and differentiated input. The final model equation has the form:

$$x_k = \frac{0.8913 - 0.8562z^{-1}}{1 - 0.9563z^{-1}} u_{k-3} \tag{5}$$

where u_{k-3} denotes water levels at the gauging station in Zawichost in time period $k-3$, and x_k denotes noise-free model prediction (Eq. 1).

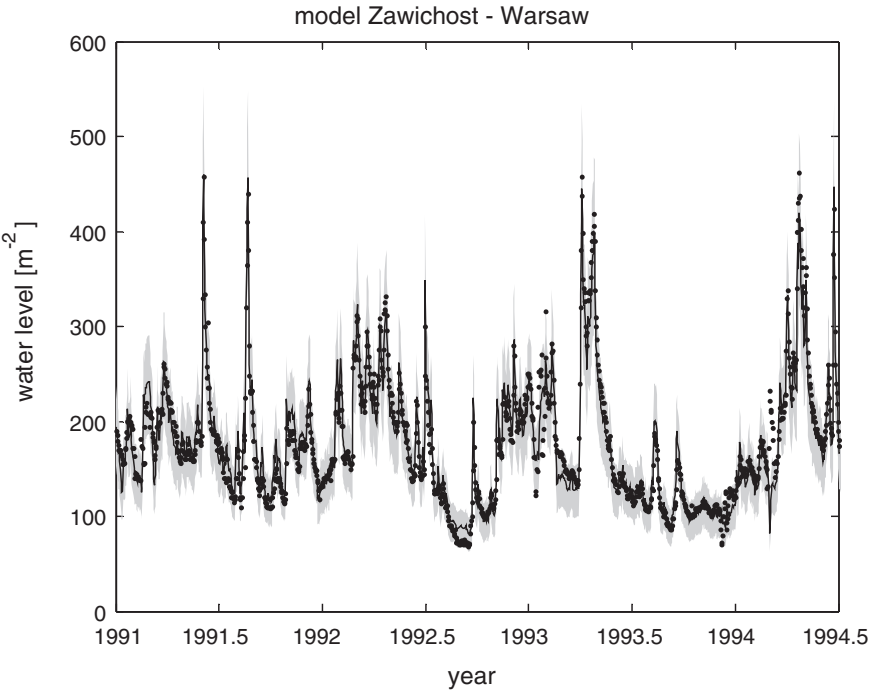

Fig. 3 Model SISO Zawichost-Warsaw Port Praski: validation stage: observations are marked by *dots*, *continuous black line* denotes the STF predictions, *grey areas* denote the 0.95 confidence limits

The model goodness of fit in the form of R_T^2 criterion is equal to 90 % for the calibration period (1985–1987). During the validation period (1988–1997) the R_T^2 was equal to 88 %. The predictions are shown in Fig. 3 by a black thin line. Apart from the water level predictions also 0.95 confidence limits are presented in the form of the shaded area. The black dots denote the observations.

This model gives three-days ahead water level predictions and can be used as the first approximation of the flood wave routing within the studied reach. Therefore, the forecast lead time is equal to the travel time of the flood wave from Zawichost to Warsaw. That lead time could be prolonged if water level observations upstream of Zawichost, e.g. at Kraków and Szczucin, were used. This would require including two additional SISO STF models, from Kraków to Szczucin and from Szczucin to Zawichost. Inclusion of these models prolongs the lead time of a forecast to five days.

3.3 The Lumped-Parameter Non-linear SISO Model for the Zawichost-Warsaw Port Praski Reach

The STF modelling approach assumes that the input and output variables are nearly Gaussian (Young 2006). In order to improve the STF model performance by justifying that assumption we introduced the Box-Cox transformation (Box and

Cox 1964) of both input and output variables (here, water levels measured at two gauging stations along the river):

$$y(\lambda) = \begin{cases} \frac{(y+\lambda_2)^{\lambda_1}}{\lambda_1}, & if \quad \lambda_1 \neq 0 \\ \log(y+\lambda_2), & if \quad \lambda_1 = 0 \end{cases} \tag{6}$$

where $\lambda = [\lambda_1, \lambda_2]$ are the Box-Cox parameters.

The STF model parameters are estimated simultaneously with the Box-Cox transformation parameters. In that sense, the obtained model has the Hammerstein-Wiener structure, well known in many areas, e.g., nonlinear filtering, audio-visual processing, signal analysis and biologic systems (Wills et al. 2013). Models of that type consist of three blocks, nonlinear static transformation of input, linear dynamics and nonlinear transformation of the output, as shown in Fig. 1 of Chap. 10, this book.

The proposed model has a form:

$$x_k^* = \frac{0.4058 \cdot u_{k-2}^*}{1 - 0.3946 \cdot z^{-1}} \tag{7}$$

where x_k^* denotes Box-Cox transformed water levels at Warsaw gauging station (Port Praski) and u_{k-2}^* denotes Box-Cox transformed water levels at Zawichost.

The Box-Cox parameters have a form $\lambda_u = [0.1376, 0.109]$ and $\lambda_y = [0.0353, 0.1103]$.

A validation of the Hammerstein-Wiener model for the Zawichost-Warsaw reach is presented in Fig. 4. The R_T^2 for the calibration was 93 % and for validation it was 90 %.

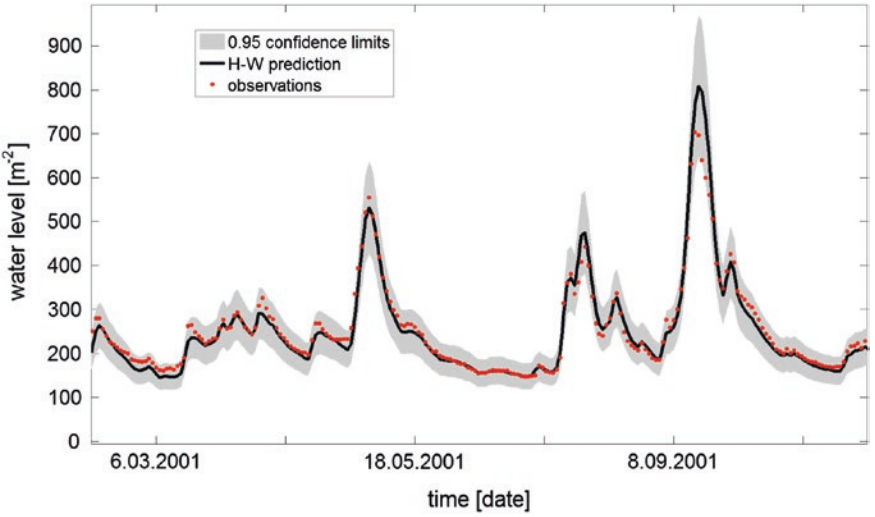

Fig. 4 Validation of the Hammerstein-Wiener model with Box-Cox transformation of water levels on the input and output of the STF model Zawichost-Warsaw [1 1 2]; *continuous line* denotes H-W model predictions; *red dots* denote observations and *shaded area* denotes 0.95 confidence limits

Using water levels instead of flows allows the problem of water balance to be neglected. Moreover, this approach can be used to elicit the form of the unknown rating curve transformation of water levels into flows if the condition for the water balance over a period of time is imposed during the optimization of the model parameters. In the present case study, the introduction of transformation gave a substantial improvement of the water level predictions but the delay time of the model was decreased by 1 day.

This example of an application of the nonlinear transformation of water levels suggests that cross-section nonlinearity can be addressed using a relatively simple approach. We applied Box-Cox transformation in order to improve the STF model predictions, but any other suitable transformation could be used. Moreover, the static transformation parameters can be adjusted to temporal changes in channel geometry. More work is required to explore the benefits that this approach can bring to solving flow prediction problems. Further in this study only an application of linear models is presented.

3.4 The Semi-distributed SISO Models for the River Reach Between Zawichost and Warsaw Port Praski

In the next step, the whole river reach was divided into five sections, following the location of gauging stations, as shown in Fig. 5. It should be noted that the division of the reach into sub-reaches depends on the goal of the modelling. When the prediction with maximum possible lead time is needed, the longest possible reach should be chosen (e.g. Zawichost-Warsaw Port Praski, for predictions of water levels at Warsaw area), as the possible lead-time increases with the length of the reach. This approach allows a quick and reliable forecast of the flow propagation along the reach to be obtained by water managers.

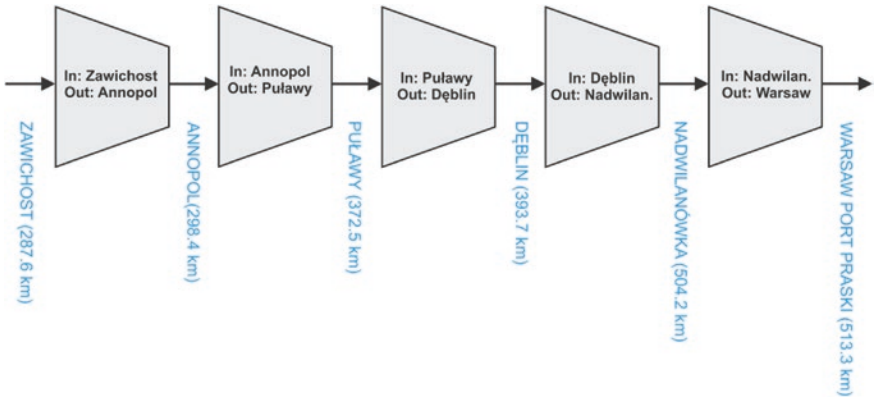

Fig. 5 The scheme of the water level forecasting system Zawichost–Warsaw Port Praski consisting of five single input single output (SISO) STF models

Table 1 SISO STF models describing transformation of water levels along the river reach

Input	Output	Model [n m δ]	T [day]	R_T^2 % calibration	R_T^2 % validation	Std [cm]
Zawichost	Annopol	[1 1 0]	0.21	97.00	97.00	8.4
Annopol	Puławy	[1 1 0]	1.41	91.00	90.00	19.5
Puławy	Dęblin	[1 1 0]	0.97	92.00	94.00	12.0
Dęblin	Nadwilanówka	[1 1 1]	0.65	91.20	92.14	20.7
Nadwilanówka	Warsaw Port Praski	[1 1 0]	1.37	92.40	92.48	9.6
Zawichost	Warsaw Port Praski	[1 2 3]	22.40	90.00	88.00	23.4

The columns from the left to right denote: input and output of the models (gauging stations), model structure, time constants T, R_T^2 criterion for calibration and validation and standard errors of predictions

Each model was calibrated using water level observations from the gauging stations. Further in this subsection the system is illustrated using the linear SISO modules. A summary of the identified model structures is given in Table 1. The model structure is presented following the notation introduced in Sect. 3.1, in the form of a vector [n m δ]. There are also given estimated model time constants (Eq. 3), the results of model calibration for the observations from the period 1985–1987 and a verification of the observations from the period 1988–1997. The last column presents the standard deviation of prediction errors.

Good calibration and validation results were obtained for all analysed reaches. The best fit during calibration was obtained for the Zawichost-Annopol sub-reach, whilst the worst results were obtained for Annopol-Puławy sub-reach. All the sub-models have the first order structure.

Model parameters differ in the sub-reaches. The model time constants vary from 0.21 day for Zawichost-Annopol sub-reach to 1.41 days for Annopol-Puławy sub-reach and 22 days for the whole Zawichost-Warsaw reach.

3.5 MISO STF Models for the River Reach Between Zawichost and Warsaw Port Praski

MISO STF model was applied to take into account major river tributaries within the studied reach. The whole reach was divided into three parts following the location of the main tributaries (Fig. 6): (i) the reach between Zawichost and Puławy with the River Kamienna tributary; (ii) the reach between Puławy and Dęblin with the River Wieprz tributary; (iii) the reach between Dęblin and Warsaw with the River Pilica tributary.

Each of the models includes a MISO STF function with two inputs and one output. The results of model structure identification and parameter estimates are given in Table 2. The model structure is presented following the notation introduced in Sect. 3.1, in the form of a vector [n m_1 m_2 δ_1 δ_2]. Similar as for the SISO model, the period 1985–1988 was used for the calibration and the years 1988–1997 were used for the validation.

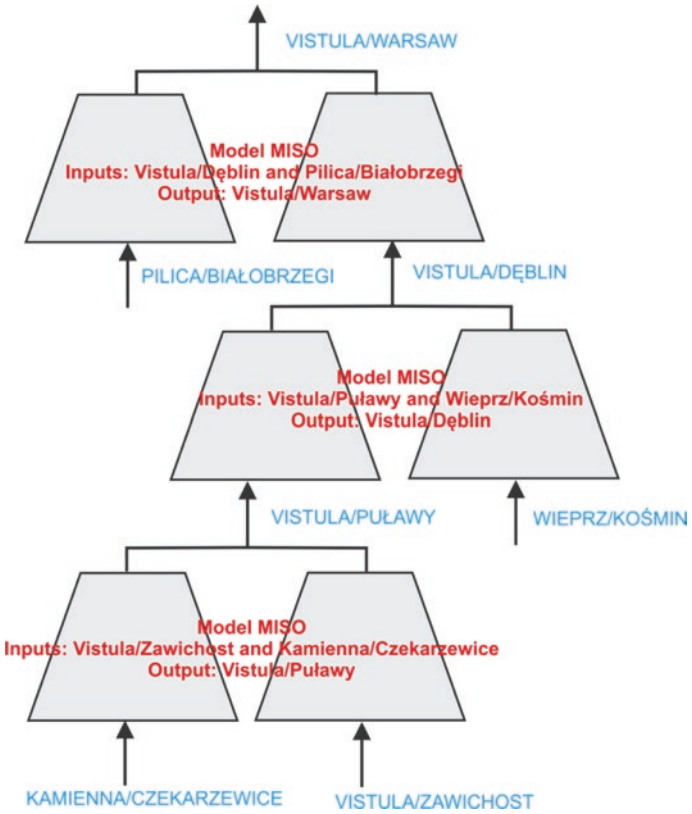

Fig. 6 The scheme of simulation model Vistula/Zawichost–Vistula/Warsaw Port Praski with main tributaries: Rivers Kamienna, Wieprz and Pilica

Table 2 MISO STF models for Zawichost-Warsaw Port Praski reach with tributaries

Inputs	Output	Model	R_T^2 % calibration	R_T^2 % validation	Std (cm)
1. Vistula/Zawichost	Vistula/Puławy	[12211]	92.00	92.60	18.30
2. Kamienna/Czekarzewice					
1. Vistula/Puławy	Vistula/Dęblin	[11111]	87.00	87.00	14.90
2. Wieprz/Kośmin					
1. Vistula/Dęblin	Vistula/Warsaw Port Praski	[11111]	89.00	89.00	28.80
2. Pilica/Białobrzegi					

Columns from left to right denote: location of input water levels; location of output water levels; model structure as given in Sect. 3.1; calibration results; validation results and standard errors of predictions

All identified models have first order dynamics and one day advective delay. All the models give very good fit to the observations. The best results were obtained for the model with Vistula/Zawichost and Kamienna/Czekarzewice water levels as input variables and Vistula/Puławy water levels as an output variable. The determination coefficient R_T^2 equals to 92 % for the calibration and 92.6 % for the validation stages. Relatively worst results were obtained for the second of the analysed MISO models (87 %). The third of the derived MISO models has water levels measured at Białobrzegi on the River Pilica and at Dęblin on the River Vistula as input variables and predicts water levels at Warsaw Port Praski gauging station on the River Vistula. The model structure [1 1 1 1 1] has the following form:

$$x_k = \frac{0.8378 \cdot u_{1,k-1}}{1 - 0.2488 \cdot z^{-1}} + \frac{0.021 \cdot u_{2,k-1}}{1 - 0.2488 \cdot z^{-1}} \tag{8}$$

where $u_{1,k-1}$ denotes water levels at the gauging station in Dęblin in time period $k - 1$, $u_{2,k-1}$ denotes water level at gauging station Białobrzegi (Pilica) in time period $k - 1$ and x_k denotes noise-free model prediction (Eq. 1).

