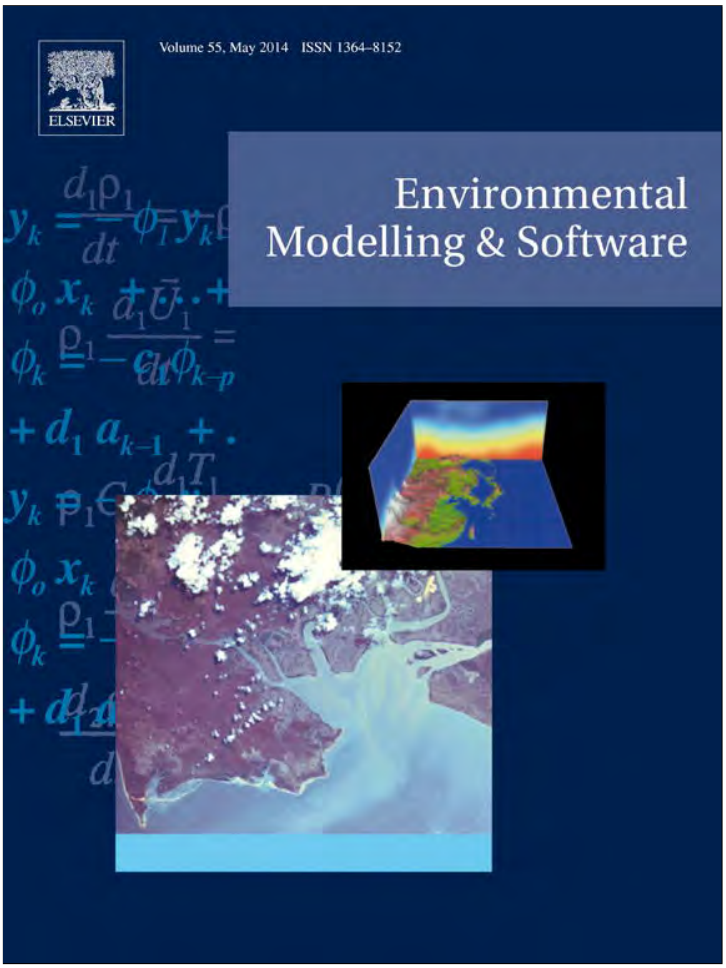

Environmental Modelling & Software 55 (2014) 250–265

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Flood inference simulation using surrogate modelling for the Yellow River multiple reservoir system

M.E. Castro-Gama [a], I. Popescu [a,*], S. Li [a], A. Mynett [a], A. van Dam [b]

[a] UNESCO-IHE, Institute for Water Education, P.O. Box 3015, 2601 DA Delft, The Netherlands
[b] Deltares, Roterdamseweg 185, 2629 HD, The Netherlands

ARTICLE INFO

Article history:
Received 15 February 2013
Received in revised form
24 January 2014
Accepted 1 February 2014
Available online 28 February 2014

Keywords:
Reservoir operation
Yellow River
Flood modeling
Surrogate modeling

ABSTRACT

The Yellow River, in China, is one of the largest hydro systems in the world. Flooding is a major problem for the river, and therefore over the last 50 years a large number of interventions have been made in its reaches and tributaries, in order to control the flooding events in the lowland area, downstream of the Huayuankou hydrological station. The development of new technologies and approaches to decision support has raised possibilities for creating new ways of managing the river and reducing loss of life, in the case of flooding, for the people living within the embankment area of the river. Given the importance of the river for the development of economic activity in China, it is essential to increase the understanding of the general flooding processes triggered by several reservoir operation scenarios, and then, after applying them to a flooding model of a specific area, to test the findings. The main goal of the research presented here is to investigate and develop the statistical inference between the operation of reservoirs on the Yellow River and a set of variables related to the downstream flooding, such as the total flooding volume and the peak discharge. The research shows that it is possible to use such inference models as decision support tools, by reducing the number of explanatory variables to be included in the simulations carried out to determine the appropriate reservoir operation.

1. Introduction

There are many models that serve as tools for decision-making in river management institutions around the world. When managing and operating single- or multi- purpose reservoirs, it is necessary to look at relevant measured variables such as reservoir storage, released discharge, reservoir water levels, water quality or total sediment in order to assess the objective of a reservoir (Hassaballah et al., 2012; Loucks and van Beek, 2005; WMO, 2008). The objectives should be analyzed during different time scales: for peak or low magnitude extreme events, seasonal approaches (e.g. winter, summer, for crops), annual approaches, or even, in some cases, for the project life span. The objectives for reservoir operation can be broad and include hydropower generation, flood management, irrigation, recreation and water supply.

The evaluation of reservoir management strategies implies a search for an optimization problem, while trying to achieve the maximum or the minimum of an objective function. The techniques used during an optimization of reservoir operations range from deterministic, stochastic, and simulation to heuristic approaches. A number of years ago Yeh (1985) studied most of the operation research techniques used for the development of reservoir operation for simplified schemes and found that these models were difficult to adapt to real operational applications, when modifications were required to respond to changes on a daily basis. Rani et al. (2010) compiled an extensive overview of the simulation and optimization techniques used for reservoir operation under climate change conditions, in view of the continuous need for adaptation strategies for flood management. Main conclusion is that nowadays optimization methods require extensive amount of data and information of the systems modeled (Celeste and Billib, 2009) and long running times to achieve calibration (Dittmann et al., 2009), therefore models can not be easily updated to be used in daily reservoir operations.

While implementing and developing any of the above approaches, one issue is mentioned every time; the "curse of dimensionality". When dealing with modeling flooding events the complexity of the task is determined by the identification of the proper variables and constraints that drive the optimal operation,

* Corresponding author. Tel.: +31 15 2151895; fax: +31 15 3122921.
E-mail addresses: m.castrogama@unesco-ihe.org (M.E. Castro-Gama), i.popescu@unesco-ihe.org (I. Popescu), s.li@unesco-ihe.org (S. Li), a.mynett@unesco-ihe.org (A. Mynett), arthur.vandam@deltares.nl (A. van Dam).

http://dx.doi.org/10.1016/j.envsoft.2014.02.002
1364-8152/

Acronyms (alphabetical order)

ALL	The set of explanatory variables used in the regression model
DoF	Degrees of Freedom
FMA	Maximum flooded area (from Huayuankou to Gaocun hydrological stations)
FMV	Maximum flooded volume (from Huayuankou and Gaocun hydrological stations)
FTA	Time to reach maximum flooded area
FTV	Time to reach the maximum flooding volume
G	Guxian reservoir
GAM	A set of 1000 Gamma Function hydrographs
G10	A set of 10,000 Gamma Functions hydrographs
HQa	Average Discharge at Huayuankou hydrological station
HQp	Peak discharge at Huayuankou hydrological station
HTa	Average time of hydrograph (centroid) at Huayuankou hydrological station
HTp	Time to peak discharge at Huayuankou hydrological station
HVL	Total hydrograph volume at Huayuankou hydrological station
INI%	Initial Water Level in a reservoir as % from total reservoir water depth
L	Luhun Reservoir
MAX%	Maximum Water Level in a reservoir as % from total reservoir water depth
MR	Multivariate Regression
Qb	Base discharge of a Gamma or Triangular inflow hydrograph in a reservoir
Qp	Peak Discharge of a Gamma or Triangular inflow hydrograph in a reservoir
r_t	correlation of training
r_v	correlation of validation
S	Sanmenxia Reservoir
SK	Skewness of the inflow Gamma type of hydrographs
SSE	Square Summ of the Estimated Residuals
SSR	Square Summ of Residuals
STO	Summ of Total errors
TRI	The set of 1000 Triangular hydrographs
Tp	Time to peak of the Gamma or Triangular inflow hydrograph
UNI	the set of 1000 Uniform (constant) hydrographs
X	Xiaolangdi Reservoir
X_{SEL}	The set of selected explanatory variables

and any reduction of the number of variables is important for the modelers in their timely response to decision makers. In order to address the issue of computational time, one of the approaches to solve reservoir operation is surrogate modeling, in which a model is build to represent the functionality of the reservoir system by using a set of explanatory variables for the reservoir operation in order to estimate the flooding in the downstream of the reservoir. The extent of flooding downstream of the reservoir system is described by a set of dependent variables. Castelleti et al. (2012) classified surrogate models in two categories; response surface modeling, in which the set of dependent variables are represented as a resulting surface and lower fidelity modeling, in which a less complex model (with less explanatory variables) is run in order to reduce the computational effort and still maintain the accuracy of the results. The main applications of surrogate modeling are mainly to be found in statistics and computer science. (Hussain et al., 2002; Gano et al., 2006; Ratto and Pagano, 2010). Castelleti et al. (2012) also points out that just recently surrogate modeling emerged in environmental related fields. In agro-ecosystems planning, Pineros-Garcet et al. (2006) used metamodeling for nitrate leaching and Audsley et al. (2008) used a surrogate for estimating the socioeconomic impact of climate change on agriculture in the UK. For water distribution system optimization Broad et al. (2005) used metamodeling for a design problem. Recently, Castelleti et al. (2010) used linearization of multiobjective surface responses for water resources emulation. Castelletti et al. (2007) used a neuro-dynamic programming approach for designing optimal reservoir network management policies. Ravazi et al. (2012), presents a review of different surrogate modeling approaches specifically for water resources. However, cases on reservoir operation for predicting flood variables were not reported by the review.

The main challenge in using surrogate modeling is the determination of the proper set of explanatory variables to be used as input for a model as shown by Bowden et al. (2005). The same authors classify the procedure for selecting the explanatory variables in three categories; knowledge based, mainly performed as an initial filter of data in an early modeling stage; statistical cross-correlation and heuristic information based selection. A discussion regarding the strategies to reduce the number of explanatory variables in case of complex design problems can be found in Shan and Wang (2010a, 2010b); Ratto et al. (2007); and Young and Ratto (2009). All presented strategies are mostly based on sensitivity analysis. Hejazi and Cai (2009) developed a heuristic approach called “modified minimum redundancy maximum relevance” (mMRMR), and coupled it with a generalized regression neural network model (GRNN) as defined by Specht (1991). The metamodel obtained was used to solve the case of day by day release forecast for 22 reservoirs in California, USA.

The approach proposed herein is a simulation-based one, which aims at supporting decision-making during a flood event. The method offers the advantage that modelers can learn the behavior of the system by testing the effect of different inflow hydrographs on the operation strategies for the reservoir in order to avoid flooding downstream. The surrogate model can offer support to the operators by allowing them, in case of extreme flooding events, to compare the incoming observed hydrograph with a known, already tested incoming hydrograph into the reservoirs. Research tries to address two main issues: will the operator be able to anticipate the behavior of the system, by using the knowledge developed by the surrogate model?; and based on the predicted behavior of the system can modelers give advice to the decision makers so that the operation strategy is adapted according to the new conditions.

The approach aims to provide a surrogate model that reduces the day-to-day modeling computational effort of reservoir operations for flood management, by building a predefined model that could be updated with new explanatory variables. This would eliminate modelers' guesswork of the process and reduce the number of explanatory variables for the state of the hydro-system.

In its current state the research is exploratory and it is not intended to be used at the YRCC. However, the implication of the results suggests that it could be implemented for flood management for real data hydrographs and that more research should be performed on flooding patterns, extending the research to computational intelligence techniques.

The paper is structured in five parts. Section One is the introduction. Section Two contains a description of the characteristics of

the Yellow River, along with the features that make it special from a hydraulic and managerial perspective. In Section Three the concept and the methodology applied to simulate the flooding events are detailed. The proposed statistical inference model for the flooding of the system is also introduced. Section Four presents the obtained results and a discussion. Finally, the conclusions and analysis show how an inference model, such as the one presented here, could be useful for decision making during flooding events and for the understanding of the cause-effect behavior of reservoir-driven rivers.

2. Case study: the Yellow River system

The study presented here looked at the middle section of the Huang He River, most commonly known as the Yellow River, in China. The river basin covers a vast area composed of three main regions (Fig. 1):

1) the upper river section, with the largest proportion of the basin, is known for its complexity with regard to water and sediment volumes. It is located between the Sanmenxia Reservoir and the spring of the river in a mountainous region in the Qinghai province. Most of the total sediment volume of the river (90%), and more than half of the total flow (56% of the water volume) originates in this area.
2) the mid Yellow River, which is characterized by the system of reservoirs developed for the management of the river in case of extreme flooding events and used for irrigation, water supply and hydropower generation in normal hydrological conditions. The river is mainly conveyed inside dykes.
3) the most downstream area of the river, as from Dongpinghu station to where the river waters are discharged into the Bohai sea.

The Yellow River has several special aspects that make it unique in the world. The main aspect is the high ratio of sediment to water volume. There is a small amount of annual normal natural runoff of the basin, of about $58*10^9$ m^3, while the annual sediment transport is $1.6*10^9$ t. Among the 13 large rivers of the world whose annual sediment transport exceeds $100*10^6$ t, the Yellow River shows the largest amount of sediment transport and average content.

Another special aspect of the Yellow River is the so-called "hanging river", as its most downstream part is usually called. This is due to the fact that the riverbed is, on average, higher than the land surface by 4–6 m. The lowland area is protected from flooding by high embankments that delimit the river banks. Around the city of Kaifeng the difference between the Yellow River bed and the land is approximately 13 m. The highest differential between the riverbed and the land occurs at Xinxiang city, which is around 20 m. This aspect is an important element to be analyzed in any flooding study of the area and has no parallel in the world.