The results of validation of this model are presented in Fig. 7.

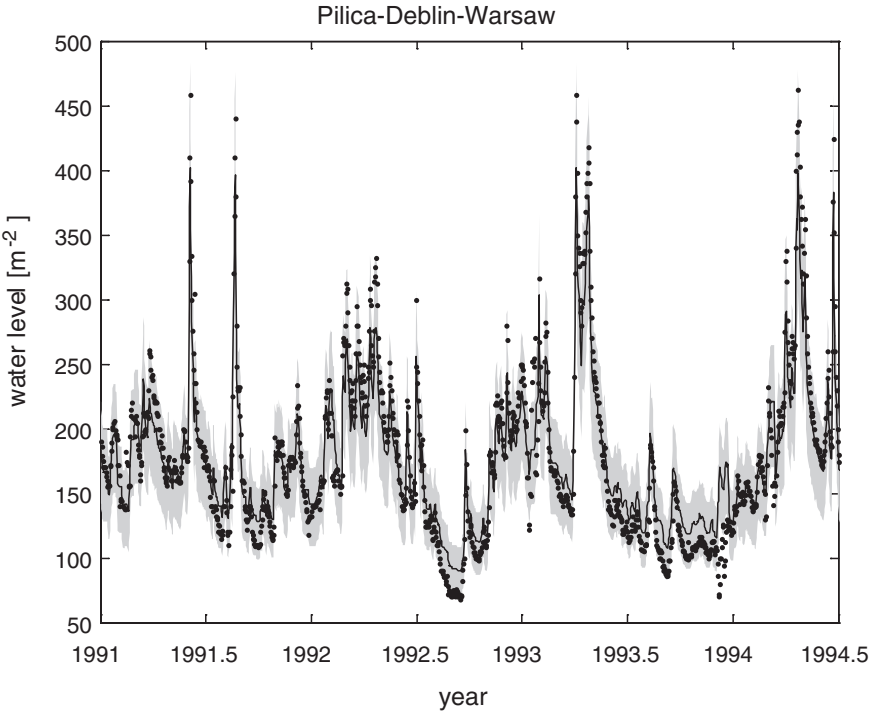

Fig. 7 Validation of MISO STF model with Vistula/Dęblin and Pilica/Białobrzegi water levels as input variables and Vistula/Warsaw Port Praski water levels as predicted variable; model structure is [1 1 1 1 1]: observations are marked by *dots*, *continuous black line* denotes the STF predictions, *grey areas* denote the 0.95 confidence bands

Comparison of SISO and MISO results shows that a more complex model does not necessarily give better model predictions. The determination criterion R_T^2 obtained during calibration of SISO model for Zawichost-Puławy is 91 % whilst MISO model for the same reach gives R_T^2 equal to 92 %. For the next reach, Puławy-Dęblin, SISO gives better results than MISO model. The determination criteria for the calibration and validation stages for SISO are 92 and 94 %, respectively, whilst MISO gives 87 % for calibration and validation stages. However, the flood wave travel times are better captured by MISO than SISO models.

4 An Integrated Simulation System and Assessment of Prediction Uncertainty

The integrated model of the whole river reach can include a combination of SISO and MISO modules, such as presented in Tables 1 and 2.

An example integrated model can have the following form:

1. Model MISO In: Vistula/Zawichost and Kamienna/Czekarzewice Out: Vistula/Puławy
2. Model MISO In: Vistula/Puławy and Wieprz/Kośmin Out: Vistula/Dęblin
3. Model MISO In: Vistula/Dęblin and Pilica/Białobrzegi Out: Vistula/Warsaw Port Praski

The lead time of the predictions for the system from Zawichost to Warsaw equals three days and can be prolonged by the MISO In: Vistula/Szczucin and San/Radomyśl Out:Vistula/Zawichost model and SISO model Vistula/Kraków-Vistula/Szczucin, giving two additional days of a lead, which gives five days of a forecast lead for Warsaw. The predictions of passage of the flood wave along the reach at gauging stations can be obtained using either SISO or MISO models as described by Romanowicz et al. (2006a, b, 2008) for the River Severn, UK, and in Romanowicz and Osuch (2008) for the River Narew.

The estimation of uncertainty of model parameters in the form of the covariance matrix together with model prediction errors are the important advantages of the stochastic transfer function approach. In the case of a stochastic system consisting of serially connected SISO or/and MISO modules, the propagation of errors follows rules of a linear stochastic system, i.e. the variance of the output is propagated into the output in the additive form (Romanowicz and Young 2003; Romanowicz et al. 2006b).

The resulting uncertainty of model predictions, in the form of prediction standard deviations, is given in Tables 1 and 2. The standard deviation errors increase with the length of the modelled reach within the range of 8.4 cm for the Vistula/Zawichost-Vistula/Annopol and 23.4 cm for the Vistula/Zawichost-Vistula/Warsaw Port Praski STF model for the validation stage.

5 Conclusions

This study presents a semi-distributed stochastic system of water level predictions for the middle Vistula River reach from Zawichost to Warsaw. We show that accurate distributed along the river reach predictions of water levels can be obtained using lumped STF models. The results of this work show that predictions of water levels at Warsaw Port Praski gauging stations can be obtained with three-to-five days lead times with an accuracy up to 20 cm std. Apart from the water level estimates, the prediction 0.95 confidence limits are also estimated. We demonstrate that the linear STF approach to water level modelling can be extended to the nonlinear Hammerstein-Wiener model, which enables nonlinear channel geometry to be taken into account. The water level forecasts produced by the STF model with a Box-Cox transformation of both input and output variables gave better fits to the observations, but a one day shorter delay time for the Zawichost-Warsaw Port Praski reach than for the linear STF model. The approach is promising, particularly in connection with its possible application within the on-line updating system (Romanowicz and Osuch 2015, this book) and the possibility of taking into account time-varying river channel geometry. The proposed forecasting system is simple and effective in comparison with much more data and computer time demanding hydrodynamic distributed models, such as MIKE11. The applied stochastic transfer function models are based on observations and therefore they require a good and reliable hydrological monitoring system, in particular in the regions of a high flood risk. On the other hand, distributed models allow the simulation along the entire river reach to be performed, i.e. give spatial interpolation of water levels, whilst the STF models give predictions only at the observation sites.

Acknowledgments This work was supported by the project "Stochastic flood forecasting system (The River Vistula reach from Zawichost to Warsaw)" carried by the Institute of Geophysics, Polish Academy of Sciences, on the order of the National Science Centre (contract No. 2011/01/B/ST10/06866). The water level data were provided by the Institute of Meteorology and Water Management (IMGW), Poland.

References

Abbott M (1979) Computational hydraulics: elements of the theory of free surface flows. Pitman, London

Box GEP, Cox DR (1964) An analysis of transformations. J Roy Stat Soc B 26:211–252

Chow V (1959) Open-channel hydraulics. McGraw-Hill Book Company, New York

Cunge J, Holly F, Verwey A (1980) Practical aspects of computational river hydraulics. Pitman Advanced Publishing Program

Huizinga HJ, Barneveld HJ, Vermeulen CJM, Solheim I, Solbø S (2005) On-line flood mapping using spaceborne SAR-images, water level deduction techniques and GIS-based flooding models in the river Rhine. In: Proceedings of ISFD-conference, Nijmegen, The Netherlands, May 2005

Kadłubowski A (2005) Szkoła Tematyczna MANHAZ, Zarządzanie Zagrożeniami dla Zdrowia i Środowiska, 26–30 Sept 2005

Madsen H, Rosbjerg D, Damgard J, Hansen FS (2003) Data assimilation in the MIKE 11 flood forecasting system using Kalman filtering, in water resources systems-hydrological risk, management and development. In: Proceedings of symposium HS02b held during IUGG2003 at Sapporo, July 2003, vol 281. IAHS Publications, pp 75–81

Mahmood K, Yevjevich V, Miller W (1975) Unsteady flow in open channels. Water Resources Publications

Osuch M (2015) Sensitivity and uncertainty analysis of precipitation-runoff models for the middle vistula basin (this book)

Park J, Obeysekera J, Van Zee R (2005) Prediction boundaries and forecasting of nonlinear hydrologic stage data. J Hydrol 312:79–94

Porporato A, Ridolfi L (2001) Multivariate nonlinear prediction of river flows. J Hydrol 248:109–122

Romanowicz R, Osuch M (2008) An integrated data based mechanistic lowland catchment model for the upper narew, vol E-9(405). Publications of the Institute of Geophysics Polish Academy of Sciences, pp 57–74

Romanowicz RJ, Osuch M (2015) Stochastic transfer function based emulator for the on-line flow forecasting. This book, chapter 10

Romanowicz RJ, Young PC (2003) Data assimilation and uncertainty analysis of environmental assessment problems—an application of transfer function and generalised likelihood uncertainty estimation techniques. Reliab Eng Syst Saf 79:161–174

Romanowicz RJ, Young PC, Beven KJ (2006a) Data assimilation and adaptive forecasting of water levels in the river Severn catchment, United Kingdom. Water Resour Res 42:W06407. doi:10.1029/2005WR004373

Romanowicz RJ, Beven KJ, Young PC (2006b) Uncertainty propagation in a sequential model for flood forecasting. In: Proceedings of symposium S7 on predictions in ungauged basins: promise and progress held during the 7th IAHS Scientific Assembly at Foz do Iguacu, Brazil, April 2005, vol 303. IAHS Publications, pp 177–184

Romanowicz RJ, Young PC, Beven KJ, Pappenberger F (2008) A data based mechanistic approach to nonlinear flood routing and adaptive flood level forecasting. Adv Water Res 31:1048–1056

Romanowicz RJ, Kiczko A, Napiórkowski JJ (2010) Stochastic transfer function model applied to combined reservoir management and flow routing. Hydrol Sci J 55:27–40

Romanowicz RJ, Osuch M, Karamuz E (2014) Stochastyczny system symulacyjny transformacji fali wezbraniowej na odcinku Wisły pomiędzy Zawichostem a Warszawą. II Krajowy Kongres Hydrologiczny. Monografie PAN. z.XX. (in Polish):115–126

Szymkiewicz R (2010) Numerical modeling in open channel hydraulics. Springer, Berlin

Taylor CJ, Pedregal DJ, Young PC, Tych W (2007) Environmental time series analysis and forecasting with the Captain toolbox. Environ Model Softw 22(6):797–814

Thielen J, Bartholmes J, Ramos MH, de Roo A (2007) The European flood alert system—Part 1: concept and development. Hydrol Earth Syst Sci 13:125–140

Wills A, Schön TB, Ljung L, Ninness B (2013) Identification of Hammerstein-Wiener models. Automatica 49(1):70–81

Young PC (1984) Recursive estimation and time series analysis. Springer, Berlin

Young PC (2001) Data-based mechanistic modeling and validation of rainfall-flow processes, In: Anderson MG, Bates PD (eds) Model validation: perspectives in hydrological science. Wiley, New York, pp 117–161

Young PC (2002) Advances in real-time flood forecasting. Philos Trans Math Phys EngSci 360:1433–1450

Young PC, (2006) Transfer function models, In: Anderson MG (ed) Encyclopedia of hydrological sciences. Wiley, Chichester

Stochastic Transfer Function Based Emulator for the On-line Flood Forecasting

Renata J. Romanowicz and Marzena Osuch

Abstract The aim of this study is the derivation of real-time updated, on-line flow forecasts at certain strategic locations along the river, over a specified time horizon into the future, based on information on the behaviour of the flood wave upstream and available on-line measurements at a site. Depending on the length of the river reach and the slope of the river bed, a realistic forecast lead time obtained in this manner may range from hours to days. The information upstream can include observations of river levels and/or rainfall measurements. The proposed forecasting system integrates distributed modelling, acting as a spatial and temporal interpolator with lumped parameter Stochastic Transfer Function models. Daily stage data from gauging stations are typically available at sites 10–60 km apart and test only the average routing performance of hydraulic models rather than their ability to produce spatial predictions. Application of a distributed flow routing model makes it possible to interpolate forecasts both in time and space.

Keywords Distributed model emulator · Water level forecasting · Stochastic transfer function · Kalman filter

1 Introduction

Flow predictions along a river reach are required for flood protection, flood risk assessment and also for the planning of water infrastructure and water management. Due to the uncertainties involved in hydro-meteorological observations and mathematical modelling, the resultant predictions are also uncertain. The

R.J. Romanowicz (✉) · M. Osuch
Institute of Geophysics Polish Academy of Sciences,
Ksiecia Janusza 64, 01-452 Warsaw, Poland
e-mail: Romanowicz@igf.edu.pl

R.J. Romanowicz and M. Osuch (eds.), *Stochastic Flood Forecasting System*,
GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-319-18854-6_10

uncertainty increases with an increase of the time horizon of the prediction—e.g. when forecasts of flow are required many days ahead. Apart from the uncertainty, the speed of forecast acquisition might also be of concern, in particular when fast preventive actions should be taken to issue flood warnings to the public, or some water management actions should be undertaken. In these cases, stochastic emulators of flood wave propagation might be very useful.

Following the definition given by Castelletti et al. (2012), an emulator can be understood as a lower-order approximation of the complex, process-based model, designed to be used as its substitute. In their study, the authors present a classification of problems that can benefit from the application of an emulator and give a design framework specifying different methods used for the reduction of complex models and emulation.

Following that definition, papers by Young et al. (1996) and Romanowicz et al. (1996) can be seen as precursors of the emulation approach.

In the work of Young et al. (1996) an outline of an integrated statistical modelling procedure was formulated that included three tools: sensitivity and uncertainty analysis using Monte Carlo techniques, dominant mode analysis and Data Based Mechanistic DBM modelling. The authors applied the approach to model the global carbon cycle in a changing climate and modelling a horticultural glasshouse for the purpose of automatic control design. The results showed that those two complex environmental models can be approximated by much simpler and more easily handled models. The basic philosophy of that work was that the observations available do not support the high degree of model complexity and much simpler models based on data and, at the same time, providing physical interpretation, are suitable to describe the processes studied.

On the other hand, Romanowicz et al. (1996) presented a distributed flow routing model emulator (although it was not called by that name). It had the form of connected nonlinear storage reservoirs similar to the Cunge (1975) approach. The model was calibrated using 2D model results as a virtual reality. In that sense, the 2D model was simplified to a discrete space form. The purpose of the model was an analysis of the uncertainty of flow predictions and the derivation of probability maps of flood inundation areas using a Generalised Likelihood Uncertainty Estimation GLUE approach (Beven and Binley 1992).

An emulator of the HEC RAS model based on the Stochastic Transfer Function STF description of water level dynamics was presented by Young et al. (2009) and was applied to the River Severn reach between Montford and Shrewsbury, UK.

Apart from the approach aimed at reducing the complex model into a simpler form, as described by Castelletti et al. (2012), Kennedy and O'Hagan (2001) introduced emulators of computer codes built in the form of a Gaussian process modelling of model output. These emulators can be used to perform sensitivity and uncertainty analyses of a model which requires too much computer time to be dealt with directly (Kennedy et al. 2006). In contrast to process-based models, the main aim of these studies is focussed on finding statistical relationships between model variables. The physical interpretation of these relationships is left to the end user.

Another application of an emulator to simplify the distributed flow routing model was presented by Romanowicz et al. (2010). In that study, the emulator based on the STF model was built using the water levels simulated by the HEC-RAS model. The emulator was applied to the design of the control scheme for the water storage reservoir discharges, aimed at maintaining the desired water levels downstream of the reservoir. The application of a much more efficient numerical scheme in the form of the STF model allowed the optimisation of the reservoir control rules to be derived in a reasonable time.

In this study we apply this type of emulator within an on-line forecasting routine that utilises a Kalman filter approach (Kalman 1960). It might be of help in cases where the water level observations are performed with varying discretisation time. That situation occurs in the gauging stations along the River Vistula. The observations are routinely done daily, but during the flood alert and flood risk situations, the observations are performed on shorter time periods—every hour, 0.5 h or even every 10 min, depending on the gauging station. Whilst that situation does not constitute a significant problem for the distributed flow routing model, the on-line updating of such system would have to be designed for irregular observation sets. The model is stochastic, enabling derivation of prediction uncertainty in a straightforward manner. The validation of the emulator and the comparison of its performance are done using available observations of water levels at the cross-section at the Warsaw Nadwilanówka gauging station.

2 Methods

The emulator of a distributed flow routing model described in this study has the form of serially or parallel connected modules representing the desired spatial characteristics of the modelled reach. Each module is calibrated using the outputs simulated by a full distributed model. The form of a module depends on the variable chosen to describe system dynamics (Romanowicz et al. 2010).

The general form of a single module is shown in Fig. 1. That form represents the class of block-oriented Hammerstein-Wiener models, well-known in engineering practice (Wills et al. 2013). They are characterised by a linear dynamic block, placed between two static (memoryless) nonlinear blocks. The basic idea behind that representation is that the process nonlinearity can be separated from the linear process dynamics. An application of this type of model to flow routing

Fig. 1 General form of the single ith module of a distributed model emulator of the Hammerstein-Wiener type; H_i denotes input variable; H_{i+1} denotes the model output

is presented by Romanowicz and Osuch, (Chap. 9, this book,), for water level predictions. The model structure may be simplified by applying only one nonlinearity, at the input or the output. In the first case we have the so-called Hammerstein model; in the latter, we have the so-called Wiener model.

The structure of the emulator presented here depends on the choice of the model variable.

In our study, the emulator is based on the distributed model simulations, available at each model cross-section and discrete in time. For clarity of explanation, let us assume that the emulator's single module is built between the 1st and nth cross-section. In the case where the emulator is built on the flow acting as a model variable, with the water levels acting as an output variable, the system's basic module will have the structure shown in Fig. 2.

The module consists of a linear dynamic part and nonlinear "inverse of a rating curve" static transformation of flow into water level. In this particular case, and further in this section, the system dynamics is modelled using the linear Stochastic Transfer Function (STF) model (Young 2006). However, other linear dynamic models can be used, e.g. a linear storage equation.

For water levels as a state variable, the emulator kth module has a form shown in Fig. 3.

The nonlinear static relation $f(\cdot)$ placed in front of a dynamic part represents possible nonlinearities between water levels at the cross-sections related to the channel geometry. This model structure was applied in flow routing, e.g., by Romanowicz et al. (2008).

As mentioned at the beginning of this section, the emulator can have a form of serially connected blocks, or can have a parallel structure. However, in the serial structure the error propagates downstream of the modelled reach and increases with

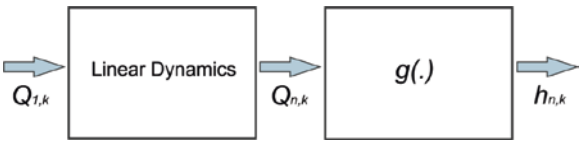

Fig. 2 Module of an emulator based on flows; $Q_{1,k}$ denotes flow at 1st cross-section at time k, for $k = 1,\ldots, T$; $Q_{n,k}$ denotes flow simulated at the nth cross-section at time k; $h_{n,k}$ denotes water levels at the nth cross-section at time k; $g(\cdot)$ denotes the inverse rating curve transformation

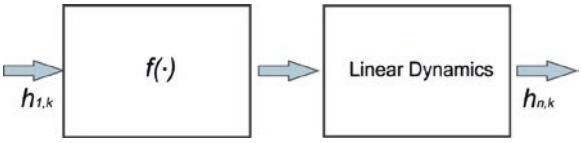

Fig. 3 Module of an emulator based on water levels; $h_{1,k}$ denotes water level at the model input at 1st cross-section at time k, $h_{n,k}$ denotes water levels simulated at the nth model cross-section at time k; $f(\cdot)$ describes process nonlinearity, e.g. related to channel geometry

an increase of the number of modules (Romanowicz and Beven 2006; Beven and Lamb 2014). The parallel structure avoids this problem, but instead the estimation errors increase with the increasing length of the reach due to the growing complexity of the flow transformation process involved (i.e., cross-sections are further apart).

2.1 Case Study

The Middle River Vistula reach stretches between Zawichost and Warsaw and is 100 km long (Fig. 4). A detailed description of the reach is presented by Gutry-Korycka et al. (2015, this book). A distributed flow routing model (MIKE11) was built for the reach based on data for the detailed river channel and floodplain geometry (Kochanek et al. 2015, this book). This model is used for the temporal and spatial interpolation of the water level observations available at only 7 cross-sections (Zawichost, Puławy, Dęblin, Gusin, Warsaw Nadwilanówka and Warsaw Port Praski) and in the form of daily instantaneous measurements of water levels. The observations span over 50 years, but they are irregular, with long periods missing either flow or level data. The flows at Zawichost are used as the model upper boundary condition and a rating curve describing the relationship between

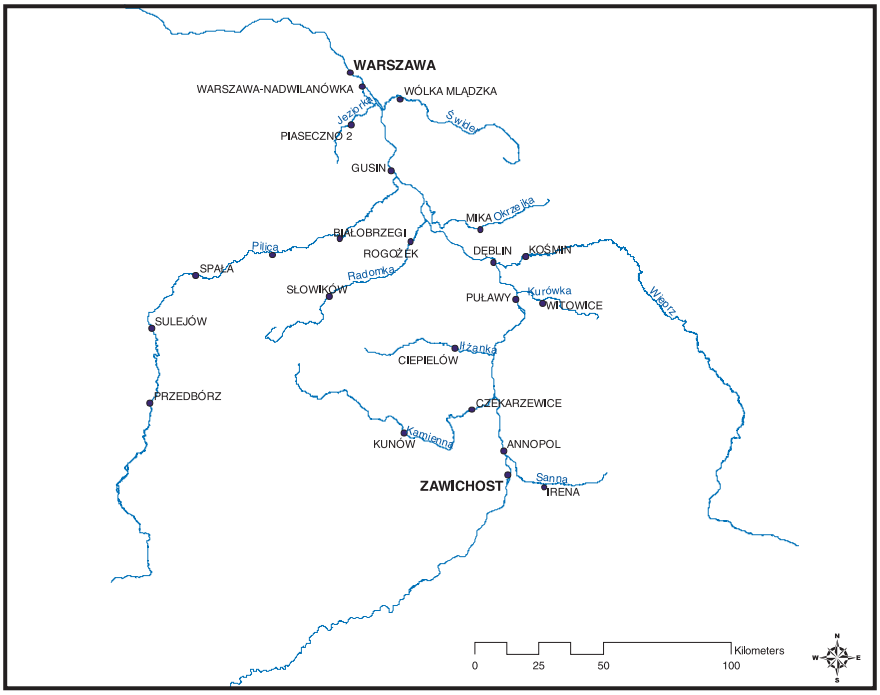

Fig. 4 The study area: middle reach of the River Vistula

flows and water levels is used as a downstream boundary condition at the Warsaw gauging station (near Port Praski) (Kochanek et al. 2015, this book). The simulated water level data were used to build stochastic emulators in the form of a linear and/or nonlinear transformation of water levels at cross-sections along the river reach. An emulator based on flows was also developed. The models are stochastic, enabling the derivation of prediction uncertainty in a straightforward manner; therefore, they are suitable for scenario analysis of the water management system under uncertain climatic conditions.

2.2 Derivation of the MIKE11 Emulator; Comparison with the STF-Based Model Derived from Observations

The distributed model simulations can be registered at each model cross-section with the minimum discretisation time depending on the model numerical solution (0.5 h in the case presented). The emulator can be based on any available input-output relationship within the simulated reach. However, in order to test the ability of an emulator to reproduce the MIKE11 simulations, we have chosen the reach between the Zawichost and Warsaw Nadwilanówka gauging stations, as an illustration. Nadwilanówka is situated at the upper end of the Warsaw reach. The Warsaw gauging station (the downstream end of the Warsaw reach at Port Praski) was not chosen, as it is strongly influenced by the downstream boundary condition (Napiórkowski and Dooge 1988). A 0.5 h discretisation time was chosen as the time step of the emulator. In the first instance, a linear STF model was applied. As the nonlinearity of the flow transformation process increases with the length of the reach, the selected example, being the longest, gives the worst-case scenario. Both water levels and flows were tested as model variables. We applied interpolated flow/water level observations at Zawichost as input variables and simulated flow/water level values at the Warsaw Nadwilanówka gauging station as an output of the STF-based emulator. The identified models have first order dynamics and 65 h advective delay both for water levels and flows acting as model variables. The emulator performance for water levels is worse than for flows, indicating that MIKE11 has linear flow dynamics. The nonlinearity of water level transformation can be attributed to the nonlinear cross-section geometry. Figure 5 presents the validation results of the Zawichost-Nadwilanówka MIKE11 emulator based on water levels. The emulator predictions are given together with 0.95 confidence limits obtained as a result of application of the STF (Romanowicz and Osuch 2015, this book).

The results of validation of the STF model emulator of MIKE11 flow predictions are presented in Fig. 6. As mentioned, the distributed model flow simulations are very well reproduced by the STF model. The accuracy of the validation expressed by the coefficient of determination R_T^2 coefficient is equal to 0.99. The flow observations shown as a red line illustrate that the MIKE11 model does not reproduce the flow timing correctly.

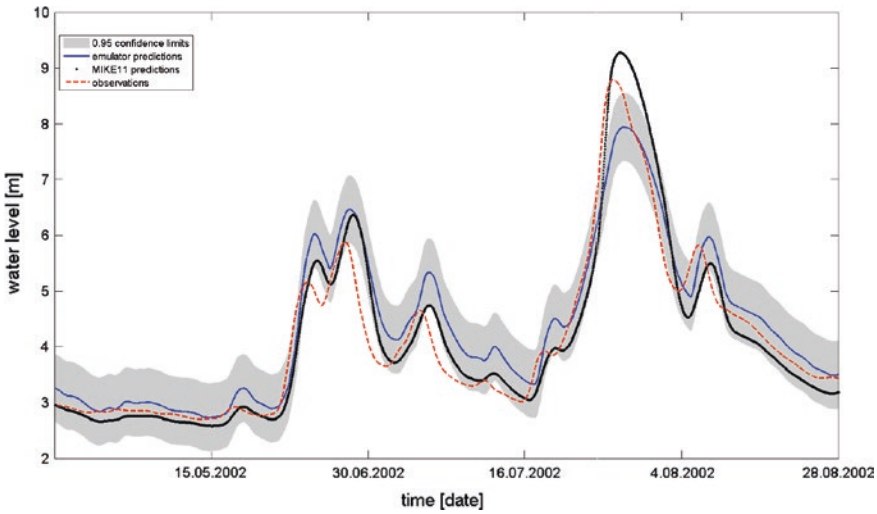

Fig. 5 The STF-based MIKE11 emulator of water level predictions, validation stage; model Zawichost-Nadwilanówka; *blue line* denotes emulator predictions; *black line* denotes MIKE11 simulations, *red dashed line* denotes observations and *shaded area* denote emulator 0.95 confidence limits

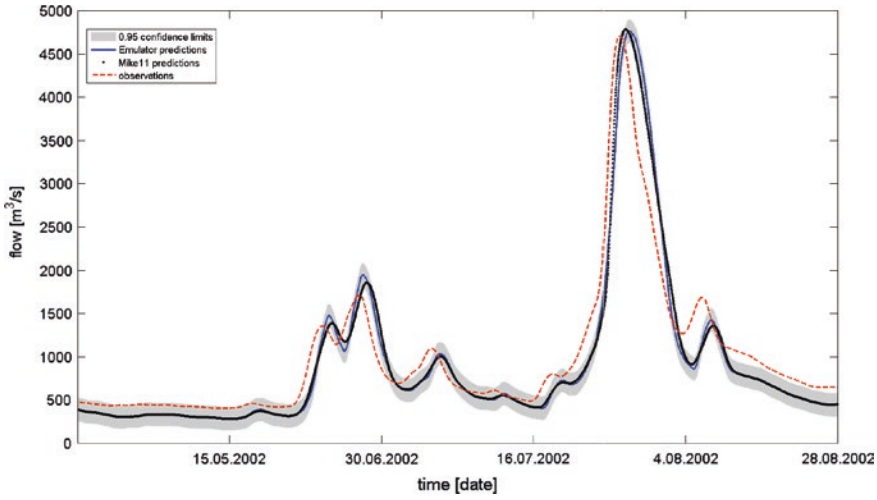

Fig. 6 The STF MIKE11 emulator of flow predictions, validation stage; model Zawichost-Nadwilanówka; *blue line* denotes emulator predictions; *dotted black thick line* denotes MIKE11 simulations; *red line* denotes observed flow rates and *shaded area* denotes emulator 0.95 confidence limits

2.3 Application of a Stochastic Emulator for On-line Data-Assimilation

Data assimilation can be understood as a process of using the available observations to decrease the model prediction errors, either off-line, during the model calibration or on-line when new observations become available.