Flooding on the Yellow River has been an important issue throughout history. Records dating from the pre-Qin Dynasty show that the river had overflowed 1590 times and changed course 26 times in the previous 2542 years which means that on average there is a flooding event twice in every three years and one course change every century (Guoying, 2010). The flooding events occur mainly in the areas from Tianjin city (between Hekou and the Sanmenxia Reservoir) and Jianghuai.

A flooding event along the river can vary in time duration, and given the catchment characteristics, it is known that flooding events can develop ranging from a week to more than one month. Adaptation strategies and reduction of flood risk have always been very important, and therefore the overall water resources management and the operation of its structures have been entrusted to YRCC (Guoying, 2010).

2.1. The Yellow River structures

In the middle reach of the Yellow River, there is a system of reservoirs that is combined with a series of dyke structures as the main elements for flood protection. The reservoir system is composed of four reservoirs (Fig. 1); two of them in parallel (Guxian and Luhun) along the Yihe and Yilohue tributaries; and two of them in series (Sanmenxia and Xiaolangdi) located on the main reach of

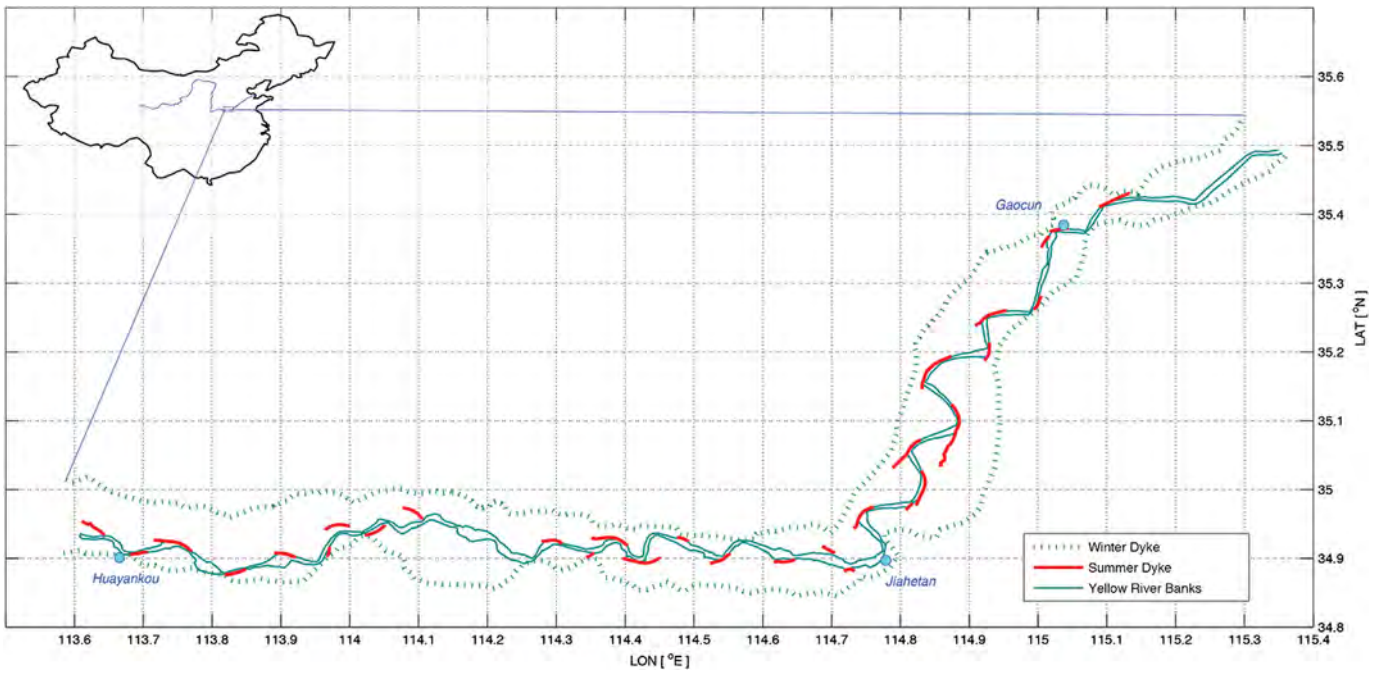

Fig. 1. The Yellow River location and the study area: the Middle Yellow River reach.

the Yellow River. The point at which the water from the four reservoirs comes together is immediately before Huayuankou, where a hydrological station of vital importance for the river management exists. According to this configuration the system can be defined as complex (Lund and Guzman, 1999). The largest reservoir in terms of storage capacity is Xiaolangdi ($101.20*10^9$ m^3) while the smallest is Guxian ($11.09*10^9$ m^3). The combined resulting capacity of $179*10^9$ m^3 and the maximum simultaneous release (discharge) of 50,156 m^3/s are amongst the largest hydro-systems in the world (Table 1). The main reach of the river downstream of the reservoirs' region is composed of a main channel and lateral embankments that are enclosed between the dykes. The main channel is not able to convey a hydrograph with a peak higher than 4,000 m^3/s, without inducing a flood event in the lateral lowland areas.

In order to reduce the risk of flooding by minimization of the flooding peak discharge, and in order to reduce the evacuation tasks in the lowland area inside the dykes, the four reservoirs on the Yellow River were designed as the primary solution at the beginning of the 20th century. In chronological order the first reservoirs built are the ones along the Yihe reaches, Luhun and Guxian, and then the largest two reservoirs, Sanmenxia and Xiaolangdi. The last reservoir built was Xiaolangdi, which entered in operation during 2001. More reservoirs are planned to be built in the future, but as of now this will take more than 10 years to become a reality, and therefore they are not taken into consideration in this study.

Downstream of the reservoirs, after the confluence of the branches of the river, at Zhengzhou city, lateral dyke structures were built in the river for flood protection. This large dyke structure (1,371 km long) allows the development of the economy outside of the main reach of the river, but it is almost impossible for Chinese people living inside the dyke area (1.97 million inhabitants) to evacuate and to avoid large flooding events without incurring loss of life (Guoying, 2010). In order to increase confidence during extreme events and to cope with discharges for flooding events with a return period of 1/100 years (16,000 m^3/s) at Huayuankou station, both dykes have been raised during the last four decades. Downstream in the river there are two storage areas, Beijingdi and Dongpinghu, which are part of the system as well, the latter originally being a natural lake. The YRCC uses these storage areas only if the input hydrograph at Huayuankou is over the established threshold of 16,000 m^3/s. These storage areas are located outside of the study area and therefore are not part of the modeling process.

2.2. Managing flooding events

When dealing with flooding events, river management authorities from the YRCC have to address the questions of when to evacuate, which areas are more susceptible to flooding and how long it would take for a flooding event to be routed out of a basin. In China, effort is constantly being made for the improvement and development of software tools for flood management (Balica et al., 2013; Gichamo et al., 2012; Popescu et al., 2010; Van et al., 2012) and currently research is being conducted to finalize an integrated tool for the management of the complex existing reservoir system (Li, 2013).

This research relates to the above aim and deals with the relations between the operation of reservoirs and flooding events.

The Yellow River uses a numerical model, also known as "digital Yellow River", for day-to-day modeling of the flow processes and it constitutes as the main decision support tool for decision and management of the reservoir system. The main outcome of such a model is a single flooding response with each run. This response can be used to forecast the amount of flooded area in the downstream river valley, or the length of time for maximum flooding extent, but because of the extensive time required for the simulation process to be performed, this cannot be done in a multi-scenario approach. In this regard it should be stated that during a flooding event operational meetings are performed almost every 4–8 h (depending on the case) and decisions are made each time on how to operate the reservoirs based on real-time forecasts. In order to operate the reservoirs properly in the case of extreme events, it is necessary to have the largest amount of information possible at all times, about the possible river behavior. If the current existing model is operated during emergency flooding events, then the model responses will be gathered too late, and the answers they provide after the decision has been made will be useless (independent of the model accuracy).

At YRCC, decision makers may not wait for the digital model results while taking decisions because the cost will be loss of life. Moreover, in addition to the computation requirements, several other explanatory variables need to be considered before running the model, such as the operation of the reservoirs, the objectives of each stakeholder at the table, and the number of possible scenarios for response provided by the digital model. Such a huge task is difficult to address in an emergency situation.

To ease this work it is common that preparation for such events is done well in advance by simulating different possible flooding events and the estimated results are taken into account.

In this paper the proposed surrogate model helps the decision makers in addressing this issue. Moreover by using such models the memory and experience of past events is kept with the model itself, with its database of results, and it is not lost whenever a new modeler takes over the task of running the model. As such decision makers do not need to depend on a small group of people who have expertise because of using the model, but they can rely on the model itself.

2.3. Managing of the Yellow River sediments

One of the main features of the Yellow river is the sediment volume and its transport. The water volume conveyed by the Yellow River is small if compared with the other main river in China, i.e. Yangtze. Yellow River carries on average only 1/17 of the Yangtze's River water volume, however the annual sediment load of Yellow River is 3 times higher than the one of the Yangtze River

Table 1
Main characteristics of the Yellow River reservoirs.

Variable	UNIT	Xiaolangdi	Sanmenxia	Luhun	Guxian	All
Latitude	DEG N	34°55'23.58″	34°49'42.99″	34°12'1.13″	34°14'22.39″	
Longitude	DEG E	112°21'51.13″	111°20'43.74″	112°10'57.25″	111°16'41.93″	
Minimum Level	masl	200	290	290	478	
Maximum Level	masl	275	335	333	551	
Minimum Storage	$10^9 m^3$	0.2	0.0	0.3	0.1	
Maximum Storage	$10^9 m^3$	101.2	54.6	12.5	11.1	179.3
Minimum Release	m^3/s	2,830.8	0.0	112.0	150.0	
Maximum Release	m^3/s	16,030.1	15,060.0	5638.6	13,427.0	50,155.7

(Guoying, 2010). The heavy sediment load of Yellow River imposes the need to be properly managed by YRCC. As a solution the Sanmenxia dam was built, on the main branch of Yellow River and put on operation on 1960 The presence of Sanmenxia reservoir changed the morphological equilibrium of Yellow River (Wang et al., 2005). Historically, it has been recorded that the riverbed can be scoured 10 m in 24 h (Wan and Wang, 1994), however with the operation of Sanmenxia reservoir a balance of sediment is maintained by using the transport capacity of the river itself to flush and carry the sediment during non flooding season (Wang et al., 2005).

The existing digital Yellow River model helped understanding better that though the morphological stability of the lower Yellow River reach is not achievable, the sediment loads are manageable due to the existing regulation policies (Li, 2013; Liu, 2012). The management of the sediments is primarily done during the summer and outside of the flooding season (Wu, 1992; Li, 2013).

During flooding events YRCC manages just the water volumes that are released through reservoir operation, in order to minimize damages in the downstream areas.

The model developed for the present research aims at addressing the operation of reservoirs in case of extreme flooding events. Sediment modeling was not included in the proposed methodology model is valid under current morphological conditions to inform daily operation under flooding emergency. After each flooding event the in situ situation needs to be analyzed and if morphological conditions have changed significantly, then model needs to be re-calibrated.

3. The surrogate model

3.1. Model description

Present research proposes a surrogate model that uses a polynomial multivariate regression meta-model in order to address the reservoir operation of the Yellow River. The selection of input explanatory variables for the regression model is done based on the decision makers knowledge of the system. Li (2013) presents in detail how decision makers meetings are taking place at YRCC and what are the most important variables to consider when dealing with floods. The initial selection of explanatory variables is 20. In order to reduce further the computational time of the model, correlation analysis is used for the selected number of explanatory variables. Stepwise (backwards or forward) approach can also be used to reduce the number of variables however when used in this research it did not show significant reduction of the number of explanatory variables (Castro Gama et al., 2013).

It is now known that climatic variability is a part of the modeling process (Milly et al., 2002; Stocker and Raible, 2005) and during simulation of flooding events the typical approaches for setting up models based on design flows seem to lack significance when these are created based only on historical records (Merwade et al., 2008; Jung and Merwade, 2012). Yellow River historical records show that the shape of the hydrographs at Huayuankou hydrological station changed in time from tall and thin to low and wide from 1962 to 2003 (He et al., 2008). In order to comply with this behavior current research uses synthetic hydrographs, of diverse shapes, to simulate extreme flood events entering the Yellow river reservoir system. The set of variables which define the hydrograph shape has been created using a Monte Carlo analysis.

Moreover, in the case of wide large rivers, such as the Yellow River (Fig. 1) or the Yangtze River, in China, and those in Colorado (Gilmore, 1999) or Tennessee, in USA, determining a representative number of explanatory variables for each reservoir becomes extremely useful. With the implementation of software modeling tools for reservoir management (Li, 2013), and given that the main objective in the operation of a particular reservoir system (such as the one analyzed here) is flood management (Guoying, 2010; Popescu et al., 2012; Jonoski and Popescu, 2011), obtaining the optimal downstream scenario of the reservoir operation makes it difficult for decision makers to perform an analysis of the magnitude and outcome of an extreme flooding event.