Following Weerts and El Serafy (2006), data assimilation procedures allow an integrated assessment of input, model structure and observation uncertainty to be performed. The on-line updating methods can be divided into four groups depending on the variable that is updated (Refsgaard 1997). These variables include inputs, state variables, parameters and predictions. There are methods where more than one variable undergoes updating, as for example in Romanowicz et al. (2006, 2008) where state variables and parameters were updated. Studies on the application of coupled and dual updating mechanisms in hydrology were published, inter alia by Weerts and El Serafy (2006) and Moradkhani et al. (2005).

The aim of the integrated stochastic flow model is the prediction of water levels with the lead-time corresponding to the time of a flood wave passage along the river reach. Prediction of water levels is most needed in those reaches which are urbanised and where the losses resulting from the flood would be highest. The assimilation of the observations of water levels at a gauging station nearest to the point of interest allows a decrease of prediction errors to be achieved. It can be performed using an adaptation of the Kalman filter gain (Kalman 1960; Young 2002; Romanowicz et al. 2006).

Following the discussion given by Romanowicz et al. (2006), the procedure of building the model with on-line assimilation of the observations is as follows:

I. identification of the structure of the STF model and estimation of its parameters;
II. formulation of the STF model in a state space form for the application of the Kalman filter allowing the assimilation of observations on-line;
III. estimation of the noise to variance ratio (the so-called NVR); estimation of hyper-parameters of the data assimilation model (its gain parameter and the model variance for the N-step ahead prediction)

An estimate of the N-step ahead prediction (here water levels) has the form:

$$y_{k+N} = \hat{p}_k \cdot \hat{x}_{k+N} + \eta_{k+N} \tag{1}$$

where $\hat{x}_{k+N}$ is the N-steps ahead estimate of water levels obtained in step (II) and η_{k+N} represents the error of that estimate; $\hat{p}_k$ denotes an estimate of the gain p_k assumed to be described by the random walk model with the variance q_k, which is one of the assimilation model hyperparameters. Under this assumption, the gain p_k can be estimated on-line using a recursive minimum square estimation algorithm (Young 2002), based on the available observations up to time k.

The application of the above-listed procedure to the STF system of the Middle River Vistula reach described in Romanowicz and Osuch (2015, this book)

provides daily water level/flow forecasts due to the fact that only daily routine observations of water levels are available at gauging stations along the river. More frequent observations are taken only in case when the danger of flooding arises.

In order to use these more frequent, but sporadic, observations we apply the STF-based emulator of the distributed model described in the previous sub-sections, with a 0.5 h discretisation time. The model utilises interpolated to 0.5 h intervals water level observations at Zawichost as an input. Verification of the model performance was done on the interpolated to 0.5 h observations of water levels at the Warsaw-Nadwilanówka gauging station. The results of validation for the 5 h ahead predictions at Warsaw-Nadwilanówka gauging station are presented in Fig. 7. The figure shows the 5 h ahead forecasts together with 0.95 confidence limits obtained from the applied Kalman filter and updating routine summarised above. There are also shown the observations (red line) and the non-updated emulator forecast is given by the blue line.

Validation of the on-line updating of water levels for lead times of 12 h is presented in Fig. 8. The forecast deteriorates with the increase of the length of the forecast lead-time. However, there is about 1 m error between the non-updated water level prediction and observations, whilst an on-line updated 12 h ahead forecast over-predicts the peak value by 20 cm. As expected, further increase of the forecast lead-time increases the forecast error. Confidence limits provided by the method for the model predictions can play an important role in the decision-making process.

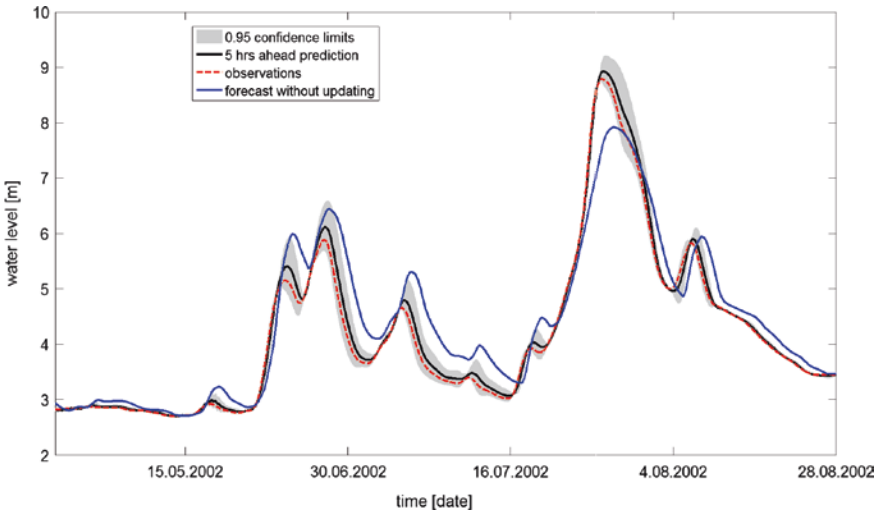

Fig. 7 Emulator-based 5 h ahead forecast of water levels for Nadwilanówka gauging station; observations are shown with the *red line*; *black line* denotes 5 h ahead forecast; *blue line* denotes forecast without updating and *shaded area* denotes the 0.95 confidence limits

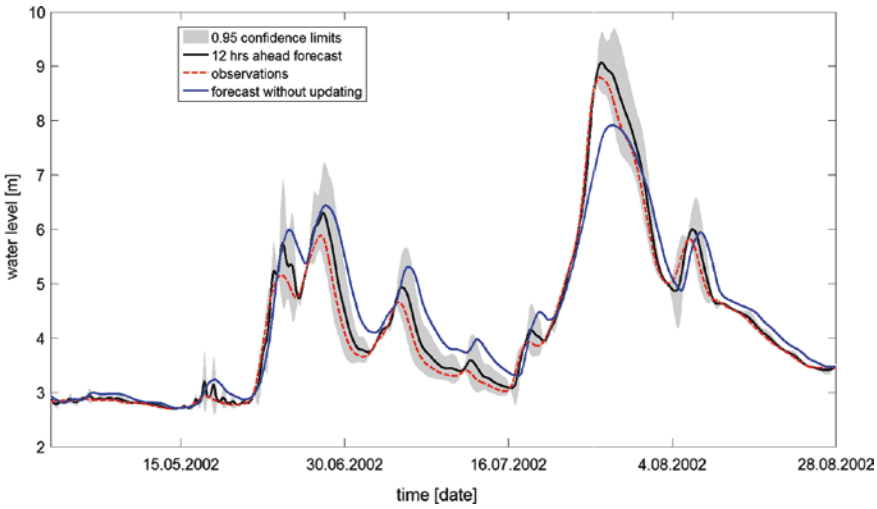

Fig. 8 Emulator-based 12 h ahead forecast of water levels for Nadwilanówka gauging station; observations are shown with the *red line*; *black line* denotes 5 h ahead forecast; *blue line* denotes forecast without updating and *shaded area* denotes the 0.95 confidence limits

3 Conclusions

The aim of this study was to present a summary of the work on the lumped-parameter emulator of a distributed flow routing model and its application to the on-line forecasting system for the Middle River Vistula. In this study we followed the on-line updating methodology described by Romanowicz et al. (2006). The main difference between the former and the present approaches lies in the nature of the case study. Namely, only daily historical observations are available for the Vistula case study, whilst hourly data were available for the River Severn case study used in Romanowicz et al. (2006).

During a flood event, observations are taken more frequently and more time-refined modelling tools are required for the forecast. The distributed model (e.g. MIKE11) provides about 3-days ahead forecasts for the Vistula reach between Zawichost and Nadwilanówka, but its prediction error is too large from a practical point of view. During the flood event, water management officers require on-line forecasts with a much smaller time step. The STF-based MIKE11 model emulator presented in this study enables an application of a Kalman-filter-based updating procedure that gives robust forecasts from 1 up to 15 h ahead lead-time. The other way forward would be to update the MIKE11 predictions on-line. However, that would require much more computer time compared to a lumped parameter emulator. The other advantage of an emulator-based approach lies in the possibility of estimation of the uncertainty of the forecast without the necessity of introducing a Monte-Carlo-based approach required in the distributed model case.

The results of our study show that a simple lumped-parameter flow routing model can perform better than a distributed flow routing model at gauged cross-sections. At the non-gauged cross-sections, the distributed model emulator provides fast and reliable responses. The study has shown that the MIKE11 model is nearly linear in flow predictions and the emulator based on a linear STF model gives sufficiently good results. Despite the nonlinear channel geometry, the water level predictions were reasonably good without nonlinear transformation of water level observations. Further work is underway towards improving the forecasts by the application of the H-W model in flood forecasting.

Acknowledgments This work was supported by the project "Stochastic flood forecasting system (The River Vistula reach from Zawichost to Warsaw)" carried out by the Institute of Geophysics, Polish Academy of Sciences, by order of the National Science Centre (contract No. 2011/01/B/ST10/06866). The rainfall and flow data were provided by the Institute of Meteorology and Water Management (IMGW), Poland.

References

Beven KJ, Binley A (1992) The future of distributed models: model calibration and uncertainty prediction. Hydrol Process 6:279–298

Beven KJ, Lamb R (2014) The uncertainty cascade in model fusion. Geological Society, London, Special Publications. doi:10.1144/SP408.3

Castelletti A, Galelli S, Ratto M, Soncini-Sessa R, Young PC (2012) A general framework for dynamic emulation modelling in environmental problems. Env Model Softw 34:5–18

Cunge JA (1975) Two-dimensional modelling of floodplains. In: Unsteady flow in open channels. 17. Water Resource Publication, Highland Ranch, Colo

Gutry-Korycka M, Mirończuk A, Hoscilo A (2015) Land cover change in the Middle Vistula Basin (this book)

Kalman R (1960) New approach to linear filtering and prediction problems. ASME Trans J Basic Eng 82(D):35–45

Kennedy MC, O'Hagan A (2001) Bayesian calibration of computer codes. JR Stat Soc Ser B Stat Methodol 63(3):425–464

Kennedy MC, Anderson CW, Conti S, O'Hagan A (2006) Case studies in Gaussian process modelling of computer codes. Reliab Eng Syst Saf 91:1301–1309

Kochanek K, Karamuz E, Osuch M (2015) Distributed modelling of flow in the middle reach of the River Vistula (this book)

Moradkhani H, Sorooshian S, Gupta HV, Houser PR (2005) Dual state-parameter estimation of hydrological models using ensemble Kalman filter. Adv Water Resour 28:135–147

Napiórkowski JJ, Dooge JCI (1988) Analytical solution of channel flow model with downstream control. Hydrol Sci J 33(6):269–287

Refsgaard JC (1997) Validation and intercomparison of different updating procedures for real-time forecasting. Nord Hydrol 28(2):65–84

Romanowicz RJ, Beven KJ (2006) Comments on generalised likelihood uncertainty estimation. Reliab Eng Syst Safe doi:10.1016/j.ress.2005.11.030

Romanowicz RJ, Osuch M (2015) Semi-distributed flood forecasting system for the River Vistula reach (this book)

Romanowicz RJ, Beven KJ, Tawn J (1996) Bayesian calibration of flood inundation models. In: Anderson MG, Walling DE (eds) Floodplain processes. Wiley, Chester, pp 333–360

Romanowicz RJ, Young PC, Beven KJ (2006) Data assimilation and adaptive forecasting of water levels in the river Severn catchment, United Kingdom. Water Resour Res 42:W06407. doi:10.1029/2005WR004373

Romanowicz RJ, Young PC, Beven KJ, Pappenberger F (2008) A data based mechanistic approach to nonlinear flood routing and adaptive flood level forecasting. Advances in Water Res 31:1048–1056

Romanowicz RJ, Kiczko A, Napiórkowski JJ (2010) Stochastic transfer function model applied to combined reservoir management and flow routing. Hydrol Sci J—Journal des Sciences Hydrologiques 55(1):27–40

Weerts A, El Serafy G (2006) Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models. Water Resour Res 42:W09403. doi:10.1029/2005WR004093

Wills A, Schon TB, Ljung L, Ninness B (2013) Identification of Hammerstein-Wiener models. Automatica 49(1):70–81

Young PC (2002) Advances in real-time flood forecasting. Philos Trans Math Phys Eng Sci 360:1433–1450

Young PC (2006) Transfer function models. In: Anderson MG (ed) Encyclopedia of hydrological sciences. Wiley, Chichester

Young PC, Parkinson SD, Lees M (1996) Simplicity out of complexity in environ-mental systems: Occam's Razor revisited. J Appl Stat 23:165–210

Young PC, Leedal D, Beven KJ (2009) Reduced order emulation of distributed hydraulic models. In: Proceedings 15th IFAC symposium on system identification SYSID09. St Malo, France

Part IV
Ensemble Forecasts, Linking Conceptual and Data-Based Models for Flow Forecasting

Adaptation of the Integrated Catchment System to On-line Assimilation of ECMWF Forecasts

Adam Kiczko, Renata J. Romanowicz, Marzena Osuch and Florian Pappenberger

Abstract Floods and low flows in rivers are seasonal phenomena that may cause several problems to society. Flow forecasts are crucial to anticipate high and low flow events. The forecasted flow is commonly given as one value, even though it is uncertain. There is an increasing interest to account for uncertainty in flood early warning and decision support systems. In response to that demand, ensemble flood forecasting has been developed using ensembles of numerical weather predictions (NWP) as a driving force for rainfall-runoff models. However, NWPs require bias correction in order to correspond to observations. This study focuses on comparison of two hydrological models and two error reduction techniques of the European Centre for Medium-Range Weather Forecasts (ECMWF). Namely, we compare an application of the conceptual HBV and a data-based mechanistic (DBM) grey-box rainfall-runoff model and two statistical methods of error correction, based on Quantile Mapping (QM), with and without seasonal adjustment. The Biała Tarnowska catchment (southern Poland) is used as a case study. The study shows that a simple, DBM model has similar prediction capabilities as the

A. Kiczko (✉)
Warsaw University of Life Sciences—SGGW, Warsaw, Poland
e-mail: adam_kiczko@sggw.pl

R.J. Romanowicz · M. Osuch
Institute of Geophysics Polish Academy of Sciences,
Ksiecia Janusza 64, 01-452 Warsaw, Poland
e-mail: Romanowicz@igf.edu.pl

M. Osuch
e-mail: marz@igf.edu.pl

F. Pappenberger
European Centre for Medium Range Weather Forecasts, Reading, UK
e-mail: Florian.Pappenberger@ecmwf.int

R.J. Romanowicz and M. Osuch (eds.), *Stochastic Flood Forecasting System*,
GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-319-18854-6_11

more complex conceptual HBV model. The use of QM downscaling techniques improves significantly the prediction skills, but seasonality can be neglected.

Keywords ECMWF forecast · Biała Tarnowska · Flood forecasting · HBV · DBM

1 Introduction

Advancing meteorological forecast methods trigger the development of hydrological warning systems. The most recent weather prediction systems provide information on the uncertainty of a forecast in the form of ensemble forecasts such as the European Centre for Medium Range Weather Forecasts (ECMWF see Buizza et al. 2007). Historic forecasts can be accessed through the THORPEX Interactive Grand Global Ensemble (TIGGE) project (Bougeault et al. 2010).

The ensemble forecast is obtained by introducing random perturbations to the model variables (i.e. initial conditions). The resulting distribution of the output is attributed to the model uncertainty (Zhu 2005). The second property is crucial for flood warning systems. Various studies have been devoted to a problem how to utilize such probabilistic nature of the input taking also into account the uncertainty of hydrological models. The most common approach is based on the assumption that the uncertainty of a weather prediction overwhelms those of rainfall-runoff modelling, which can be therefore neglected (Pappenberger et al. 2011). The example applications are given by e.g. Gouweleeuw et al. (2005).

However, prediction skills of flow forecasts might depend equally strongly on the uncertainty of runoff models (Demirel et al. 2013). Many efforts have been made to include both sources of uncertainty, weather forecasts and hydrological models as well as initial conditions. The first combined approach, named cascading model uncertainty, was given by Pappenberger et al. (2005) and further developed by Cloke and Pappenberger (2009), He et al. (2009) and others.

The use of Numerical Weather Prediction (NWP) output in hydrological modelling requires application of downscaling techniques to link different spatial scales. The weather forecasts are issued for large computational cells of NWPs, while hydrological models are mostly conditioned on data collected with a much finer resolution or direct point measurements. A number of downscaling tools might be used to apply a correction to the weather forecasts and improve ensemble dispersion. These include statistical correction methods which can be applied either to meteorological ensembles or to hydrological forecasts, or to both input and output variables of a hydrological model (Zalachori et al. 2012).

The development of the stochastic NWP, that utilizes the forecast uncertainty, should allow for improving the prediction skills of the hydrological warning systems. This is a conclusion from various previous studies; however, it is still necessary to investigate the applicability of this approach in different case studies

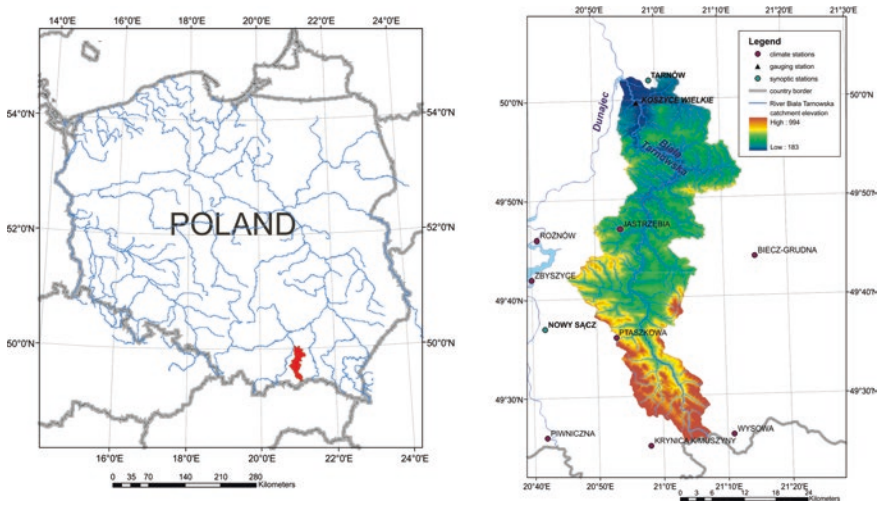

Fig. 1 Study area: Biała Tarnowska catchment

(Cloke and Pappenberger 2009). The objective of this work is to test two different modelling approaches with the European Centre for Medium-Range Weather Forecasts (ECMWF) for the purpose of hydrological ensemble forecasting applied to the Biała Tarnowska catchment (southern Poland, Fig. 1). It is the first application of the ECMWF forecasts in Poland. The first modelling technique involves the statistical downscaling of the forecast ensemble and a conceptual rainfall-runoff model HBV (Bergström 1995). It is a common design of the flood warning model. In the second, approach, the grey-box model following the Data Based Mechanistic (DBM) methodology of Young (2001) is applied instead of the HBV model. It is the first time that the DBM model is applied with the ECMWF forecasts. The DBM model has much simpler structure than the conceptual HBV model that might be an advantage in an on-line forecasting. For both modelling concepts, an analysis of the influence of estimation uncertainty on the forecast performance was performed. The hydrological models are customarily used for flow forecasting for lead times up to several days using meteorological forecasts of the European Centre for Medium-Range Weather Forecasts (ECMWF) as input. In the present application we analyze the efficiency of the warning system with discharge forecasts issued up to seven days ahead, as the longer lead times have not provided satisfactory forecast quality.

The two modelling approaches are compared in a sense of a Relative Operating Characteristic ROC (Peterson et al. 1954). Their performance is tested for different lead-times.

Next section presents the study area, followed by the methodology, numerical analysis, the assessment of forecasts and a discussion of the results.

2 Study Area

Biała Tarnowska catchment is situated in southern Poland, in the Carpathian Mountains. The highest peaks within the catchment reach almost 1000 m a.s.l. It is a typical mountainous river with slopes up to 8.6 ‰ in the upper part and about 0.9 ‰ in the lower part. The catchment area to the Koszyce Wielkie gauging station, where stage observations are taken from, equals 956.9 km^2. Biała Tarnowska catchment is characterized by high precipitations and a large variation of runoff during summer.

3 Methodology

The forecasts used in this study consist of perturbed ECMWF forecasts available from the THORPEX Interactive Grand Global Ensemble (TIGGE, Bougeault et al. 2010). The current ensemble model has a resolution of ~16 km. However, it is still coarse with respect to requirements of hydrological runoff models, conditioned mostly on weighted point measurements of meteorological variables. These types of data differ from the NWP forecast, averaged in space, by a mean and also variation (Maraun et al. 2010). This has to be overcome by an application of downscaling methods, which allows a link between the larger spatial scale of the NWP and a local scale of the hydrological model to be established. In hydrology, the most widely used are techniques of the Statistical Downscaling (SD), where the downscaling consists of identifying of a transformation that allows to unify the properties of the forecast ensemble with the observation data. The Dynamic Downscaling (Boé et al. 2007) is another downscaling approach where circulation models with a finer resolution are used. In the present study, two methods of the SD are applied, the Quantile Mapping (QM) with and without seasonal adjustment. In order to take into account uncertainty of hydrological model parameters and their structure, we test two different rainfall-runoff models. As the discharge from the basin is also influenced by the snow accumulation, the first is a conceptual HBV model (Bergström 1995), where this term is well developed. The second is a data-based (grey box) DBM model, with transparent structure that reduces the computational cost and therefore requires less computer time for the uncertainty estimation, which is advantageous in the real time flow forecasting.

The perturbed ECMWF forecasts have a form of 51 ensembles with an issue time every 12 h and lead-time up to 15 days. Due to the fact that only daily historical observations of rainfall, temperature and discharge/water levels are available, the approach will be illustrated using ECMWF forecast with the lead time of multiples of 24 h periods.

The HBV and DBM solutions depend on initial conditions. They are especially important during the winter period, when water is accumulated in the snow cover. Here, to include the water conditions of the basin in the flood forecasting model,

Table 1 Experimental tests of the flood forecasting system for the Biała Tarnowska catchment

No,	Name	Runoff model	Runoff model uncertainty	NWP downscaling method
1	DBM, Raw	DBM	no	None/raw
2	DBM QM nonseasonal	DBM	no	Quantile Mapping, without distinction of the seasons
3	DBM QM seasonal	DBM	no	Quantile Mapping, conditioned separately for winter and summer season
4	DBM U, Raw	DBM	yes	None/raw
5	DBM U QM nonseasonal	DBM	yes	Quantile Mapping, without distinction of the seasons
6	DBM U QM seasonal	DBM	yes	Quantile Mapping, conditioned separately for winter and summer season
7	HBV, Raw, 1 PS	HBV	no	None/raw
8	HBV, QM seasonal, 1 PS	HBV	no	Quantile Mapping, conditioned separately for winter and summer season
9	HBV, QM nonseasonal, 1 PS	HBV	no	Quantile Mapping, without distinction of the seasons
10	HBV, Raw, 6 PS	HBV	yes	None/raw
11	HBV, QM Seasonal, 6 PS	HBV	yes	Quantile Mapping, conditioned separately for winter and summer season
12	HBV, QM nonseasonal, 6 PS	HBV	yes	Quantile Mapping, without distinction of the seasons

the simulation period was extended backwards by one or three months, depending on the season. In a result, before reaching the forecast time, the model state variables should not depend on initial values any longer.

Twelve different experimental tests of the bias correction and rainfall-runoff modelling techniques applied during the performance tests are presented in Table 1.

3.1 Quantile Mapping

The direct outputs from meteorological forecasting systems are biased and especially precipitation is inadequate for reliable hydrological predictions. To overcome this problem, error correction is applied to both temperature and precipitation ECMWF forecasts. In this paper a Quantile Mapping (QM) method was applied.

The QM method was proposed by Panofsky and Brier (1968) and is widely used as a downscaling technique for ensemble forecasts (e.g. Boé et al.

2007; Wetterhall et al. 2012). The main advantage of the QM, comparing to the simple, delta change error-correction technique, is that it affects not only a mean value but also quantiles of distribution. Wetterhall et al. (2012), who analyzed different statistical downscaling methods, showed that QM provide a similar improvement of the flood forecast system as a Direct Method, where forecast is transformed by a constant factor.

In the QM method, the cumulative distribution functions (cdfs) are derived for previous forecasts and observation sets (Boé et al. 2007). The issued forecast is then transformed in two stages: (i) using the cdf of the simulations, variables are mapped to the quantiles, (ii) which then are converted to the corrected forecast by the inverse cfd for observation sets. Extreme values, especially in the case of a short training set, might fall beyond the ranges of the empirical cdf, bounded by minimum and maximum observation values. To overcome this issue, Boé et al. (2007) proposed an extrapolation using a constant correction, calculated as a difference of last quantiles for simulation and observation cdfs. This approach is applied in the present study.

For the Biała Tarnowska catchment, the cdfs were identified for 2006–2008 period which is called later a training period. We corrected precipitation and air temperature time series independently for each NWP grid located in the catchment area.

The bias of forecasts depends on a circulation pattern (Boé et al. 2007). Due to the short training period, conditioning of the QM on weather patterns would significantly reduce the sample size used to derive the cdf making impossible proper bias correction. In this paper, the QM cdfs were elaborated separately for winter and summer seasons to include seasonal variation of the forecast bias (Maraun et al. 2010).

Figure 3 shows an example hydrograph for the flood event from May 2010. The ensemble forecasts were bias corrected using QM seasonal approach (separately for winter and summer periods).

3.2 HBV Model

The HBV model, developed by Bergström (1995), is a conceptual model of a basin runoff (Wetterhall 2014). It incorporates an efficient snow accumulation and melting module, which is an important issue in the case of the River Biała Tarnowska, where a spring thaw causes a significant flood thread. The model parameters were initially identified using Differential-Evolution algorithm (Storn and Price 1997) and a sum of square errors, as an objective function. The description of the model is presented in Chap. 5, this book.

The uncertainty of model parameters was estimated following the Generalized Likelihood Uncertainty Estimation approach (Beven and Binley 1992). The procedure was applied for time series of the observed precipitation, air temperature and flows for the 2004–2006 period. The meteorological data was available for five

stations: Nowy Sącz, Tarnów, Biecz-Grudna, Krynica and Wysowa. The discharge series were taken from the Koszyce Wielkie gauging station.

The parameters were sampled uniformly with the mean values equal to the derived optimal values of parameters and the variance chosen following the information on parameter spread near the optimum obtained from the optimisation process. The transformation of a priori distribution to a posteriori is achieved using likelihood function, defined as an exponential function of model residuals, in the form proposed by Romanowicz and Beven (2006). However, the required size of the Monte Carlo (MC) sample, that is necessary to estimate the a posteriori distribution of parameters, leads to the solution that is computationally costly. It reduces the applicability of the model in real-time forecasting.

To increase the numerical efficiency, but keeping the information on model uncertainty, we decided to reduce the MC sample size by clustering. In a parameter space, six classes were defined, according to the Euclidean distance of normalized parameters values. The reduced parameter sample was elaborated by taking the most probable sets, in a sense of a likelihood function value based on the optimisation criterion, from each class. The probabilistic interpretation of the parameter ensemble is not straightforward; however, it reflects the model uncertainty resulting from the parameter variation. As a result of clustering, six sets of parameters were derived. Therefore, the dimension of a parameter sample used in simulations was equal to six (6 PS). The applied concept is similar to the approach presented by Pappenberger and Beven (2004).

To analyze how taking into account the uncertainty of the model parameters improves the performance of the flood warning system, a single set of parameters (1 PS) was also considered. In that case, the model has a clearly deterministic form and should reflect the common approach of neglecting parametric uncertainty. The set of parameters was chosen as the one with the highest value of the likelihood function in the Monte Carlo sample, the same that was used for the clustering. It ensures that the probabilistic and deterministic forms of the HBV were elaborated for the same goodness of fit criteria.

3.3 Grey-Box DBM Model

The concept of Data Based Mechanistic approach consists of an application of simplified data-based dynamic models that can still have some physical interpretation (Young 1999, 2001).