The methodology presented here is demonstrated on the Yellow River. With the insights given by the polynomial regression model, it was possible to forecast the behavior of the river in the case of extreme flooding events at the Yellow River Huayuankou hydrological station.

The study carried out analysis in two steps:

- an analysis and selection of explanatory variables of the reservoirs and incoming hydrographs which influence the pattern of the hydrographs at Huayuankou hydrological station; and
- an analysis of the flooding pattern (e.g. area, volume) by modeling the floods with an integrated 1D-2D unstructured grid model up to the most downstream observation station located at Gaocun.

For the second subject of analysis, a new multivariate regression was performed between the explanatory variables of the first approach and the results of flooding events due to the 1D-2D model between Huayuankou and Gaocun. In similar reviewed studies (Schmitz and Cullman, 2008; Saavedra et al., 2010), the use of large hydrological databases was used in combination with computational intelligence methods, but due to the lack of this database or a hydrological input in the current research, the operation of reservoirs is explicitly driven by the incoming hydrographs. It is important to mention that, ensemble forecasting for online reservoir operations was not used in the research presented here, because the hydrological input is not different to the incoming hydrographs. In general, the second analysis showed that this is an ongoing field of study because low correlations were obtained between the explanatory variables of the reservoirs and the flooding variables.

Fig. 2 presents the schematization of the simulation model and its components along with the surrogate model. A set of explanatory variables is generated for the operation of the reservoir system using Monte Carlo simulation. Inputs to the reservoir system are flood hydrographs that are routed by the reservoirs through operation rules. Releases from the reservoirs are routed to Huayuankou hydrological station. The routing generates a set of output hydrograph variables. The obtained hydrographs are used as upstream boundary condition for a 2D flood model between Huayuankou and Gaocun hydrological stations. This generates a second set of output variables related to flooding.

A multivariate regression model is built between the set of explanatory variables and the sets of output variables related to hydrographs at Huayuankou and the flooding variables. The input variable selection for the multivariate regression is performed using correlation analysis. The best fitted models are selected based on analysis of variance of the regressions and model training correlation.

3.2. Mathematical formulation

Given a river system of reservoirs, *Xall* is a vector of n explanatory variables that are the input for the operation of a reservoir system and for a flood routing simulation model:

$$Xall = \{X_1, X_2, X_3, \ldots, Xn\} \tag{1}$$

Let Y be the set of m variables as a total set of the variables describing the reservoir output hydrographs (Y_H) and the set of

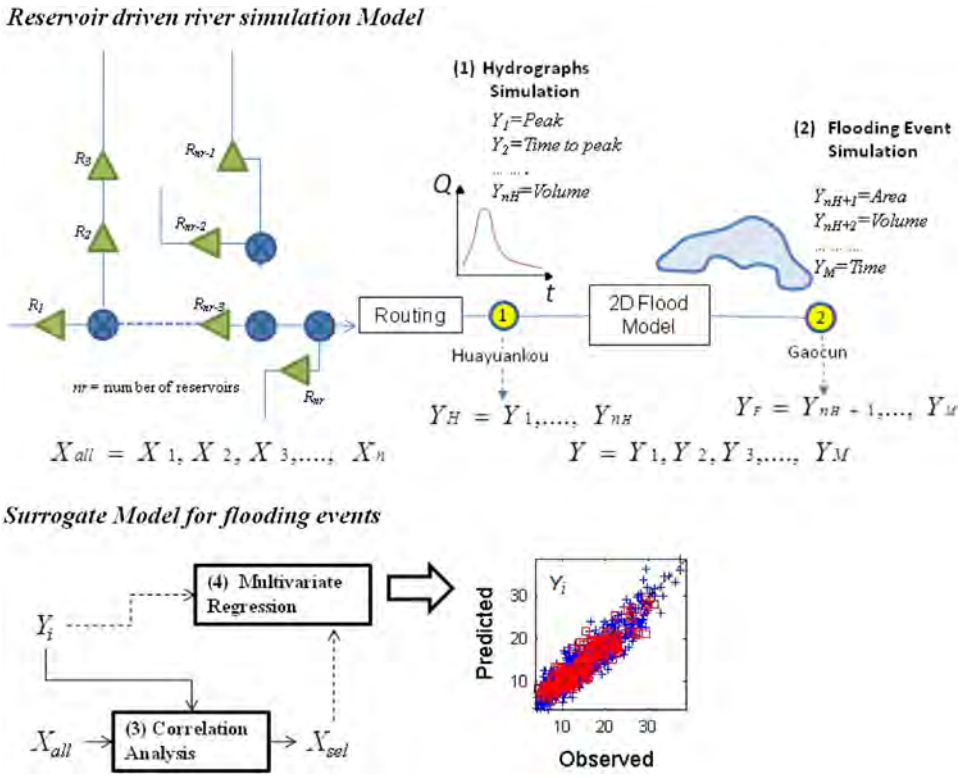

Fig. 2. The surrogate model.

variables describing the flooding (Y_F), as they are obtained by running a simulation model with *Xall* as input parameters:

$$Y = \{Y_H \mathrm{U} Y_F\} = \{Y_1, Y_2, Y_3, \ldots, Y_m\} \quad (2)$$

Corresponding to *Xall* there is a selected set of explanatory variables *Xsel*, for which a maximum correlation (ρ) with each Y_j can be obtained:

$$X_{sel} = \{X_i \in X_{all} | \rho(X_{sel}, Y_j) = \mathrm{MAX}\} \quad \text{with} \quad j = 1, \ldots, m \quad (3)$$

With each set of selected explanatory variables *Xsel*, a multivariate regression model is created for each *Yj* and a forecast $\widehat{Y}_j$ is determined:

$$\widehat{Y}_j = f(X_{sel}) \quad (4)$$

Several multivariate regression models are created with the same set of explanatory variables *Xsel*, and the selected best fitted model is the one that provides the lowest square sum of residuals.

$$\mathrm{SSR} = \sum_{S=1}^{n_s} \left(\widehat{Y}_j - Y_j\right)_S^2 \quad \text{with} \quad S = 1, \ldots, n_s \quad (5)$$

where n_S is the number of samples considered.

3.3. Limitations of the surrogate model

There are several limitations when using a surrogate model as the one presented above. The first limitation is given by the input variable selection that will be addressed by using only correlation analysis. The major disadvantage of this technique is that is able to detect only linear dependence between variables, denying the existence of non-linear relationships. However other techniques such as heuristic input selection have the disadvantage that with the increase of the dimensionality the computational effort becomes higher.

The second limitation of the proposed surrogate is that the best fitted multivariate regression model of the surface response of *Y* can only be used inside the range of variability of the *Xsel*. However, as presented in the discussion section, a comparison with more sophisticated tools provided similar results, while this approach has the advantage of a reduced complexity of the obtained model.

4. Application of the surrogate model to Yellow River

The application of the surrogate model for the Yellow River is detailed below. The steps followed in the application of the model are highlighted in Fig. 2.

4.1. Explanatory variables

A subset of explanatory variables, related to two main inputs for the models, is used; the type of hydrographs and the operation of the reservoirs. The input variable subset was selected after discussions with specialists from the YRCC (Li, 2013).

4.1.1. Incoming hydrographs

Climatic variability is one of the factors explored for the selection of hydrographs as input for the reservoir system. This is done for two main reasons; one is to try to address the curse of dimensionality (Labadie, 2004; Yeh, 1985; Wurbs, 1993) for reservoir operations by incorporating simple variables in the subsequent regression model, and the second is partly due to a possible lack of data in the regression model based on historical records, and because a subset of explanatory variables can effectively reduce the requirements without reducing the accuracy of the flood forecast. Present study did not use rainfall-runoff modeling, Three kinds of

upstream hydrographs were defined: a) Gamma functions (NERC, 1975; Todini, 2007), b) triangular functions; and c) uniform values.

In the case of Gamma functions, four explanatory variables are required in order to represent the base flow (*Qb*), peak flow (*Qp*), time to peak flow (*Tp*) and skewness of the hydrograph (*SK*), for a total of 12 explanatory variables, because Sanmenxia and Xiaolangdi are located in series, so only three hydrographs are required in each simulation.

In the case of triangular hydrographs, only 9 explanatory variables are used because the skewness coefficient becomes irrelevant.

In the case of uniform type of hydrographs, only 3 explanatory variables are useful. A comparison with steady flow could be made to measure the relevance of the peak discharge and the time to the peak discharge in the analysis section of this study.

The ranges for the base discharges and peak discharges of the corresponding basins are based on hydrological values provided by the YRCC, and are presented as box-plots in Fig. 3. The following variables are represented; *Qb*: Base flow of the input hydrograph; *Qp*: Peak discharge of the input hydrograph; *Tp*: Time to the peak of input hydrograph; *Sk*: Skewness, shape of hydrograph; *G*: Guxian Reservoir; *L*: Luhun Reservoir; *S*: Sanmenxia Reservoir; and *X*: Xiaolangdi Reservoir.

A total of 1000 combinations of input hydrographs were synthetically generated by random generation in the ranges specified. The same explanatory variables were used for each type of incoming hydrograph (Gamma, triangular, and uniform). The variable selected for the uniform value was the peak discharge.

The total time, of each generated hydrograph, is 60 days, with values defined at intervals of 2 h. After the 60 days, the base flow was taken as a uniform value until reaching 100 days, for each of the considered hydrographs.

4.1.2. Reservoir variables

Two explanatory variables were defined for each of the four existing reservoir operations, the initial reservoir level (INI%), and the maximum reservoir levels (MAX%). They both incorporate the flood management, and it results in a total of 8 explanatory variables. This variable is usually described in literature as the Flood Limited Water Level (FLWL) (Li X.et al., 2010; Li S., 2008). Both variables were taken as percentages. The range of initial reservoir was selected to be larger than 50% and less than 99%, to take into account reservoir operations that are related to an almost full reservoir. Given this restriction, the maximum capacity of the reservoirs (MAX%) was restricted to the range of 75%–100%. The ranges of values used for INI% and MAX% are presented in Fig. 3. A set of 1000 pairs of initial and maximum levels was randomly generated.

4.2. Operation of reservoirs

The specialists at the YRCC suggested that the reservoir system is operated based on an explicit reservoir level pool method (Chow et al., 1988). There are three basic operation rules, which are performed based on the existing reservoir volume at the moment that an event occurs. Firstly, if the reservoir is full then the input is released, and secondly, if the level given by the balance is lower than the minimum, the release is restricted to a minimum ecological flow. This ensures a minimum flow in the summer period and/or ecological flow (Dong, 2007). Thirdly, between the maximum and the minimum water levels in the reservoir, the goal is to store as much water as possible. The present model is developed in accordance with these rules.

4.3. Flood routing between branches (1D)

To connect the discharges between releases of the reservoirs, a Muskingum–Cunge model was applied (Cunge, 1969; Todini, 2007). The data used for this model are based on data surveys performed every year on the river by the YRCC. The time discretization used for the routing scheme is 2 h, for a total simulation period of 100 days. The result of the connection of all branches and reservoirs is a hydrograph at Huayuankou.

After obtaining the hydrograph, variables such as the peak discharge of the hydrograph (*HQp*), the time to peak discharge of

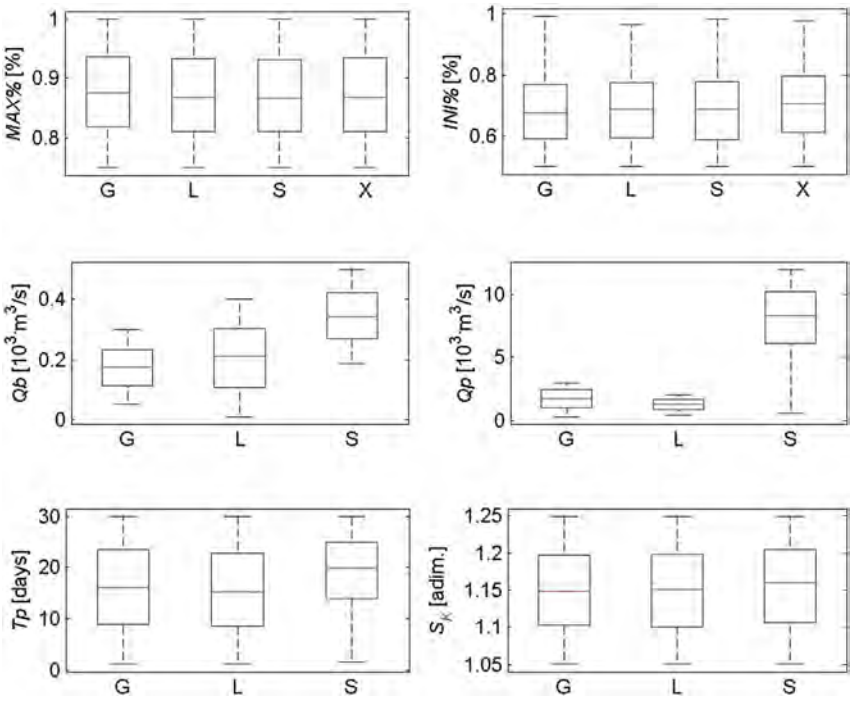

Fig. 3. Explanatory variables.

the hydrograph (*HTp*), the average discharge of the hydrograph (*HQa*), the average time of the hydrograph (*HTa*) and the total volume of water routed at Huayuankou (*HVL*) were estimated. This set of variables is called hydrograph variables.

The results in each case vary, depending on the type of incoming hydrographs into the reservoirs.

4.4. Flood modeling methodology

The resulting hydrographs at Huayuankou provide the input for a flooding model that covers the downstream part of the river, until the Gaocun station.

The total number of flooding models used was 339. The models were run for a total combined runtime of approximately 2000 h, split into several computers of different configurations and processing capacities. The results were stored in separated netCDF files. Finally, the flooding variables were calculated. The new variables comprise: 1) the maximum flooded area (*FMA*), 2), the time required for this maximum area to be flooded (*FTA*), 3) the maximum flooding volume (*FMV*) and 4) the time required for maximum flooding volume to occur (*FTV*).

4.4.1. Flooding model (2D)

A coarse rectangular grid model has been developed at YRCC for flooding simulations, however the flooding extent were not well represented by the model (Guoying, 2010). In the present research the alternative was to change the basic network and reduce the number of cells without loss of accuracy. In order to do this an unstructured grid approach was selected. The advantage in the use of unstructured grids is that it allows an increase in the refinement for narrow areas, while using coarse refinement in areas where solutions are not fundamentally required (Hartanto et al., 2011; Kernkamp et al., 2011). The modeling tool used was D-FLOW FM*βeta*,[1] which has a Finite Volume Method solution solver of the vertically integrated 1D2D shallow water equations (Kramer and Stelling, 2008; Kernkamp et al., 2011).

The flooding model represents the flooding between Huayuankou and Gaocun and inside the dykes, for a total river length of 240 km. No storage areas are located inside the modeled area.

The data for the creation of network are based on freely available Shuttle Radar Topography Mission data of the area (CIAT, 2004) and transformed into a projected system of coordinates, UTM 50. The use of this information for long wide rivers has been tested before for the Amazon River (LeFavour and Alsdorf, 2005) and Mississippi (Quinn et al., 2010; Muste et al., 2010).

A 3 m water level in the river was provided as the initial condition of the model. The calibration of this model was based on knowledge of the solution of a flooding event of 3400 m^3/s, where no flooding occurred outside of the main channel of the river. The obtained water levels are the initial condition for each flood simulation.

The Manning roughness of the system was set to 0.025 to match data from a preexisting model of the river. No distinction was made between the main channel and embankments. Coincidently this value is also suggested for the Amazon (LeFavour and Alsdorf, 2005).

The upstream boundary condition is set to be the hydrograph obtained at Huayuankou due to routing of the hydrographs that are obtained due to reservoir operation. For the downstream boundary condition, a rating curve at Gaocun has been developed by the YRCC based on several surveys.

[1] Developed by Deltares Software Centre.

4.4.2. Calibration of flooding events

Several studies on floods have focused on calibration of known extreme events with different approaches, by varying the sets of information and sources of data from the topographic, hydraulic and hydrological point of view, while at the same time trying to reduce the sources of uncertainty (Dinh et al., 2012; Gichamo et al., 2012; Hunter et al., 2005; Merwade et al., 2008; Moya Quiroga et al., 2013; Schumann et al., 2009). Design hydrographs or recorded data were used to keep the uncertainty of the input out of the scope, and the calibration was performed for roughness coefficients in the main channel and lowland areas of the river. The results obtained from these studies are often presented as a flood map. In general, two mutually exclusive approaches were taken to classify the results into deterministic or probabilistic maps. A critical discussion of these approaches was reviewed (Schumann et al., 2009), but there is no case study with a reservoir-driven system calibrated for spatial flooding under different operation scenarios. This particular omission is of extreme importance for the YRCC authorities. If a calibrated hydraulic model is already available to decision makers, then knowing which explanatory variables of the reservoir are more relevant to avoid flooding in lowland areas is more useful rather than just knowing the probability of flooding in a particular cell of the model.

4.4.3. Correlation analysis

Having the set of explanatory variables (*Xall*), reservoir output hydrograph and flooding variables, the simulation set was split in two sets: 800 value sets for training of the model; and 200 for validation of the results.

With the training group of simulations, an analysis of correlation was performed between explanatory variables and the resulting hydrograph variables.

The hypothesis of no correlation against the alternative that there is a non-zero correlation was used. The significance of the correlation was based on *p*-values bigger than 0.05, that creates a selection of relevant explanatory variables for each hydrograph and flooding variable. The new subset of explanatory variables is called *Xsel* (selected). The set varies depending on the hydrograph variable, the flooding variable and the type of hydrograph.

4.5. Multivariate regression

A multivariate linear regression model was developed between the set of selected explanatory variables (X_{sel}) as predictors, and the hydrograph (Y_H) and flooding variables (Y_F) as dependent variables (Y). The data was partitioned using a cross-validation type splitting with hold-out, as detailed by Bennett et al. (2013), into a training set of 800 samples and a validation set of 200 samples.

In order to obtain the correlation coefficients $\widehat{\beta}$, the least squares estimation was used:

$$\widehat{\beta} = \left(Z_{sel}^T \cdot Z_{sel}\right) \cdot Z_{sel}^T \cdot U \tag{6}$$

Based on eq. (6), four kinds of regression models were defined

1) Linear: $U = Y$, and $Zsel = Xsel$;
2) Log X: $U = Y$, and $Zsel = \log(Xsel)$;
3) Log Y: $U = \log(Y)$, and $Zsel = Xsel$;
4) *Log–Log*: $U = \log(Y)$, and $Zsel = \log(Xsel)$.

An analysis of variance was made for every regression model which included the degrees of freedom (k = number of predictor variables X_{SEL} or X_{ALL}), sum of squares of regression (*SSR*), sum of squares of errors (*SSE*), total sum of squares (*SSTO*), mean square of

regression (*MSR*), variance of model/mean square error (*MSE*), test statistic for ANOVA (F_{ratio}) and $f(0.05,k,n-k-1)$, the coefficient of determination R^2, the adjusted coefficient of determination R^2_{adj}, and the correlation *r*. The single best type of regression model was selected for each type of incoming hydrograph and for each dependent hydrograph and flooding variable, based on *MSE* and *r*. These regression models were chosen as the best model for each variable. A total of 5 best models were found for each dependent variable based on each type of hydrograph. The correlation (*r*) shows a linear relationship between original and modeled regression data, while the MSE allows for taking into account the bias in the regression model (Bennett et al., 2013).

An additional set of 10,000 simulations based on Gamma hydrographs (*G10*) was developed to test the regression models while changing the total number of simulations or sample points. The final selection of regression models for each variable was based on the validation results and *F* statistics obtained.

Subsequently these best models were tested with the validation sets to check their performance and capabilities.

The process was repeated also taking into account all the explanatory variables, to test the significance of the increase in fitness of the regression models.

4.6. Assumptions, uncertainties and limitations of the simulation model

The research presented herein proposes a method that entailed several assumptions. These assumptions, along with the uncertainties related to the model inputs, structure and observations may lead to limitations in the application of the proposed methodology.

The main assumptions of the methodology are: the initial set of explanatory variables must not be correlated, in order to guarantee that the surrogate MR is not built with cross-correlated variables; the surrogate has on average zero residuals, to guarantee that the model errors are random; sediment is not important in case of extreme flooding events; roughness is the same in the main channel and the floodplain; and the initial condition in the river is assumed to be the same in every simulation corresponding, i.e. a flow of 3,400 m^3/s.

A correlation analysis of the explanatory variables show that there is a moderate correlation between the initial and maximum volume of a reservoir, however this is a correlation imposed by the constraint that comes from the physical definition of a reservoir, i.e. the minimum volume of a reservoir cannot be larger than its maximum volume.

The simple routing of the flow, using the Muskingum–Cunge method, between the toe of the reservoirs and the Huayuankou station, instead of using a fully dynamic numerical solution of the Saint Venant equations, is an assumption that was used because of the mild slope of the reaches and taking into consideration the constant YRCC effort to keep up to date the Muskingum–Cunge parameters.

All assumptions along with the model structure lead to uncertainties. In the model structure one of the main sources of uncertainty lies with the flood model (2D), which was not calibrated using an extreme flood event, however this was not possible because no comparable records exist for the simulated events. A second source of uncertainty in the physically based flood model structure is the lack of knowledge of the variability of the main river reach. This could change the development of patterns of inundation, due to erosion and accretion of the banks.

From the perspective of the surrogate model the main source of uncertainty is the need to fix a predefined model structure (Specht, 1991), i.e. the linear relation between explanatory and dependant variables. However such relations may greatly overlook the existence of non-linear relationships (Poveda and Mesa, 1997; Wood, 1997; Harnold et al., 2001).

As a consequence of the assumptions and uncertainties one of the main limitation of the reservoir model is related to the use of triangular and Gamma hydrographs, which are having one singular peak along the whole simulation domain. If a double peak hydrograph is used as an input for each reservoir, a different filling pattern would occur. The filling of the reservoir with a first peak of the hydrograph would limit the amount of available volume to capture the second peak and consequently the release routed at Huayuankou would be higher than for the case of the single peak hydrographs. This will have an effect in the flooding, because the main channel downstream of Huayuankou will not be able to convey the same amount of water, spreading much faster the water in the lowland area. This limitation can be overcome by simulating the system with hydrographs from historical records, which are with different number of peaks.

The influence of the initial condition in the main reach of the Yellow river is also a limitation of the physically based model because it influences the amount of water that is flooding the banks inside the river dykes. The analysis can be extended to take into consideration variable initial river water levels, which needs to be used as an additional explanatory variable.

5. Results and discussion

The results have been split into two categories because of the differences found during the modeling results. The first category is the suggested model between the reservoir system and the resulting hydrographs at Huayuankou, and the second category is the model between the reservoir system and the flooding variables up to Gaocun.

5.1. Reservoir explanatory variables and hydrographs at Huayuankou

The ranges of the variables of the hydrographs at Huayuankou are presented in Fig. 4, A–E. Following the proposed model, a hypothesis on the correlation of explanatory variables that are more relevant for each hydrograph variable is obtained. This result is presented in Table 2 in columns 1 to 7. Each row represents the correlation of explanatory variables with respect to the hydrograph's variables at Huayuankou ($r_{P\text{-}H}$). The bold values represent significant *p*-values for each hydrograph variable.

For the peak discharge at Huayuankou (*HQp*), a total of 6 explanatory variables showed *p-value* significance. The peak discharge entering Sanmenxia shows the highest correlation (0.834). For *HQp* the time to the peak discharge at Sanmenxia (0.160), the maximum storage at Sanmenxia (0.079) and Luhun (−0.098) reservoirs and the peak discharge (0.116) and skewness of hydrograph (0.091) entering Guxian (0.018) have a significant correlation as well.

The time to peak discharge at Huayuankou (*HTp*) shows a significant correlation with 7 explanatory variables; the highest correlation is with the time to peak of the hydrograph entering Sanmenxia (0.901) and the maximum inverse correlation is to the peak discharge at the same reservoir (−0.234). In addition, this variable shows another interesting inverse correlation with the initial storage capacity at Xiaolangdi Reservoir.

The peak discharge and time to peak discharge at Huayuankou have simultaneous correlation with the peak discharge and time to peak discharge entering Sanmenxia. They also share a significant correlation with the time to peak of the discharge at Guxian and the maximum storage capacity at Sanmenxia. So the main driver of

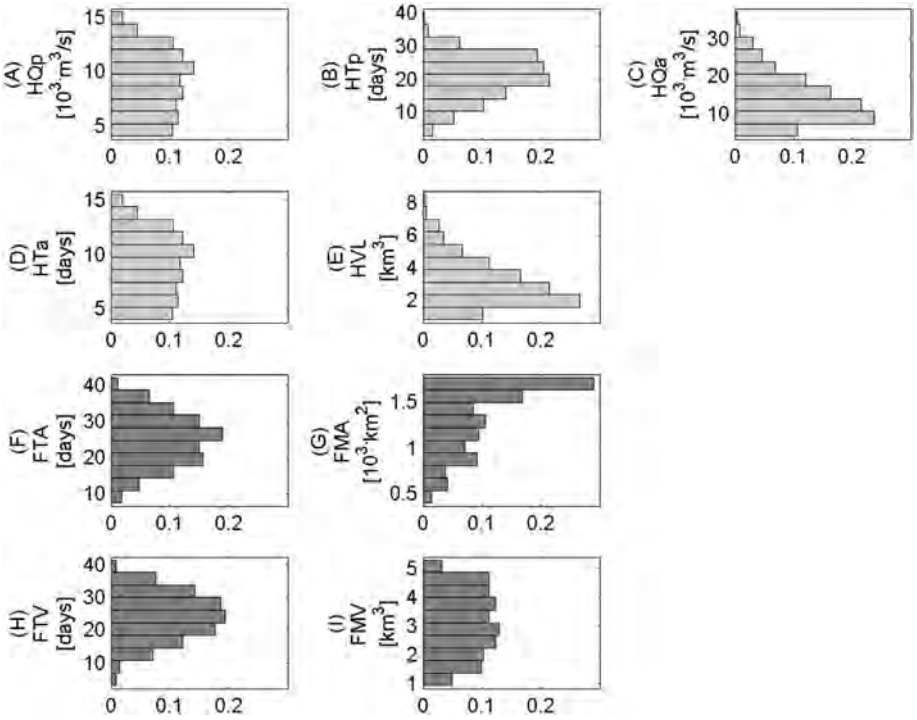

Fig. 4. Histogram of hydrographs at Huayuankou and flooding variables at Gaocun (Gamma input hydrographs).

these two variable areas is mostly related to the Sanmenxia and Guxian reservoirs.