The DBM model applied in this work includes three modules. These are: the Stochastic Transfer Function module describing linear flow dynamics, effective rainfall module based on the Soil Moisture Accounting (SMA) approach (Jolley 1995) and a snow-melt module based on the model introduced by Castelletti et al. (2005), presented schematically in Fig. 2.

Optimization of all model parameters was performed simultaneously using standard MATLAB optimization procedures and a sum of square errors, the same

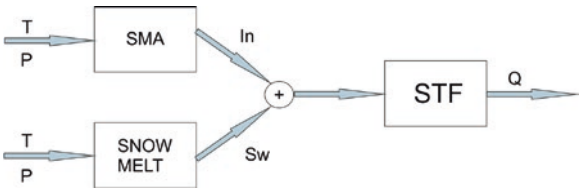

Fig. 2 The schematic presentation of a *grey-box*, DBM rainfall-runoff model; *T* denotes temperature, *P* denotes precipitation, *In* denotes effective rainfall, *Sw* denotes water delivered by snowmelt and *Q* denotes the flow

as that used for the HBV model optimisation. The choice of an optimisation routine has been dictated by the comparative simplicity of the DBM model, but any other method could be used. The DBM model uncertainty was assessed using MC sampling from the model parameter space taking into account parameter covariance structure and the DBM model. The clustering was not necessary because the model is much more computationally efficient than the HBV.

3.4 An Assessment of Flow Forecast Performance

The performance of the flow forecasts, evaluated using different bias correction methods and two rainfall-runoff models, was assessed using the score of the Relative Operating Characteristic (ROC, Peterson et al. 1954).

The score is defined as an area under the ROC curve, which expresses the probability that system correctly anticipates the occurrence of an event (hit rate), against a probability of a false indication (false alarm rate). That means, the higher the ROC curve the better, and it should lie above the diagonal of the ROC plot for the forecast to be useful. The solution of a probabilistic forecast system has a form of a probability that certain event will occur and warning is issued if the assumed threshold is reached. The ROC curve represents the dependency of hit-rate and false-alarm rate for each threshold value (Mason and Graham 2002).

In this approach, the process description takes a binary form, i.e., an event, non-event, and a score is a measure how the system discriminates between events and non-events. The technique was initially developed in the field of signal-detection theory and at present is widely used in other areas. The application of ROC in the forecast system was discussed in detail by Mason and Graham (2002) and can be found in recent studies on flood warning systems like Demargne et al. (2010), Wetterhall et al. (2012) and others.

The events were defined using the local flood defence criteria. For the Koszyce Wielkie gauging station, two flood warning levels are specified: (i) flood warning for the discharge above 100 m^3/s (ii) flood alert for the discharge above 174 m^3/s. The bias correction methods and two rainfall-runoff models were tested

for predicting these two warning levels, at different forecast lead-times, from 1 to 7 days ahead. The score was computed for the 2007–2010 (01-Nov-2007–31-Oct-2010) period, during which the most noticeable flood events occurred in 2010.

4 Results and Discussion

Ensemble of hydrographs is the outcome of each of twelve test-cases presented in Table 1, as shown, for example, in Fig. 3, where the observed discharge is compared with "spaghetti" ensemble predictions for the 2010 flood event.

For each approach the ROC was determined, measuring the efficiency in predicting the warning flows ($Q_a = 100$ m^3/s) and alert flows ($Q_t = 174$ m^3/s). The above threshold values are used in a local flood management strategy.

Additionally, the ROC was computed for seven different forecast horizons: starting from first to seventh day, then from second to seventh day and up to the prediction at seventh day only. The forecast is successful, if an extreme flow (both observed and predicted) occurred in the given forecast timespan. If the system predicts warning or alert flows at any time step within a given time horizon and the discharges above threshold levels have been truly observed, the solution is considered as a hit. Taking the above in consideration, the longer the time horizon is used, the less demanding the test.

It should be noted that the hit-solution as defined above could be different, e.g., by a mean, from observed values. This result is justified, as only the capability to

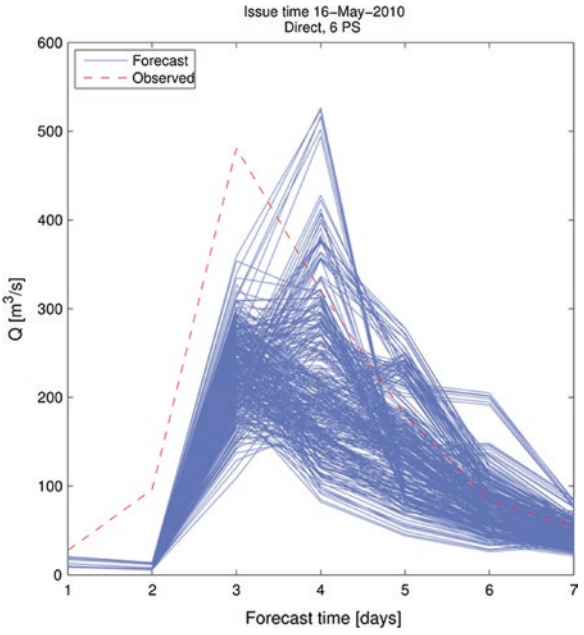

Fig. 3 An example ensemble hydrograph for the flood event from May 2010 (*blue lines*); *red dotted line* denotes observed flow

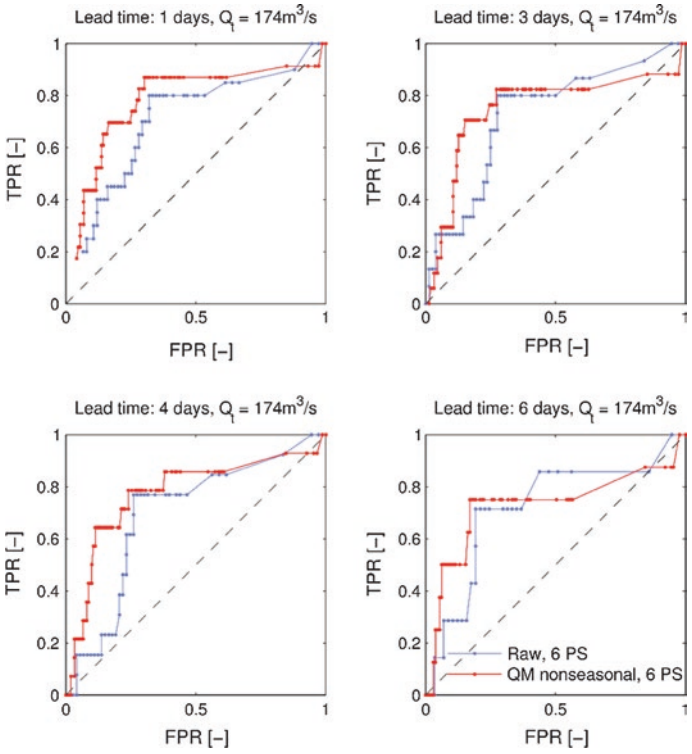

Fig. 4 The ROC plots showing the performance of the flood warning system in identifying flood alert events (discharge above $Q_t = 174$ m^3/s) for the HBV model with uncertainty estimation. The *blue curves* were calculated for the system based on raw forecast ensembles (untransformed) and *red* on transformed using the quantile mapping method, derived separately for summer and winter seasons; *x-axis* denotes false positive rate and *y-axis* denotes true positive rate of the classifier

forecast the threshold values is measured. This might be seen in the example given in Fig. 3. The prediction is considered as successful, as the majority of ensemble members exceeded threshold values and flows above the alert levels are considered as very probable.

In Fig. 4 an example ROC for two models is shown. It provides information on the probability of true and false predictions at certain threshold levels. In the presented plots, the ROC was determined for the HBV model with the uncertainty estimation, supplied by two different data sets: transformed NWPs output through the QM ("HBV, QM Seasonal, 6 PS") and untransformed one (Raw). The performance of the approach based on corrected predictions is better than the second one, as can be seen in higher ratio of hits (TPR) than false-alarms (FPR).

In the next step, by an integration of the ROC, a score measure of overall performance of tested models was computed. The results for the predictions of flood

Table 2 The ROC score measuring the efficiency of selected methods in predicting the flood warning flows ($Q_a = 100\ m^3/s$) for different time horizon (d–days)

	ROC score						
	1-7 [d]	2-7 [d]	3-7 [d]	4-7 [d]	5-7 [d]	6-7 [d]	7-7 [d]
DBM, Raw	0.671	0.659	0.644	0.670	0.661	0.631	0.629
DBM QM nonseasonal	0.772	0.743	0.727	0.728	0.73	0.727	0.706
DBM QM seasonal	0.755	0.727	0.707	0.711	0.688	0.679	0.611
DBM U, Raw	0.804	0.792	0.763	0.767	0.795	0.817	0.753
DBM U QM nonseasonal	**0.849**	**0.831**	**0.806**	**0.816**	**0.847**	0.776	**0.892**
DBM U QM seasonal	0.833	0.819	0.789	0.797	0.765	**0.820**	0.888
HBV, Raw, 1 PS	0.628	0.607	0.548	0.525	0.486	0.496	0.562
HBV, QM seasonal, 1 PS	0.795	0.769	0.743	0.735	0.728	0.760	0.724
HBV, QM nonseasonal, 1 PS	0.814	0.79	0.758	0.746	0.727	0.743	0.744
HBV, Raw, 6 PS	0.811	0.793	0.751	0.755	0.740	0.768	0.755
HBV, QM Seasonal, 6 PS	0.802	0.773	0.740	0.719	0.713	0.707	0.630
HBV, QM nonseasonal, 6 PS	0.787	0.758	0.728	0.711	0.695	0.695	0.603

With bold font and grey background, the maximum values were emphasized

warning flows are given in Table 2 and for flood alerts in Table 3. In the first case, the best forecasts were obtained using DBM model, run on the forecasts transformed with the simple QM. The seasonal QM correction did not improve the performance of predictions, neither for the DBM nor for the HBV rainfall-runoff models. It can be noted that high scores for flood warning predictions were obtained for models with the raw NWP forecasts as an input.

The results are different for flood alert flows (Table 3). The highest scores were reported for the HBV model, including deterministic and probabilistic approaches, and a deterministic (without uncertainty modelling) DBM models, for the input data transformed using the QM method. At this point it is hard to indicate the most appropriative method, as both methods provided similar scores. Moreover, the results are similar in the case of two approaches to the QM. It can be only noted that predictions based on raw data have significantly worse performance.

The advantage of the DBM models with uncertainty modelling, which could be noted for flood warning predictions, is not maintained for the flood alerts. In the HBV approach, the uncertainty term improves the predictions of the flood alerts

Table 3 The ROC score measuring the efficiency of selected methods in predicting the flood alert flows ($Q_t = 174$ m^3/s) for different time horizon (d–days)

	ROC score						
	1-7 [d]	2-7 [d]	3-7 [d]	4-7 [d]	5-7 [d]	6-7 [d]	7-7 [d]
DBM, Raw	0.630	0.628	0.619	0.584	0.639	0.617	0.562
DBM QM nonseasonal	0.807	0.828	0.829	0.782	0.837	**0.819**	0.701
DBM QM seasonal	0.785	0.814	0.814	0.766	0.806	0.763	0.721
DBM U, Raw	0.675	0.688	0.641	0.589	0.554	0.537	0.526
DBM U QM nonseasonal	0.792	0.810	0.815	0.810	0.804	0.784	0.646
DBM U QM seasonal	0.786	0.805	0.791	0.792	0.795	0.802	0.703
HBV, Raw, 1 PS	0.741	0.719	0.633	0.583	0.500	0.361	~0
HBV, QM seasonal, 1 PS	**0.833**	**0.850**	**0.833**	0.809	0.835	0.803	0.735
HBV, QM nonseasonal, 1 PS	0.815	0.836	0.817	0.793	0.814	0.788	0.720
HBV, Raw, 6 PS	0.719	0.730	0.740	0.706	0.759	0.727	0.637
HBV, QM seasonal, 6 PS	0.809	0.814	0.791	**0.814**	**0.836**	0.806	0.765
HBV, QM nonseasonal, 6 PS	0.782	0.779	0.750	0.772	0.775	0.720	**0.772**

With bold font and grey background, the maximum values were emphasized

with longer lead times. However, it is not a case for the remaining tests, where the single realization model has similar scores.

5 Conclusions

The results give an insight into different aspects of methods applied in the flood prediction system for the Biała Tarnowska basin.

1. The study shows that much simpler, grey-box DBM model has similar prediction capabilities as the conceptual HBV model. Furthermore, in the case of flood warnings, the DBM simulations fed with corrected by QM method ECMWF forecast surpass the other tested approaches.
2. The uncertainty term improves significantly the performance of the DBM approach in predictions of flows higher than flood warning. A similar effect, but on a smaller scale, can be noted for the HBV model in flood alert predictions for longer lead times. On the other hand, the results do not show if runoff

uncertainty modelling has a positive impact on basin outflow predictions. The DBM model provides much less accurate predictions of flood alert flows, if the uncertainty term is present. That conclusion points out that the DBM model has a tendency to under-prediction, which can be related to its simplicity. If optimised on maximum flow, its performance would be better for flood alert but worse for the low and medium flows (Romanowicz et al. 2006).

3. The use of downscaling techniques, such as QM, improves the prediction skills significantly. This conclusion does not provide any assessment of the impact of input uncertainty on predictions, as it was not assessed here.
4. In the present study, conditioning the QM on two seasons did not improve the performance of either DBM or HBV rainfall-runoff models.
5. Further research is needed on introducing additional score measures (Ferro 2014) and estimation of uncertainty of their results. Updating of the runoff model would improve the overall forecast performance.
6. The performance of the warning system could be further improved by extending the ensemble forecast using outputs from the other NWP models, provided by different meteorological centres, also available in the TIGGE project.

Acknowledgments This work was financed by the project "Stochastic flood forecasting system (The River Vistula reach from Zawichost to Warsaw)" carried out by the Institute of Geophysics Polish Academy of Sciences for the National Science Centre (contract no. 2011/01/B/ST10/06866). The flow data were provided by the Institute of Meteorology and Water Management (IMGW), Poland.

References

Beven K, Binley A (1992) The future of distributed models: model calibration and uncertainty prediction. Hydrol Process 6(3):279–298

Bergström S (1995) The HBV model. In: Singh VP (ed) Computer models of watershed hydrology. Water Resources Publications, Highlands Ranch, pp 443–476

Boé J, Terray L, Habets F, Martin E (2007) Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. Int J Climatol 27(12):1643–1655

Bougeault P, Toth Z, Bishop C, Brown B, Burridge D, Chen DH, Ebert B, Fuentes M, Hamill TM, Mylne K, Nicolau J, Paccagnella T, Park Y-Y, Parsons D, Raoult B, Schuster D, Silva Dias P, Swinbank R, Takeuchi Y, Tennant W, Wilson L, Worley S (2010) The THORPEX interactive grand global ensemble. Bull Am Meteorol Soc 91:1059–1072

Buizza R, Bidlot J-R, Wedi N, Fuentes M, Hamrud M, Holt G, Vitart F (2007) The new ECMWF variable resolution ensemble prediction system (VAREPS). Q J R Meteorol Soc 133:681–695. doi:10.1002/qj.75

Castelletti A, Pianosi F, Soncini-Sessa R, Young PC (2005) Data-based mechanistic modelling of a snow affected basin. In: Proceedings of the 16th IFAC World Congress. doi:10.3182/20050703-6-CZ-1902.02171,2170

Cloke H, Pappenberger F (2009) Ensemble flood forecasting: a review. J Hydrol 375(3–4):613–626

Demargne J, Brown J, Liu Y, Seo D-J, Wu L, Toth Z, Zhu Y (2010) Diagnostic verification of hydrometeorological and hydrologic ensembles. Atmos Sci Lett 11(2):114–122

Demirel MC, Booij MJ, Hoekstra AY (2013) Effect of different uncertainty sources on the skill of 10 day ensemble low flow forecasts for two hydrological models. Water Resour Res 49:4035–4053

Ferro CAT (2014) Fair scores for ensemble forecast. Q J R Meteorol Soc 140(683):1917–1923
Gouweleeuw BT, Thielen J, Franchello G, De Roo APJ, Buizza R (2005) Flood forecasting using medium-range probabilistic weather prediction. Hydrol Earth Syst Sci 9(4):365–380
He HY, Cloke HL, Wetterhall F, Pappenberger F, Freer J, Wilson M (2009) Tracking the uncertainty in flood alerts driven by grand ensemble weather predictions. Meteorol Appl 16(1):91–101
Jolley TJ (1995) Large-scale hydrological modelling—the development and application of improved land-surface parameterisations for meteorological models. Unpublished Ph.D. thesis, Imperial College of Science, Technology and Medicine, London, UK
Maraun D, Wetterhall F, Ireson AM, Chandler RE, Kendon EJ, Widmann M, Brienen S, Rust HW, Sauter T, Themeßl M, Venema VKC, Chun KP, Goodess CM, Jones RG, Onof C, Vrac M, Thiele-Eich I (2010) Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. Rev Geophys 48(3):RG3003. doi:10.1029/2009RG000314
Mason SJ, Graham NE (2002) Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: statistical significance and interpretation. Q J R Meteorol Soc 128(584):2145–2166
Pappenberger F, Beven KJ (2004) Functional classification and evaluation of hydrographs based on multicomponent mapping (Mx). Int J River Basin Manage 2(2):89–100
Pappenberger F, Beven KJ, Hunter NM, Bates PD, Gouweleeuw BT, Thielen J, de Roo APJ (2005) Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). Hydrol Earth Syst Sci 9(4):381–393
Pappenberger F, Thielen J, Del Medico M (2011) The impact of weather forecast improvements on large scale hydrology: analysing a decade of forecasts of the European Flood Alert System. Hydrol Process 25:1091–1113. doi:10.1002/hyp.7772
Peterson W, Birdsall T, Fox W (1954) The theory of signal detectability. Inf Theor Trans IRE Prof Group 4(4):171–212
Panofsky HA, Brier GW (1968) Some applications of statistics to meteorology. The Pennsylvania State University Press, University Park, Pa, p 224
Romanowicz RJ, Beven KJ (2006) Comments on generalised likelihood uncertainty estimation. Reliab Eng Syst Safe 91(10–11):1315–1321. doi:10.1016/j.ress.2005.11.030
Romanowicz RJ, Young PC, Beven KJ (2006) Data assimilation and adaptive forecasting of water levels in the river Severn catchment, United Kingdom. Water Resour Res 42:W06407. doi:10.1029/2005WR004373
Storn R, Price K (1997) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 11:341–359. doi:10.1023/A:1008202821328
Young PC (1999) Data-based mechanistic modelling, generalized sensitivity and dominant mode analysis. Comput Phys Commun 115:1–17
Young PC (2001) Data-based mechanistic modeling and validation of rainfall-flow processes. In: Anderson MG, Bates PD (eds) Model validation: perspectives in hydrological science. Wiley, New York, pp 117–161
Wetterhall F (2014) HBV—The most famous hydrological model of all? An interview with its father: Sten Bergström. http://hepex.irstea.fr/the-hbv-model-40-years-and-counting/. Accessed 3 March 2015
Wetterhall F, Pappenberger F, He Y, Freer J, Cloke H (2012) Conditioning model output statistics of regional climate model precipitation on circulation patterns. Nonlinear Process Geophys 19(6):623–633
Zalachori I, Ramos M-H, Garcon R, Mathevet T, Gailhard J (2012) Statistical processing of forecasts for hydrological ensemble prediction: a comparative study of different bias correction strategies. Adv Sci ERes 8:135–141
Zhu Y (2005) Ensemble forecast: a new approach to uncertainty and predictability. Adv Atmos Sci 22(6):781–788

Part V
End-User Interests

Social Aspects in Flood Risk Assessment

Dorota Rucinska

Abstract Most scientists recognize social vulnerability to natural hazards as an integral part of the risk, taking into account the population, its diversity and specificity. At the same time, we observe new phenomena such as risk amnesia, as well as the incorporation of the aspect of our sense of security associated with natural hazards (e.g. flooding) into our assessment of quality of life, the natural environment and landscape. This analysis of social vulnerability to flooding in the Middle Vistula Valley uses thirteen variables corresponding to the characteristics of the local community and the local authorities' coping capacity at a municipality and county level. Spatial distribution of social vulnerability to natural hazards in the case study area has been shown on a map.

Keywords Social vulnerability · Spatial distribution · Natural hazards · Flood risk

1 Introduction

The social aspect is an important element in the assessment of risk of losses incurred as a result of natural hazards, including flood risk. Most authors recognize vulnerability as an integral part of the risk (Davidson 1997; Bollin et al. 2003; Peduzzi et al. 2002; Schmidt-Thomé 2006; Wood 2011). Risk estimation methods have evolved over the years. Various mathematical equations were developed: the total risk was calculated taking into account overall vulnerability (Alexander 2000; UN/ISDR 2002; UNDP/ISDR 2004), exposure (Dilley et al. 2005) and coping capacity (Hahn 2003; Villagrán De Leon 2006). Another example is the risk triangle (Crichton 1999) which also takes into account the element of vulnerability.

D. Rucinska (✉)
Faculty of Geography and Regional Studies, University of Warsaw, Warsaw, Poland
e-mail: dmrucins@uw.edu.pl

R.J. Romanowicz and M. Osuch (eds.), *Stochastic Flood Forecasting System*,
GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-319-18854-6_12

Meanwhile, social vulnerability itself focuses on the population, taking into account its diversity and specificity (Cutter 1996).

The growing anthropogenic impact on the environment and technological progress are behind the fact that we no longer observe nature closely on a daily basis. We have a sense of dominance over it, and we base our knowledge and behaviour on information provided by mass media. These changes have contributed to the emergence of a new phenomenon—natural hazards amnesia. The hazard amnesia of people living in a given area manifests itself in their forgetting of past dangers the area has been subject to. The condition is also connected with a lack of knowledge about natural hazards, combined with increasingly dynamic migration. In natural hazard amnesia, the time lapse between events is of importance. In spite of repeated flooding in Poland in recent decades, we are still seeing inadequate preparation and reactions on the part of the population. Mouth-to-mouth messages have lost their significance as a centuries-old means of communication, and have been replaced by the mass media, which do not fulfil the same functions. The public lacks the ability to identify the at-risk areas and to prepare and respond to emergencies. In Poland, classroom education is insufficient in this respect, and the successive reduction of hours spent discussing natural hazards in high school geography lessons is not conducive to improving the situation (Rucinska 2012). Today, natural hazard amnesia indicates the need for a new approach to the problem of natural hazards in social education, also when it comes to flooding.

The nature of the study of how natural hazards are perceived poses numerous methodological problems. At the same time, it has been suggested that studies should aim to identify what the threat signifies to humans who come into contact with it (Saarinen 1974), and what people think and feel about potential threats (Golant and Burton 1976). In the past, research was only concerned with the perception of people who directly experienced the phenomenon or lived in the affected area. Nowadays, the dynamic process of migration and the movement of people bring about the need to study the indirect experience (resulting from verbal and visual experience from the electronic transmission of past events) and the perception of natural hazards, including that of people not living in the affected area. The results of such studies can provide important information for risk assessment.

The existence of a sense of security taking into account natural hazards in our perception of the environment was noted by Wojciechowski (1986). He observed that the perception of threats is linked to the assessment of quality of life (the quality of the environment and the landscape in which we live), which has arrived via an assessment of the likelihood of the presence of threats, catastrophes and natural disasters (Wojciechowski 2004). Variations in individual assessments of the environment and ensuing decisions about the choice of residence or place of stay are most often justified by previous experience of natural disasters, attitude towards nature, attachment to the region of residence and individual personality traits.

Human behaviour does not change automatically in line with economic and cultural changes (Brooks 1973), but tends to be individual (Ittelson 1974), which indicates a need to develop behavioural models and to implement them in risk reduction

processes. Human reactions vary (Hewitt and Burton 1971), and are associated with information about the threat, as well as the role and response of the authorities and the effects caused by extreme phenomena (Walmsley and Lewis 1997). These factors combine with an element of uncertainty and doubt, and the stimuli received sometimes conflict with each other. Individual experience contributes to the final perception of natural hazards, significantly improving awareness and the ability to not ignore warnings. Increased awareness affects the ability to take protective measures, also on an individual level (Kates 1971), and plays an important role in minimising the effects of dangerous natural phenomena (Hanson et al. 1979). The significance of individual factors cannot be overestimated, however, because human behaviour can vary greatly (Kates 1962), as can the reasons for ignoring threat warnings (Palm 1981). Age constitutes one of the most important factors. Individual sensitivity increases with age (Abney and Hill 1966). It changes in groups with higher social status and a higher level of education (Baker and Patton 1974). People's reactions are also affected by the frequency, severity and intensity of a phenomenon, and to what degree it is typical of a particular area (Kates 1971).

2 Selection of Characteristics and Social Factors Used in Risk Assessment; Individual, Group and Institutional Characteristics

A dualism characterised by opposing directions occurs among individuals and groups of people, resulting in an increase or decrease of social vulnerability to natural hazards. Moreover, socio-economic factors can also be distinguished, the presence of which is independent of the type of natural hazards and can be included in the risk analysis. These include factors determining the behaviour and ability to make decisions and to act: age, health, education, social status and sometimes sex. In the case of children and the elderly who are physically and/or mentally weaker, age contributes to a reduced capacity to react and to act appropriately in a dangerous situation. Another factor is physical and/or mental health, which in some cases may be associated with age. Education also plays a role; because better educated individuals are characterized by a greater knowledge and awareness of making decisions, as well as behaviour that is appropriate to the situation. Social status is also a fourth factor, with higher social status entailing more frequent preventive actions, for example those related to investment in the protection or modernisation of existing buildings. When it comes to sex, it is generally agreed that women are physically weaker and do not cope as well as men in the face of danger. At the same time, there exists a belief that they are mentally stronger in extreme situations, especially when they have children in their care. In view of the context of the study and the fact that women are more likely to have children or the elderly in their care (particularly applicable in Poland) and at the time of the event may be taking care of family members (a situation similar to the case of working nurses), this was considered a fifth feature impacting the ability to act. Moreover, individual personal characteristics, factors associated with the

community, its organization and management are also relevant to vulnerability to natural hazards.

When it comes to community, account should be taken of groups of foreigners choosing particular locations to settle down, and social minorities. Both of these situations can be observed in Poland, for example in Warsaw, where a large number of foreigners live. The Warsaw region stands out when it comes to groups of foreigners. In the case of ethnic minorities, special attention should be paid to the Roma community, which functions in very traditional ways and does not generally observe the customs of Polish communities. However, there is a lack of information on their location.

When it comes to an analysis of the region or local studies, there are national laws in place, as well as flood protection stemming from regional regulations and management; these features have therefore not been included in the present study as they will not affect the variation of social vulnerability involved in social risk. It should be noted, however, that variations in crisis management in Poland's seven hydrological regions can account for variations in social vulnerability to flooding.

The selected characteristics can be divided into those that deepen social weaknesses and those that build resilience and resistance. Moreover, they should be regarded in the context of vulnerability and coping capacity resulting from individual actions or the organisation and functioning of society. Although at the moment we have no access to individual data pertaining to the preventative actions undertaken by the inhabitants of the region, data on social organisation and functioning can help complete the present account.

3 Research Area and Methods

An analysis of overall social vulnerability of the region was carried out for the physico-geographical Middle Vistula Valley region, lying in the Mazowsze province and the northern part of the Lubelskie province. The administrative organisation of the province is characterized by a subdivision into districts and even smaller municipalities. The focus was set on the smallest administrative units, using statistical GUS data from the 2011–2013 period. Within the Mazowsze province, municipalities along the main research area were analysed, from the central districts of Warsaw to the town of Puławy.

Because the borders of the physico-geographical and administrative regions do not coincide, the administrative regions of the Middle Vistula Valley were used as a basis for the estimation of social vulnerability. This approach gives the preliminary illustration of the phenomenon. It should be noted that the estimates obtained correspond to the value averaged over the whole administrative unit. Therefore, the smaller the area the larger the error of the estimate, as can be seen, e.g., for the municipalities Chynów and Grabów on Pilica.

In the classification, data on social characteristics from the inhabitants of the Middle Vistula Valley or data on the activities undertaken there by state services

are used. These characteristics determine social sensitivity and the population's potential coping capacity before and during a crisis situation resulting from a risk of, or actual flooding. The following factors were used to preliminarily estimate social sensitivity to natural hazards: population density, as an exposure, pre-productive (14 years old or less) and post-productive age as a whole, (total number of women aged 60 and over and men aged 65 and over), health (using its inverse in the form of deaths due to a variety of diseases), the number of women, social migrations, the presence of foreign tourists. To describe coping capacity, the data used included unemployment, data on persons benefiting from social assistance and persons receiving child benefits. Other factors considered include earnings within the administrative unit, the detection of selected crimes, the health service (the presence of hospital rescue units and emergency air rescue teams), expenditure on security and fire protection. Unfortunately, because of the lack of data on the percentage of people with higher education and financial status, these could not be included in the analysis.