For the average discharge at Huayuankou, 9 explanatory variables were found to be significant. The highest correlation found was for the time to peak discharge (0.563) and peak discharge (0.441) at Sanmenxia. This variable is mostly correlated to the incoming hydrograph at Sanmenxia.

For the average time of the hydrograph, 10 explanatory variables are significant, being the same two Sanmenxia hydrograph variables as for the other hydrograph variables, the initial and maximum storages at Sanmenxia and Xiaolangdi, the initial storage, time to peak and peak discharge at Guxian and the time to peak discharge at Luhun. In this case the maximum correlation found is again for the time to peak discharge at Sanmenxia (0.784), while the maximum inverse correlation found was with the initial storage capacity at Xiaolangdi (−0.232). This means that the faster the last reservoir is filled, the faster the flooding event will develop.

The volume of the hydrograph at Huayuankou was found to have a significant correlation to 11 of the variables, the four discharges variables at Sanmenxia, time to peak discharges and peak discharge incoming Luhun, three related with the incoming hydrograph at Guxian (except base flow) and the initial storage at Sanmenxia Reservoir. The maximum inverse correlation in this case occurs with

Table 2
Correlation of reservoir variables and descriptor variables ($r_{P\text{-}H}$) based on Gamma input hydrographs.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
		HQp	HTp	HQa	HTa	HVL	FTA	FMA	FTV	FMV
MAX%	G	0.028	−0.046	0.022	0.029	−0.015	0.023	−0.033	0.007	−0.033
	L	**−0.098**[a]	−0.007	**−0.080**	0.007	**−0.112**	−0.054	−0.021	−0.046	−0.056
	S	**0.079**	**0.130**	0.041	**0.140**	0.044	−0.092	−0.049	−0.102	−0.037
	X	0.035	**0.074**	−0.002	**0.075**	−0.004	0.012	−0.017	0.021	−0.031
INI%	G	0.037	−0.052	−0.041	**−0.088**	0.010	−0.079	0.032	−0.076	0.041
	L	−0.013	−0.011	−0.061	−0.056	−0.001	−0.022	−0.010	−0.021	−0.025
	S	0.048	−0.066	**0.072**	**−0.088**	**0.085**	0.020	−0.048	0.005	−0.030
	X	0.006	**−0.193**	0.040	**−0.232**	0.051	−0.017	0.004	−0.010	−0.021
G	Qb	0.030	−0.029	0.012	0.002	0.025	0.003	−0.066	−0.008	−0.076
	Qp	−0.002	−0.006	0.058	**−0.102**	**0.085**	−0.033	0.067	−0.018	0.052
	Tp	**0.116**	**0.136**	**0.233**	**0.219**	**0.186**	−0.013	0.057	−0.018	0.084
	Sk	**0.091**	−0.034	**0.086**	−0.027	**0.095**	−0.012	−0.043	−0.027	−0.056
L	Qb	0.038	0.000	0.027	0.047	0.060	0.053	−0.082	0.042	−0.070
	Qp	0.032	0.040	0.060	−0.049	**0.070**	−0.027	−0.004	−0.019	−0.014
	Tp	0.054	**0.073**	**0.154**	**0.141**	**0.115**	0.030	−0.031	0.022	−0.022
	Sk	−0.029	−0.008	−0.036	−0.004	−0.014	0.004	0.031	−0.016	0.042
S	Qb	0.026	−0.015	**0.090**	0.067	**0.091**	−0.049	−0.082	−0.057	−0.077
	Qp	**0.834**	**−0.234**	**0.459**	**−0.224**	**0.470**	0.027	−0.076	0.009	−0.054
	Tp	**0.160**	**0.901**	**0.567**	**0.784**	**0.566**	−0.058	0.051	−0.046	0.057
	Sk	−0.016	−0.059	**0.222**	0.038	**0.231**	−0.113	−0.082	−0.122	−0.075

[a] Bold letters represent correlation with *p*-value significance.

Table 3
ANOVA of multivariate regression analysis for the best models at Huayuankou.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
HydType	Selection	Variable	Model	Nparam	DoF	SSR	SSE	STO	MSR	MSE	Fratio	f(0.05,k,n-k)	R^2	R^2_{adj}	r
GAM	SEL	HQp	LOG_y	6	793	5037.0	1483.6	6520.6	839.5	1.871	448.7	2.110	0.773	0.771	0.880
GAM	SEL	HTp	LINEAR	7	792	28576.5	4018.3	32594.8	4082.4	5.074	804.6	2.021	0.877	0.876	0.936
GAM	SEL	HQa	LOG_y	9	790	1302.3	301.2	1603.5	144.7	0.381	379.5	1.892	0.812	0.810	0.903
GAM	SEL	HTa	LOG_y	10	789	11058.7	3287.8	14346.5	1105.9	4.167	265.4	1.843	0.771	0.768	0.878
GAM	SEL	HVL	LOG_y	11	788	27071.5	5109.3	32180.8	2461.0	6.484	379.6	1.801	0.841	0.839	0.918
GAM	ALL	HQp	LINEAR	20	779	5358.6	1162.0	6520.6	267.9	1.492	179.6	1.584	0.822	0.817	0.901
GAM	ALL	HTp	LINEAR	20	779	29021.9	3572.9	32594.8	1451.1	4.587	316.4	1.584	0.890	0.888	0.944
GAM	ALL	HQa	LOG_y	20	779	1443.0	160.5	1603.5	72.1	0.206	350.2	1.584	0.900	0.897	0.949
GAM	ALL	HTa	LOG_y	20	779	11360.9	2985.6	14346.5	568.0	3.833	148.2	1.584	0.792	0.787	0.890
GAM	ALL	HVL	LOG_y	20	779	29492.6	2688.2	32180.8	1474.6	3.451	427.3	1.584	0.917	0.914	0.958
G10	SEL	HQp	LOG–LOG	8	7991	49427.5	13724.7	63152.2	6178.4	1.718	3597.3	1.940	0.783	0.783	0.890
G10	SEL	HTp	LINEAR	9	7990	306400.4	37323.8	343724.2	34044.5	4.671	7288.0	1.881	0.891	0.891	0.944
G10	SEL	HQa	LOG_y	15	7984	12570.8	1869.1	14439.9	838.1	0.234	3579.8	1.668	0.871	0.870	0.934
G10	SEL	HTa	LOG_y	16	7983	109708.5	33317.9	143026.5	6856.8	4.174	1642.9	1.645	0.767	0.767	0.876
G10	SEL	HVL	LOG_y	16	7983	256345.5	33556.8	289902.2	16021.6	4.204	3811.5	1.645	0.884	0.884	0.941
G10	ALL	HQp	LINEAR	20	7979	51374.0	11778.2	63152.2	2568.7	1.476	1740.1	1.572	0.814	0.813	0.902
G10	ALL	HTp	LINEAR	20	7979	306507.0	37217.2	343724.2	15325.3	4.664	3285.6	1.572	0.892	0.892	0.944
G10	ALL	HQa	LOG_y	20	7979	12891.5	1548.4	14439.9	644.6	0.194	3321.6	1.572	0.893	0.893	0.946
G10	ALL	HTa	LOG_y	20	7979	110299.5	32727.0	143026.5	5515.0	4.102	1344.6	1.572	0.771	0.771	0.878
G10	ALL	HVL	LOG_y	20	7979	263132.8	26769.5	289902.2	13156.6	3.355	3921.5	1.572	0.908	0.907	0.954
TRI	SEL	HQp	LOG–LOG	4	795	4874.2	1674.2	6548.4	1218.6	2.106	578.6	2.383	0.744	0.743	0.869
TRI	SEL	HTp	LINEAR	6	793	29971.8	4721.5	34693.3	4995.3	5.954	839.0	2.110	0.864	0.863	0.930
TRI	SEL	HQa	LOG_y	6	793	993.6	508.1	1501.7	165.6	0.641	258.5	2.110	0.662	0.659	0.818
TRI	SEL	HTa	LOG_y	12	787	10502.8	3713.1	14215.9	875.2	4.718	185.5	1.764	0.739	0.735	0.860
TRI	SEL	HVL	LOG_y	6	793	19754.8	10076.8	29831.7	3292.5	12.707	259.1	2.110	0.662	0.660	0.819
TRI	ALL	HQp	LOG–LOG	17	782	5214.8	1333.6	6548.4	306.8	1.705	179.9	1.636	0.796	0.792	0.895
TRI	ALL	HTp	LINEAR	17	782	30297.8	4395.4	34693.3	1782.2	5.621	317.1	1.636	0.873	0.871	0.935
TRI	ALL	HQa	LOG_y	17	782	1126.5	375.2	1501.7	66.3	0.480	138.1	1.636	0.750	0.745	0.868
TRI	ALL	HTa	LOG_y	17	782	10719.9	3496.0	14215.9	630.6	4.471	141.1	1.636	0.754	0.749	0.869
TRI	ALL	HVL	LOG_y	17	782	22657.5	7174.2	29831.7	1332.8	9.174	145.3	1.636	0.760	0.754	0.873
UNI	SEL	HQp	LINEAR	2	797	4406.5	2141.4	6547.9	2203.3	2.687	820.0	3.007	0.673	0.672	0.820
UNI	SEL	HTp	LINEAR	5	794	6452.1	27863.4	34315.5	1290.4	35.092	36.8	2.225	0.188	0.183	0.434
UNI	SEL	HQa	LINEAR	4	795	345.3	1149.2	1494.5	86.3	1.446	59.7	2.383	0.231	0.227	0.481
UNI	SEL	HTa	LOG_y	8	791	3841.9	10419.3	14261.3	480.2	13.172	36.5	1.950	0.269	0.262	0.522
UNI	SEL	HVL	LINEAR	3	796	7260.7	22195.0	29455.8	2420.2	27.883	86.8	2.616	0.247	0.244	0.497
UNI	ALL	HQp	LINEAR	11	788	4587.8	1960.2	6547.9	417.1	2.488	167.7	1.801	0.701	0.697	0.837
UNI	ALL	HTp	LINEAR	11	788	6818.1	27497.4	34315.5	619.8	34.895	17.8	1.801	0.199	0.188	0.446
UNI	ALL	HQa	LINEAR	11	788	395.7	1098.8	1494.5	36.0	1.394	25.8	1.801	0.265	0.255	0.515
UNI	ALL	HTa	LOG_y	11	788	3924.3	10337.0	14261.3	356.8	13.118	27.2	1.801	0.275	0.265	0.528
UNI	ALL	HVL	LINEAR	11	788	8223.3	21232.5	29455.8	747.6	26.945	27.7	1.801	0.279	0.269	0.528

the maximum storage at Luhun Reservoir (−0.112). In general, the time to peak discharges and peak discharges of the three incoming hydrographs have relevance but adding the initial storage at Sanmenxia makes the operation of this reservoir the most relevant to the calculation of the volume of the resulting hydrographs.

5.1.1. Regression models up to Huayuankou

The multivariate models were built with this subset of selected explanatory variables. After that the models with all the explanatory variables were also built to compare the performance of the selection. An additional simulation of 10,000 hydrographs was generated for Gamma input hydrographs (*G10*) for a total of 160 models. The best-fitted models for each hydrograph variable based on input type hydrograph, were selected in each case. A total of 40 models are presented in Table 3 with the ANOVA table of regression. The column header of Table 3 is detailed in the acronym list of the paper.

By analyzing the results presented in Table 3 the following comparisons can be made:

1) Models with all the independent variables give better r and less *MSE* than the models with selected variables, but the reduction of the F ratios is not of significance, and nor is it near to the f statistic for any of the models. This result shows that there is no need to increase the number of explanatory variables from the ones selected by correlation, because in reality the combined effect of additional variables is not significant. Considering operational issues, less information is required for performing inference in the outcome hydrograph at Huayuankou while preserving the fitness of the results, which also saves time in decision-making. In most of the cases the variables related to the incoming hydrographs to the reservoirs in real-time operation are provided through forecasting modeling, so a link between these two could easily be achieved.
2) Models with Gamma input hydrographs show better fitted models than the models based on triangular and uniform hydrographs, independent of the number of explanatory variables *SEL* or *ALL*. However, in this case the models built with triangular hydrographs show comparable results to the Gamma type. This can be explained in principle for the effect of the skewness coefficient that seems to be not relevant for most of the resulting variables of the hydrographs. For uniform hydrographs the results are low compared either with Gamma and triangular input hydrographs, meaning that the importance of the peak discharge and the time to peak discharge as input to the reservoir system cannot be avoided. In an operational approach this means that a good forecast of the incoming flows at the reservoirs is required, because the average inflow does not provide enough information about the resulting hydrographs at Huayuankou.