Using the theoretical vulnerability concept (White et al. 2005; Villagrán De Leon 2006) the index of overall social vulnerability to natural hazards was calculated over the area of potential flood risk, within the administrative units.

The value of synthetic vulnerability index (SVI) is defined as the ratio of exposure and social susceptibility product to coping capacity, taking into account the characteristics of variables with a division into stimulants and anti-stimulants presented in Table 1.

Social sensitivity (rows 1–6 in Table 1) and the coping capacity (rows 7–13, Table 1) were expressed as an average of the group of normalized indices. The total social vulnerability index is given by:

Table 1 Variables characterizing vulnerability in the Middle Vistula Valley

1.	Population density per km^2	Stimulant
2.	Number of women	Stimulant
3.	Total population in pre-productive and post-productive age	Stimulant
4.	Lack of health: total number of deaths due to diseases and other factors	Stimulant
5.	Registrations of residence: from cities, from the countryside, and from abroad	Stimulant
6.	Nights spent in tourist accommodation by foreign tourists	Stimulant
7.	Hospital rescue units and air emergency rescue units	Anti-stimulant
8.	Expenditure on security and fire safety in administrative units	Anti-stimulant
9.	Registered unemployed	Stimulant
10.	Poverty: persons receiving social assistance	Stimulant
11.	Poverty: children for whom parents claim benefits	Stimulant
12.	Detectability of economic crime	Anti-stimulant
13.	Earnings per inhabitant in the county	Anti-stimulant

$$SVI = \frac{\frac{1}{n}\sum_{i=1}^{n=6} a_i}{\frac{1}{m}\sum_{i=1}^{m=7} b_i};$$

where $a_i, i = 1, \ldots, n$ denotes the social sensitivity attribute and $b_i, i = 1, \ldots, m$ denotes the coping capacity.

The risk increases with increasing values of social vulnerability index. The high risk threshold is defined at the value of mean plus standard deviation.

The spatial variation in the vulnerability of the population of the Middle Vistula Valley was carried out using the ArcGIS software. A five-point and colorful scale was used in cartogram natural breaks (Jenks).

4 Results and Summary

This study's aim was the choice of variables that characterise social sensitivity to the natural hazards, including floods, and coping capacity of society in Polish conditions. These two attributes were used to estimate spatial distribution of social vulnerability indices in the region of Middle Vistula Valley, and presented in the form of a map.

The spatial distribution of the social vulnerability to natural hazards index in the Middle Vistula Valley is presented in Table 2 and in Fig. 1.

An analysis of administrative regions of the Middle Vistula Valley has shown that the most socially vulnerable regions include: Dęblin in Lubelskie voivodship, Celestynów and town Puławy. High values of social vulnerability are present in the region stretching from the west, from the municipality Kozienice (1.21) up to eastern part of municipality Ryki (1.06) and Dęblin (2.44), reaching the highest value of that index within the studied region. Heading down the Vistula River, the communes with the highest vulnerability to natural hazards include Celestynów (2.43), Józefów (1.62) and Otwock (1.57). Higher than average vulnerability occurs also in Środmieście (1.35), Mokotow (1.28), Wilanów (1.23), and situated on the right side of the River Vistula: Praga Południe (1.4), Karczew (1.26). The smallest vulnerability occurs in the municipalities: Chynów (0.02), Żyrzyn (0.04), Grabów on Pilica (0.09) and Osieck (0.13).

Social vulnerability in the surveyed area differs from the arithmetic mean (0.86) by a standard deviation value equal to 0.67. Units with heightened social vulnerability there are: Dęblin, Celestynów, Puławy City, Józefów, and Otwock. We can expect an increase of disaster risk on the tested area. There is a necessary preventive activity to focus on social aspects of habitants. Regions with social vulnerability lower than average occur in Chynów, Żyrzyn, Grabów on Pilica and Osieck. In those regions, social factors do not influence the risk related to natural hazards.

In the Middle Vistula Valley, communes with a medium social vulnerability prevail (within the 0.67 confidence limits). A very high population density usually influences the high rate of social vulnerability, as in the case of the most densely

Table 2 The values of social vulnerability index (SVI) in the Middle Vistula Valley

No.	Township/districts/municipality	SVI
1.	Chynów	0.02
2.	Żyrzyn	0.04
3.	Grabów nad Pilicą	0.09
4.	Osieck	0.13
5.	Trojanów	0.19
6.	Gniewoszów	0.23
7.	Maciejowice	0.24
8.	Sobienie-Jeziory	0.24
9.	Magnuszew	0.27
10.	Puławy, distr.	0.36
11.	Głowaczów	0.38
12.	Konstancin Jeziorna, distr.	0.40
13.	Wilga	0.48
14.	Góra Kalwaria, distr.	0.52
15.	Garwolin	0.53
16.	Stężyca	0.74
17.	Sieciechów	0.75
18.	Garbatka-Letnisko	0.78
19.	Wawer	0.86
20.	Ryki	1.06
21.	Kozienice	1.21
22.	Wilanów	1.23
23.	Warka	1.25
24.	Karczew	1.26
25.	Mokotów	1.28
26.	Śródmieście	1.35
27.	Praga Południe	1.40
28.	Otwock	1.57
29.	Józefów	1.62
30.	Puławy, city	2.14
31.	Celestynów	2.43
32.	Dęblin	2.44

populated Warsaw districts, such as Praga Południe, Śródmieście and Mokotów (with population density in the range from 6000 to 8000 people per square kilometer of land area). However, the administrative regions of Wilanów, Warka, Karczew and Puławy City, characterised by a population density that is several times lower (does not exceed 1000 people per sq. km of land area), show a relatively high index of 1.23 and 1.40, respectively (Fig. 1). The analysis has shown the importance of the complex of social characteristics, together with population density, which is regarded in many studies as the only measure of social vulnerability in the estimates included in flood risk assessment.

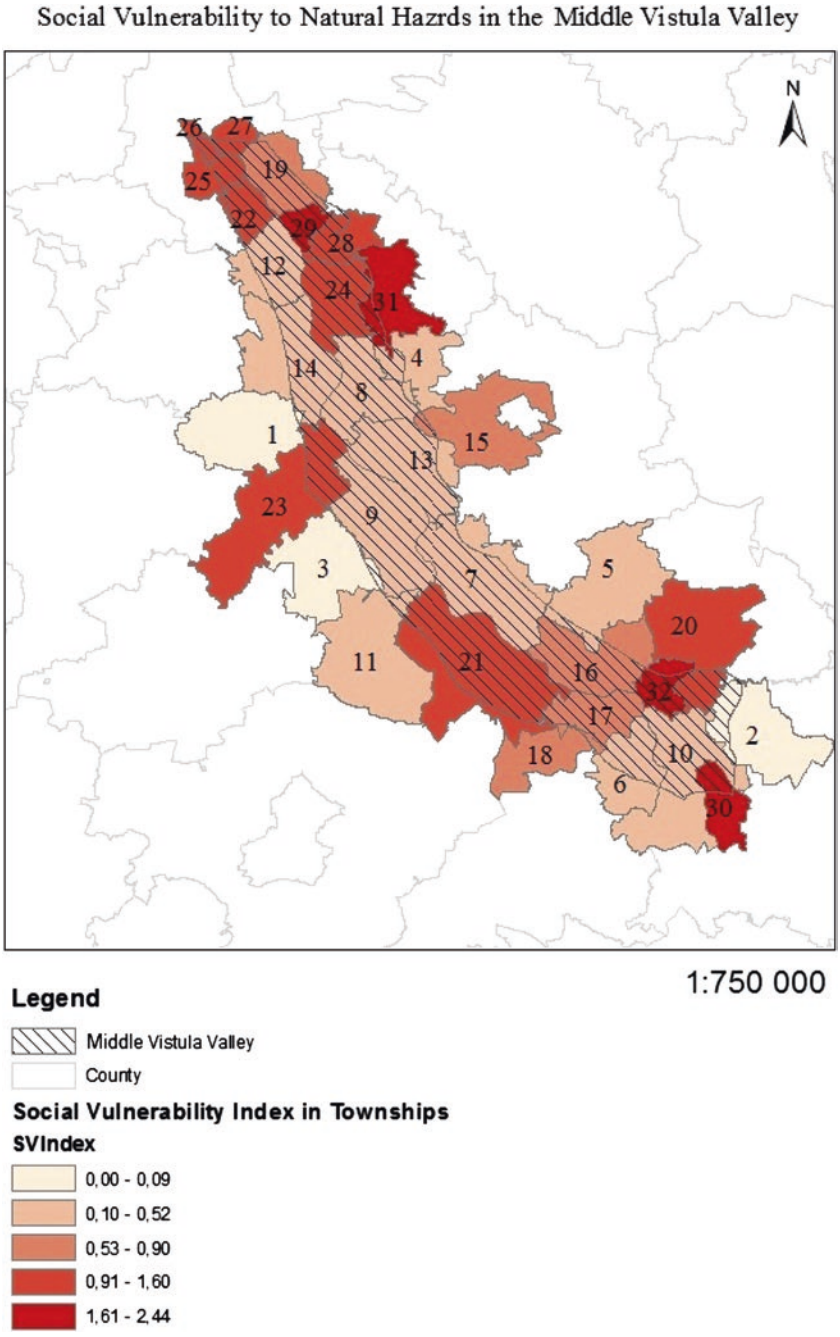

Fig. 1 Social vulnerability to natural hazards in the Middle Vistula Valley

The main novelty of the study consists of the selection of variables that characterize the social sensitivity and the coping capacity in Polish conditions. Moreover, the regions of the Middle Vistula Valley that show high and low social vulnerability were recognized. The study allows the activities and funds devoted to the reduction of risk from flooding to be allocated focusing on the social aspects.

At the current stage, the analysis of social vulnerability requires further search for the best variables to describe it.

References

Abney FG, Hill LB (1966) Natural disasters as political variable: the effect of a hurricane on an urban election. Am Polit Sci Rev 60:974–981

Alexander DE (2000) Confronting catastrophe: new perspectives on natural disasters. Terra Publishing, Harpenden, UK, and Oxford University Press, New York

Baker EJ, Patton DJ (1974) Attitudes toward hurricane hazard on the Gulf Coast. In: White GF (ed) Natural hazards: local, national, global. Oxford University Press, New York

Bollin C, Cárdenas C, Hahn H, Vatsa KS (2003) Disaster risk management by communities and local governments. Inter-American Development Bank, Washington, DC. Available at http://www.iadb.org/sds/doc/GTZ%2DStudyFinal.pdf

Brooks RH (1973) Differential perception of drought in north-eastern Brazil. Proceedings of the Association of American Geographers 5:31–34

Crichton D (1999) The risk triangle. In: Ingleton J (ed) Natural disaster management. Tudor Rose, London, pp 102–103

Cutter S (1996) Vulnerability to environmental hazards. Prog Human Geogr 20(4):24–29

Davidson R (1997) An urban earthquake disaster risk index. Report No. 121, The John A. Blume Earthquake Engineering Center, Stanford University, Stanford

Dilley M, Chen RS, Deichmann U, Lerner-Lam AL, Arnold M (2005) Natural disaster hotspots. A global risk analysis. World Bank

Golant S, Burton I (1976) A semantic differential experiment in the interpretation and grouping of environmental hazards. In: Moore GT, Golledge RG (eds) Environmental knowing, Dowden. Hutchinson and Ross, Stroudsburg, pp 364–374

Hahn H (2003) Indicators and other instruments for local risk management for communities and local governments. Document prepared as part of the documents related to the Project: Local Risk Management for Communities and Local Governments. The German Technical Cooperation Agency, GTZ, for IADB

Hanson S, Vitek JD, Hanson PO (1979) Natural disaster: long-range impact on human response to future disaster threats. Environ Behav 14:171–185

Hewitt K, Burton JS (1971) The hazardousness of a place: a regional ecology of damaging events. Department of Geography Research Paper 6, University of Toronto

Ittelson WH (1974) An introduction to environmental psychology. Holt, Rinehart and Winston, New York

Kates RW (1962) Hazard and choice perception in flood plain management, Chicago, Paper 78

Kates RW (1971) Natural hazards in human ecological perspective: hypotheses and models. Econ Geogr 47

Palm R (1981) Public response to earthquake hazard information. Ann Assoc Am Geogr 71:389–399

Peduzzi P, Dao H, Herold C (2002) Global risk and vulnerability index. Trends per year (GRAVITY) United Nations development programme. Bureau of Crisis Prevention & Recovery (UNDP/BCPR)

Rucinska D (2012) Ekstremalne zjawiska przyrodnicze a świadomość społeczna. Warszawa, Uniwersytet Warszawski, Wydział Geografii i Studiów Regionalnych, p 219

Saarinen T (1974) Problems in the use of a standardised questionnaire for cross—cultural research on perception of natural hazards. In: Natural hazards: local, national, global, Oxford University Press, London, pp 180–184

Schmidt-Thomé P (ed) (2006) ESPON Project 1.3.1 – Natural and technological hazards and risks affecting the spatial development of European regions. Geological survey of Finland

UN/ISDR (2002) Living with risk: a global review of disaster reduction initiatives. Preliminary version prepared as an interagency effort co-ordinated by the ISDR Secretariat, Geneva, Switzerland

UNDP/ISDR (2004) Visions of risk a review of international indicators of disaster risk and its management. A report for the ISDR inter-agency task force on disaster reduction working group 3: risk, vulnerability and disaster impact assessment, December 2004. http://www.unisdr.org/2005/HFdialogue/download/tp1-visionsofrisk.pdf. Accessed 12 Sept 2013

Walmsley DJ, Lewis GJ (1997) Geografia człowieka. Podejścia behawioralne, PWN, Warszawa

White P, Pelling M, Sen K, Seddon D, Russell S, Few R (2005) Disaster risk reduction. A development concern, DFID

Villagrán De Leon JC (2006) Vulnerability. A conceptual and methodological review, Studies of University Research, Counsel, Education, No. 4/2006—Publication Series of UNU-EHS

Wojciechowski KH (1986) Problemy percepcji i oceny estetycznej krajobrazu, UMCS, Habilitation degree thesis (Rozprawa habilitacyjna) Wydział Biologiii Nauk o Ziemi, 28, Lublin

Wojciechowski KH (2004) Miejsce postrzeganego krajobrazu w całościowym ujęciu jakości życia. Annales Universitatis Mariae Curie-Sklodowska, Lublin-Polonia, Vol. LIX, 13. Sectio B, Zakład Ochrony Środowiska, Instytut Nauk o Ziemi UMCS, Lublin

Wood N (2011) Understanding risk and resilience to natural hazards: US geological survey fact sheet 2011–3008, 2 p