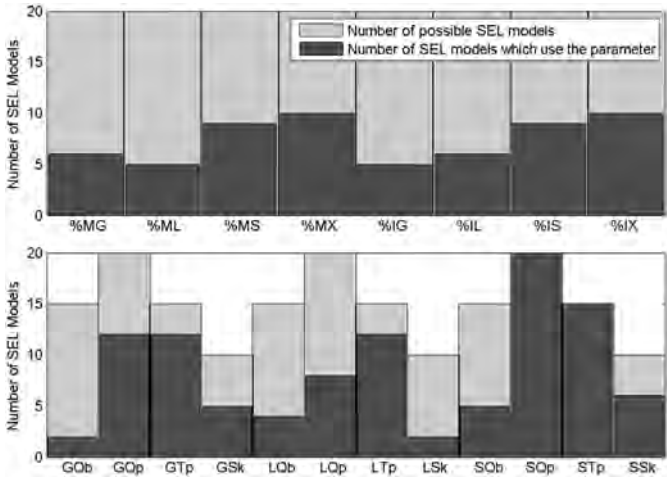

Fig. 5. Number of SEL models.

3) An increase in the number of simulations for the same hydrograph type preserving the actual model structure (*GAM* vs. *G10*) does not show an important increase in *r* or decrease in *MSE*. Even if the models show that the number of selected variables becomes dependent on the threshold value in the hypothesis test to correlation, for comparison purposes the selection made was based on the same threshold of *p*-values$<$0.05. The resulting models for the case where all variables are selected provide similar results, as shown by the R^2_{adj}, being the lowest differences between the ones for *HQp* and *HVL*. The best fitting model types are the same for the two different simulations, which shows that the structure of the regression is suitable for the same hydrograph type.
4) The most dominant regression model type for *HQa* is a *LogY*, showing that there is only the need for a logarithmic transformation, almost independent of the selection of variables or input hydrographs. A similar result can be found for *HTa*.
5) Taking into account the number of fitted models (*SEL*) in which every explanatory variable appears (Fig. 5), the most relevant explanatory variable is the peak discharge at Sanmenxia (20 models), followed by the time to peak discharge in the same location (15 models). Each selected model uses the two variables if the type of hydrograph requires it (the uniform type does not require times to peak). Third are the time to peak discharges at Guxian and Luhun, and the peak incoming discharge at Guxian (12 models). The least relevant explanatory variables by number of models selecting them are the base discharges of the incoming hydrographs at Guxian and the skewness of the hydrograph at Luhun (2 models).
6) The validation for input Gamma hydrographs and *SEL* variables (Fig. 6) shows that the models fit the data with error results that are allocated between the range of (+-) 20%. The lowest variability is presented by the *HTa* and *HTp*. The largest dispersion of error is presented for low values of *HQp*. Indeed, the models are intended to serve for peak values so this is not considered a deficiency of the resulting models, but rather as a consequence of the initial selection of simulations after routing which binds the minimum magnitude of the peak discharge.
7) After performing the multivariate regression of the explanatory variables, results were found to have similar fitness to the ones obtained using more sophisticated tools such as ANN or ANFIS that were presented in other studies (Peters et al., 2006; Ghalkhani et al., 2012; Li X. et al., 2010) in terms of hydrograph behavior.

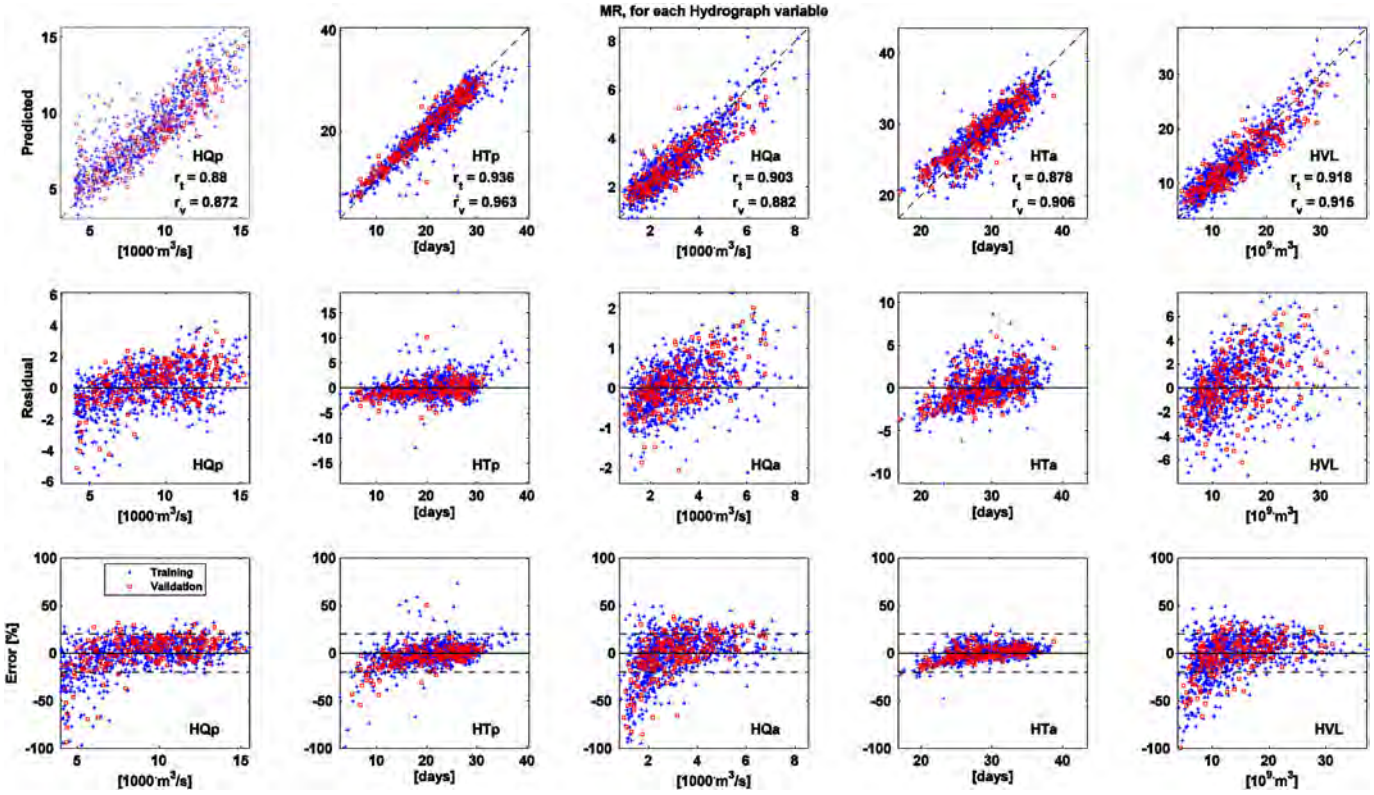

Fig. 6. Dispersion, residuals and percentage of error for Multivariate Regression models ($\boldsymbol{X_{SEL}}$).

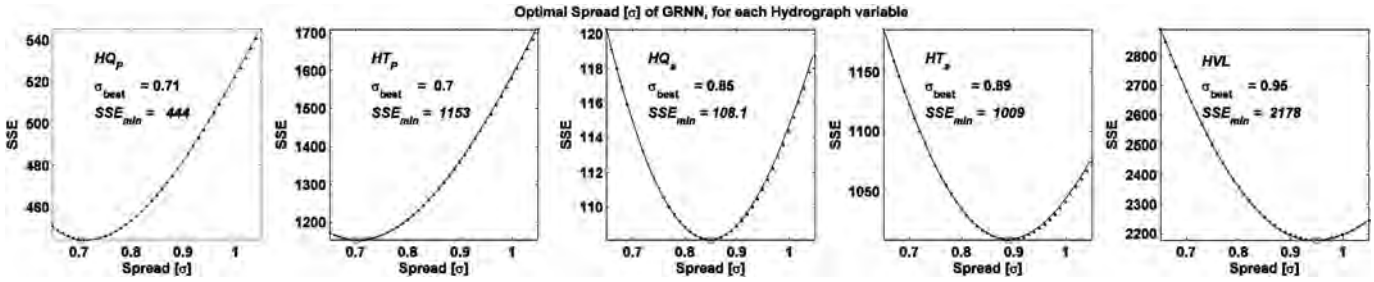

Fig. 7. Optimal spread parameter found for each GRNN model (X_{SEL}).

8) An additional comparison between the multivariate regression model and a generalized regression neural network (GRNN) as presented by Specht (1991) was performed. The advantage of GRNN is that an *a priori* specification of the regression model is not required. The input explanatory variables provided to the GRNN were the same as for the multivariate regression models (X_{sel}). The output variables were the ones obtained for Gamma hydrographs. A calibration of the spread parameter of the GRNN was performed by trial and error (see Fig. 7) Results of the best spread parameters and its corresponding SSE are presented for each hydrograph variable in Fig. 7. The considered spread parameter corresponds to the minimum point of the convex curve. Fig. 8, shows the obtained GRNN models for each hydrograph variable with optimal spread. The first row shows the scatter plots of training (800 simulations) and validation sets (200 simulations), the correlation is presented for each hydrograph variable. Table 4 shows that for every hydrograph variable the values of r_t are better than the ones obtained for the multivariate regression using SEL input variables, however when using the validation dataset the multivariate regression shows better performance for all the hydrograph variables. This result shows that overfitting appears due to the particularities of the analyzed case.
10) The second row of Fig. 8 shows the residuals with respect to the observed values. The residuals for the training set are smaller compared to the ones of the validation set. This shows a poor convergence of the GRNN, for this case study, in particular, even after optimizing its spread.
11) The third row of Fig. 8 shows that most of the errors for the training set appear to be in the range of ±20%, while there is a larger variation of the errors for the validation set.
12) A second GRNN model for each hydrograph variable was created using all the variables and optimizing the spread parameter. The correlation of the training dataset is 1.0 for every variable in the GRNN, however the correlation for the validation shows a reduced performance (see Table 4).

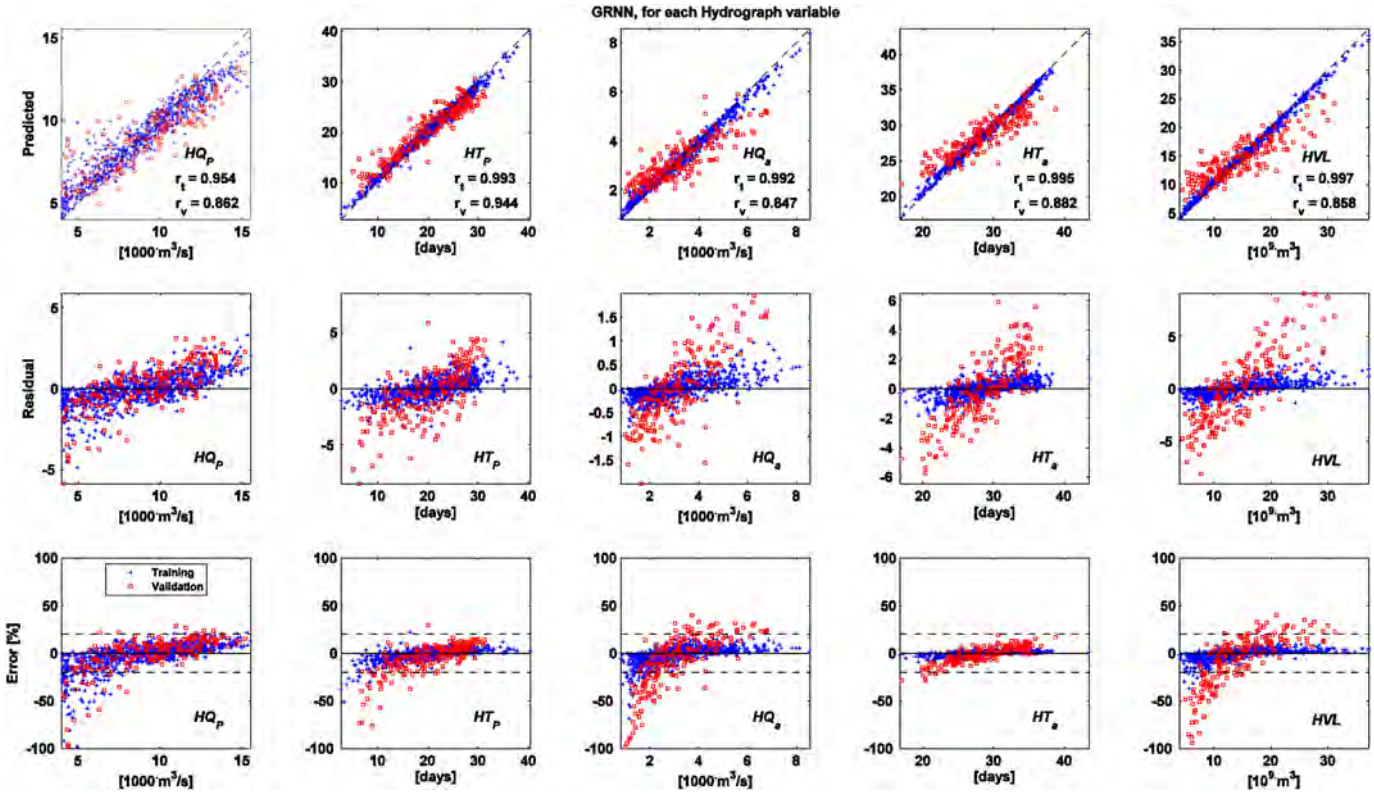

Fig. 8. Dispersion, residuals and percentage of error for GRNN models (X_{SEL}).

Table 4
GRNN and MR correlations for SEL and ALL variables using Gamma input hydrographs.

	Model	HQp	HTp	HQa	HTa	HVL
Training	GRNN X_{sel}	0.954	0.993	0.992	0.995	0.997
	GRNN X_{all}	1.000	1.000	1.000	1.000	1.000
	MR X_{sel}	0.880	0.936	0.903	0.878	0.918
	MR X_{all}	0.901	0.944	0.949	0.890	0.958
Validation	GRNN X_{sel}	0.862	0.944	0.847	0.882	0.858
	GRNN X_{all}	0.780	0.859	0.788	0.802	0.799
	MR X_{sel}	0.872	0.963	0.882	0.906	0.915
	MR X_{all}	0.896	0.968	0.946	0.909	0.953

13) Comparing the number of coefficients of regression, the reduction in complexity is an advantage of the multivariate regression models.
14) A comparison between the two GRNN models shows that GRNN with input selection performs poorly as compared to its corresponding GRNN with ALL variables for training. The advantage of the input selection appears when the correlation of the two GRNN's is compared in the validation, where the selection guarantees a better performance for every hydrograph variable (Table 4).
15) Present research shows that the model considered and the GRNN model performs in a similar way, however the later one requires a considerable effort for calibration in order to increase its performance, therefore the present type of models might be easier to use for particular cases, as it is the one considered here.

5.2. Flooding variables and regression models up to Gaocun

With the simulations that fall within 2% of the expected value for total volume of water at Huayuankou (*HVL*), a new set of regression models was developed up to Gaocun.

The ranges of the results are presented in Fig. 4 (F–I). The variables *FTA* and *FTV* show a less uniform behavior than the other two. One of the most interesting results is that the *FMA* has the highest frequency for the largest values; this shows the susceptibility of the Yellow River during flooding events, and that large proportions of the area inside the dykes are at risk. In contrast the *FMV* shows that frequencies of ~10% occur for a wide range of combinations of variables.

The correlation selection shows that there is no significance in the explanatory variables to the flooding variables (Table 2, columns 8–11). Given these circumstances, a selection of variables cannot be made for dependent flooding variables. This shows the complexity of identifying explanatory variables for the quantification of flooding events in a spatial approach. The only possible alternative was to construct the regression models with all the original explanatory variables as predictors of flooding variables, but even then the regression models would lack operational significance. Therefore more research should be carried out in the future for the Yellow River.

6. Conclusions

The main variables to take into account for the development of reservoir operation flooding in the mid Yellow River are related to the incoming hydrographs at the Sanmenxia Reservoir. This result is well known at YRCC due to the scale of the inflows reaching this reservoir as compared to the Guxian and Luhun reservoirs.

A comparison of the selection of relevant variables based on correlation as opposed to the use of all the explanatory variables showed that it is not necessary to calculate every possible regression model with every combination of selected explanatory variables to obtain very accurate results. In this case of 20 explanatory variables, the total number of regression models could have been ~10^6, when indeed only eight were analyzed in each case for each variable. In the real operative reservoir system of the YRCC, many more variables are taken into account, hence based on the obtained results the methodology can be extended with more data.

The regression models between the independent reservoir variables and the dependent variables of the hydrographs at Huayuankou show that it is possible to use the proposed approach in a very complex system like the one studied.

The best regression models for hydrograph variables were the ones for the total volume while the worst performance models were for the peak discharge at Huayuankou. This shows the difficulty in the inference and forecasting of this variable. These models can be used in the future research to develop confidence intervals of hydrograph responses.

In the case of the regression models between the input variables and the flooding variables, it was not possible to find a satisfactory structure. This shows once again the complexities of the Yellow River and the amount of uncertainty in the understanding of the processes during flooding events. In particular case of the Yellow River the poor performance of the regression model for the flooding variables is due to the fact that flooding takes place only in the space between the dykes. Consequently two sub-processes developed in the 2D flood model. First the water floods out of the hanging main channel and splits in branches, concentrating near the winter dykes. Secondly is the reflection of the water from the dykes towards the main channel, as well as the spreading along the dykes. The combination of the two processes makes it almost impossible to perform inference of flooding variables because there is not enough space for the flood to propagate; it is confined within the limit of the dykes.

Future research of the inference model on the Yellow River should address the fact that there is a positive correlation between the maximum storage capacity at Sanmenxia and the peak discharge at Huayuankou. The expectation is that these two variables should be negatively correlated. Furthermore the research could be extended to application of computational intelligence techniques for regression of flooding variables, such as artificial neural networks, radial basis, support vector machines and fuzzy based rules.

It is recognized that there are three characteristics of the flooding model that need to be improved in future developments. The first one is the need for integration with a hydrological model in the area because the internal area of the flooding model is almost 2,000 km^2, and no additional volume was added to the simulations for a period of 100 days. Second, the model is not yet able to take into account the infiltration process, so basically all of the water that enters as a hydrograph either goes out or continues to represent the flooding area. The third is that it is almost impossible to obtain smooth input hydrographs like the ones created in the generation with the three types of hydrographs (i.e. Gamma, Triangular and Uniform), but also the regression models, at least up to Huyankou, were able to show the relevance of the time to peak of the hydrographs at Sanmenxia.

Current research contributes to the understanding of complex river systems like the Yellow River and can be extended to other major rivers in the world.

Acknowledgments

The authors are grateful to Deltares, for providing the funding to carry out the research. Data was kindly provided by Shengyang Li who is a member of the YRCC.

References

Audsley, E., Pearn, K.R., Harrison, P.A., Berry, P.M., 2008. The impact of future socioeconomic and climate changes on agricultural land use and the wider

environment in East Anglia and North West England using a metamodel system. Clim. Change 90 (1–2), 57–88.
Balica, S.F., Popescu, I., Beevers, L., Wright, N.G., 2013. Parametric and physically based modelling techniques for flood risk and vulnerability assessment: a comparison. Environ. Model. Softw. 41 (3), 84–92.
Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environ. Model. Softw. 40 (2), 1–20.
Bowden, G.J., Dandy, G.C., Maier, H.R., 2005. Input determination for neural network models in water resources applications. Part 1-Background and methodology. J. Hydrol. 301, 75–92.
Broad, D.R., Dandy, G.C., Maier, H.R., 2005. Water distribution system optimization using metamodels. J. Water Resour. Plan. Manag. 131 (3), 172–180.
Castelletti, A., de Rigo, D., Rizzoli, A., Soncini-Sessa, R., Weber, E., 2007. Neuro-dynamic programming for designing water reservoir network management policies. Control Eng. Pract. 15 (8), 1001–1011.
Castelleti, A., Galelli, S., Ratto, M., Soncini-Sessa, F., Young, P.C., 2012. A general framework for dynamic emulation modelling in environmental problems. Environ. Model. Softw. 34, 5–12.
Castro Gama, M., Popescu, I., Shengyang, L., Mynett, A., 2013. Modeling the inference between upstream inflow hydrographs and downstream flooded areas in a reservoir driven system. In: Proceedings of AiroWinter Conference, Operational Research for Development, Sustainability and Local Economies, Champoluc, Italy, January 2013.
Celeste, A., Billib, M., 2009. Evaluation of stochastic reservoir operation optimization models. Adv. Water Resour. 32, 1429–1443.
Chow, V., Maidment, D., Mays, L., 1988. Applied Hydrology. Editorial McGraw-Hill, New York, p. 712.
CIAT, 2004. Void-filled Seamless SRTM Data V1. International Centre for Tropical Agriculture. Consortium for Geospatial Information. King College of London as seen on. http://srtm.csi.cgiar.org, and http://www.ambiotek.com/topoview. Last time visited 2011/10/03.
Cunge, J.A., 1969. On the subject of a flood propagation computation method (Muskingum method). J. Hydrol. Res. Delft, Neth. 7 (2), 205–230.
Dinh, N.Q., Balica, S., Popescu, I., Jonoski, A., 2012. Climate change impact on flood hazard, vulnerability and risk of the Long Xuyen Quadrangle in the Mekong Delta. J. River Basin Manag. 10 (1), 103–120.
Dittmann, R., Froehlich, F., Pohl, R., Ostrowski, M., 2009. Optimum multi-objective reservoir operation with emphasis on flood control and ecology. Nat. Haz. Earth Syst. Sci. 9, 1973–1980.
Dong, W., 2007. Evaluation of the Flow Regime in the Lower Yellow River Delta. China (M.S. thesis UNESCO-IHE). WSE-HI, Delft, The Netherlands, 2007–09.
Gano, S.E., Kim, H., Brown, D.E., 2006. Comparison of three surrogate modeling techniques: datascaper, kriging, and second order regression. In: Proceedings of the 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference. 6–8 September 2006. Portsmouth, VA.
Ghalkhani, H., Golian, S., Saghafian, B., Farokhnia, A., Shamseldin, A., 2012. Application of surrogate artificial intelligent models for real-time flood routing. Wat. Environ. J. http://dx.doi.org/10.1111/j.1747–6593.2012.00344.x.
Gichamo, Z.G., Popescu, I., Jonoski, A., Solomatine, D.P., 2012. River cross section extraction from ASTER global DEM for flood modeling. Environ. Model. Softw. 31 (5), 37–46.
Gilmore, A., 1999. A Study of Optimization of Reservoir Operations of the Colorado River (M.S. thesis). Science Department of Civil, Environmental, and Architectural Engineering, University of Colorado.
Guoying, L., 2010. Ponderation and Practice of the Yellow River Control System, vol. 1. Yellow River Conservancy Press. ISBN 7-80621-727-4/TV331.
Harnold, T.I., Sharma, A., Sheather, S., 2001. Selection of kernel bandwidth for measuring dependence in hydrologic time series using the mutual information criteria. Stoch. Environ. Res. Risk Assess. 15, 310–324.
Hartanto, I.M., Beevers, L., Popescu, I., Wright, N.G., 2011. Application of a coastal modelling code in fluvial environments. Environ. Model. Softw. 26 (12), 1685–1695.
Hassaballah, K., Jonoski, A., Popescu, I., Solomatine, D.P., 2012. Model-based optimisation of downstream impact during filling of a new reservoir: case study of Mandaya/Roseires reservoirs on the Blue Nile River. Water Resour. Manag. 26 (2), 273–293.
He, H., Yu, Q., Zhou, J., Tian, Y.Q., Chen, R.F., 2008. Modelling complex flood flow evolution in the middle Yellow river basin, China. J. Hydrol. 353, 79–92.
Hejazi, M.I., Cai, X., 2009. Input variable selection for water resources systems using a modified minimum redundancy maximum relevance (mMRMR) algorithm. Adv. Water Resour. 32, 582–593.
Hunter, N.M., , et al.Bates, P.D., Horritt, M.S., De Roo, P.J., Werner, M.G.F., 2005. Utility of different data types for calibrating flood inundation models within a GLUE framework. Hydrol. Earth Syst. Sci. 9 (4), 412–430.
Hussain, M.F., Barton, R.R., Joshi, S.B., 2002. Metamodelling: radial basis functions, versus polynomials. Eur. J. Oper. Res. 138 (1), 142–154.
Jonoski, A., Popescu, I., 2011. Distance learning in support of water resources management: an online course on decision support systems in river basin management. Water Resour. Manag. 26 (5), 1287–1305.
Jung, Y., Merwade, V., 2012. Uncertainty quantification in flood inundation mapping using generalized likelihood uncertainty estimate and sensitivity analysis. J. Hydrol. Eng. 17 (4), 507–520.
Kernkamp, H., Dam, A., Stelling, G., Goede, E., 2011. Efficient scheme for the shallow water equations on unstructured grids with the application to the continental shelf. Ocean. Dyn. 61 (8), 1175–1188.
Kramer, S., Stelling, G., 2008. A conservative unstructured scheme for rapidly varied flows. Int. J. Numer. Meth. Fluids 58 (2), 183–212.
Labadie, J., 2004. Optimal operation of multireservoir systems: state of the art review. J. Water Resour. Plann. Manage 130 (2), 93–111.
LeFavour, G., Alsdorf, D., 2005. Water slope and discharge in the Amazon River estimated using the shuttle radar topography mission digital elevation model. Geophys. Res. Lett. 32, L17404. http://dx.doi.org/10.1029/2005GL023836.
Li, X., Guo, S., Liu, P., Chen, G., 2010. Dynamic control of flood limited water level for reservoir operation by considering inflow uncertainty. J. Hydrol. 391 (1–2), 124–132.
Li, S., 2008. Software System Requirements for Multi-reservoir-based Flood Control and Management. A Case Study for the Yellow River in China (M.Sc thesis). UNESCO IHE, Delft, The Netherlands.
Li, S., 2013. Adaptive Multi-reservoir-based Flood Control and Management for the Yellow River, towards a Next Generation Software System (PhD thesis). WSE, TU, Delft, The Netherlands.
Liu, G.W., 2012. On the geo-basis of river regulation in the lower reaches of the Yellow. River. Sci. China Earth 55 (4), 530–544.
Loucks, D.P., van Beek, E., 2005. Water Resources Systems Planning and Management. An Introduction to Methods, Models and Applications. Studies and report in Hydrology. UNESCO publishing. ISBN 92-3-103998-9.
Lund, J., Guzman, J., 1999. Some derived operating rules for reservoirs in series or in parallel. J. Water Resour. Plann. Manage 125 (3), 143–153.
Merwade, V., Olivera, F., Arabi, M., Edleman, S., 2008. Uncertainty in flood inundation mapping: current issues and future directions. J. Hydrol. Eng. 13 (7), 608–620.
Milly, P.C.D., Wetherald, R.T., Dunne, K.A., Delworth, T.L., 2002. Increasing risk of great floods in a changing climate. Nature 415, 514–517. http://dx.doi.org/10.1038/415514a.
Moya Quiroga, V., Popescu, I., Solomatine, D., Bociort, L., 2013. Cloud and cluster computing in uncertainty analysis of integrated flood models. J. Hydroinform. 15 (1), 55–70.
Muste, M., Quinn, P.F., Hewett, C.J.M., Popescu, I., Basu, N.B., Kumar, P., Franz, K., Merwade, V., Arnold, W., Potter, K., 2010. Initiation of the upper Mississippi river Basin observatory. In: ASCE Proceedings, Innovation in Watershed Management under Land Use and Climate Change 39, pp. 1270–1281.
NERC, 1975. London. Natural Environmental Research Council. Flood Studies Report, vol. 1–5. ISBN: 978 190 6 698003.
Peters, R., Schmitz, G., Cullmann, J., 2006. Flood routing modelling with artificial neural networks. Adv. Geosci. 9, 131–136.
Pineros-Garcet, J.D., Ordonez, A., Roosen, J., Vanclooster, M., 2006. Metamodelling: theory, concepts and application to nitrate leaching modelling. Ecol. Model. 193 (3–4), 629–644.
Popescu, I., Jonoski, A., van Andel, S.J., Onyari, E., Moya Quiroga, V.G., 2010. Integrated modelling for flood risk mitigation in Romania: case study of the Timis-Bega river basin. Int. J. River Basin Manag. 8 (3–4), 269–280.
Popescu, I., Jonoski, A., Bociort, A., 2012. Decision support systems for flood management in the Timis Bega catchment. Environ. Eng. Manag. J. 11 (12), 2305–2311.
Poveda, G., Mesa, O.J., 1997. Feedbacks between hydrological processes in tropical south america and large-scale ocean–atmospheric phenomena. J. Clim. 10, 2690–2702.
Quinn, P., Hewet, C., Popescu, I., Muste, M., 2010. Towards new types of water-centric Collaboration: Instigating the upper Mississippi river Basin observatory process. Water Manag. 163, 39–51.
Rani, D., Moreira, M.M., 2010. Simulation-Optimization modeling: a survey and potential application in reservoir systems optimization. Wat. Resour. Manage 24, 1107–1138.
Ratto, M., Pagano, A., Young, P., 2007. State dependent parameter metamodelling and sensitivity analysis. Comp. Phys. Commun. 177 (11), 863–876.
Ratto, M., Pagano, A., 2010. Using recursive algorithms for the efficient identification of smoothing spline anova models. AStA Adv. Stat. Anal. 94 (4), 367–388.
Ravazi, S., Tolson, B.A., Burn, D.H., 2012. Review of surrogate modeling in water resources. Water Resour. Res. 48, W07401. http://dx.doi.org/10.1029/2011WR011527.
Saavedra, V.O.C., Koike, T., Yang, K., Graf, T., Li, X., Wang, L., Han, X., 2010. Decision support for dam release during floods using a distributed biosphere hydrological model driven by quantitative precipitation forecasts. Water Resour. Res. 46 (10), 1–13.
Shan, S., Wang, G.G., 2010a. Survey of modeling and optimization strategies to solve high dimensional problems with computationally expensive black-box functions. Struct. Multidisc. Optim. 41, 219–241.
Shan, S., Wang, G.G., 2010b. Metamodeling for high dimensional simulation based design problems. J. Mech. Des. 132 (5), 1–11.
Schumann, G., Bates, P.D., Horritt, M.S., Matgen, P., Pappenberger, F., 2009. Progress in integration of remote sensing–derived flood extent and stage data and hydraulic models. Rev. Geophys. 47 (4), 1–20.
Schmitz, G.H., Cullman, J., 2008. PAI-OFF: a new proposal for online flood forecasting in flash flood prone catchments. J. Hydrol. 360, 1–14.
Specht, D.F., 1991. A general regression neural network. IEEE, Trans. Neural Networks 2 (6), 568–576.
Stocker, T.F., Raible, C.C., 2005. Water cycle shifts gear. Nature 434, 830–833.
Todini, E., 2007. A mass conservative and water storage consistent variable parameter Muskingum-Cunge approach. Hydrol. Earth. Sci. Discuss. 11 (4), 1549–1592.

Van, P.D.T., Popescu, I., van Griensven, A., Solomatine, D., Trung, N.H., Green, A., 2012. A study of the climate change impacts on fluvial flood propagation in the Vietnamese Mekong Delta. Hydrol. Earth Syst. Sci. 16 (12), 4637–4649.

Wan, Z.H., Wang, Z.Y., 1994. Hyperconcentrated Flow. In: IAHR Monogr. Ser. A.A. Balkema, Brookfield, Vt.

Wang, G., Wu, B., Ya, Z., 2005. Sedimentation porblems and management strategies of Sanmenxia reservoir, Yellow river, China. Water Resour. Res. 41, W09417. http://dx.doi.org/10.1029/2004WR003919.

WMO, 2008. Reservoir Operation and Managed Flows, a Tool Integrated Flood Management. In: Flood Management Tools Series. World Meteorological Organization. APFM Technical Document No. 10.

Wood, E.F., 1997. Effects of soil moisture aggregation on surface evaporative fluxes. J. Hydrol. 190, 397–412.

Wu, B.S., 1992. A study on real-time optimal operation of reservoirs for flood control for the lower Yellow River (in chinese). J. Wuhan Univ. Hydraulic Electr. Eng. 25, 97–101.

Wurbs, R., 1993. Reservoir-system simulation and optimization models. J. Water. Resour. Plann. Manage. ASCE 119, 455–472.

Yeh, W.-G., 1985. Reservoir management and operation models: a-state-of-the-art Review. Water Resour. Res. 21 (12), 1797–1818.

Young, P.C., Ratto, M., 2009. A unified approach to environmental systems modeling. Stoch. Environ. Red. Risk Assess. 23, 1037–1057.

Hydrol. Earth Syst. Sci., 18, 1225–1237, 2014
www.hydrol-earth-syst-sci.net/18/1225/2014/
doi:10.5194/hess-18-1225-2014

Hydrology and Earth System Sciences
Open Access

Challenges in modelling river flow and ice regime on the Ningxia–Inner Mongolia reach of the Yellow River, China

C. Fu[1], I. Popescu[2], C. Wang[2,4], A. E. Mynett[2,3], and F. Zhang[5]

[1]NARI Group Corporation, water conservancy and hydropower technology branch company, Nanjing 211000, China
[2]UNESCO-IHE, Institute for Water Education, 2601DA, Delft, the Netherlands
[3]Delft University of Technology, Faculty of CiTG, 2600GA, Delft, the Netherlands
[4]Hydrology Bureau, Yellow River Conservancy Commission, Zhengzhou 450004, China
[5]Yellow River Institute of Hydraulic Research, Zhengzhou 450003, China

Correspondence to: I. Popescu (i.popescu@unesco-ihe.org)

Received: 22 July 2013 – Published in Hydrol. Earth Syst. Sci. Discuss.: 11 October 2013
Revised: 31 January 2014 – Accepted: 18 February 2014 – Published: 28 March 2014

Abstract. During winter the Yellow River in China is frequently subjected to ice flood disasters. Possible dike breaking due to ice floods poses a serious threat to the part of the region located along the river, in particular the Ning–Meng reach (including Ningxia Hui and the Inner Mongolia autonomous regions). Due to its special geographical location and river flow direction, the ice dams and jams lead to dike breaking and overtopping on the embankment, which has resulted in huge casualties and property losses throughout history. Therefore, there is a growing need to develop capability in forecasting and analysing river ice floods. Research into ice floods along the river is taking place at the Yellow River Conservancy Commission (YRCC). A numerical model is one of the essential parts of the current research going on at the YRCC, which can be used to supplement the inadequacies in the field and lab studies which are being carried out to help understand the physical processes of river ice on the Yellow River. Based on the available data about the Ning–Meng reach of the Yellow River, the YRCC river ice dynamic model (YRIDM) has been tested for capabilities to conduct ice flood forecasting. The YRIDM can be applied to simulate water level, discharge, water temperature, and ice cover thickness under unsteady-state conditions. Different scenarios were designed to explore the model uncertainty for two bounds (5 and 95 %) and probability distribution. The YRIDM is an unsteady-state flow model that can show the basic regular pattern of ice floods; hence it can be used as an important tool to support decision making. The recommendation is that data and research should be continued in order to support the model and to measure improvements.

1 Introduction

River ice is a natural phenomenon that can be commonly seen in the cold regions of the world. River ice plays an important role in cold regions, being the main transportation means in wintertime: the ice roads and ice bridges caused by river ice in the northern regions of Canada, Russia, and the USA (Alaska), where the population is sparse, are a positive effect of the ice formation in cold regions (Petrow et al., 2007; Rojas et al., 2011). However, on the negative side, river ice can cause ice flooding, hamper hydropower generation, threaten hydraulic structures, hinder water supply and river navigation, and has many other important detrimental effects.

When it comes to floods caused by river ice, the Ning–Meng reach of the Yellow River, China, cannot be ignored. The ice flood of the Yellow River basin is one of the most dangerous situations of all the Chinese rivers. According to historical data, different ice flood events have been recorded in the past during the winter on the main stream and the tributaries of the Yellow River. These floods have caused huge casualties and property losses, especially on the Ning–Meng reach of the Yellow River. Since 1949 there have been 29 ice flood hazards, nine of them causing the levee to breach in the Ning–Meng reach.

Ice regime forecasting is an efficient method for ice flood control, regulation and decision making, and mathematical modelling is playing an essential part in the development of river ice research. Mathematical models can supplement the inadequacies of field and lab studies to help understand

the physical processes of river ice, and at the same time, they can also be a tool to help design and plan engineering projects. They have become valuable tools for exploring the research area of river ice, such as understanding the physical processes and simulating river ice phenomena, and even forecasting ice floods (Dahkle et al., 2012).

The development of the economy and society with the resulting climate change and human activities have changed the characteristics of ice regimes, especially ice disasters during the freezing or breaking-up periods, attracting more and more attention from water authorities and local governments. Hence, it is very important to know ice regime characteristics and use mathematical models to enable ice forecasting, ice flood prevention and ice flood alleviation.

At the YRCC a 1-D numerical model, named YRIDM has been developed to simulate ice regimes. It also contains data on observed data including hydrological, meteorological, and ice regime data, covering about fifty years of measurements. The available model and collected field data provide an opportunity to examine the possibility of ice flood forecasting. The choice of the 1-D model is given by the availability of data at YRCC. Currently there is no DEM available for the Ningxia–Inner Mongolia reach and the available collected data allowed for the development of a one-dimensional model. As the data collection continues a more complex model will be made available, by extending the existing model to a one-dimensional–two-dimensional (1-D2-D) one. It is important to use such a model in forecast conditions, in order to determine what are the possible times to have ice-formation and consequently flooding.

The purpose of this paper is to apply the YRIDM to the Ning–Meng reach of the Yellow River to examine the accuracy of the model.

This paper first describes the current state of the art in ice flood modelling, followed by a description of the case study. The modelling application is detailed next. Modelling of ice floods on the Ning–Meng reach is discussed with respect to model set-up, calibration, verification and uncertainty analysis. This paper ends with a conclusion section.

2 Ice flood modelling state of the art

2.1 Ice flood research

Ice floods in rivers are generated due to ice jams, ice dams, ice melt and snow melt (Liu et al., 2000). The ice flood process cannot be separated from the whole river ice process, although it mainly involves the ice-covered period and breaking-up period.

As summarized by Shen (2006), during the past fifty years engineering and environmental issues have largely driven the development of river ice research, and significant achievement has been made during this period. However, Beltaos (2008) thought that although there has been remarkable progress in understanding and quantifying the complex river ice processes, many problems concerning river ice still remain today or are only partially solved.

Shen (2006) classifies the river ice research area into two parts: firstly, the energy budget methods and water temperature distribution calculations before and during the freezing-up period and secondly, the evolution of frazil ice, frazil floc, anchor ice and ice dams. Currently the water temperature distribution is well understood (e.g. Shen and Chiang, 1984), and the mechanisms of super cooling and frazil ice formation are also relatively well understood (Osterkemp, 1978; Daly, 1984). However, the evolution of frazil ice, frazil floc, anchor ice and ice dams needs to be studied further (Ye et al., 2004), the transitional conditions among different ice run regimes are not clearly understood (Hammar et al., 2002), and also knowledge on the mechanism of ice pans and ice floe formation is limited. In addition, a complete analytical formulation of the mechanical break-up should be developed.

Beltaos (2008) showed the challenges and opportunities for research into river ice processes. The main challenge is to avoid or decrease the negative influence of river ice processes and to make sure that the positive influence is not affected by human activities (Beltaos, 2008). In order to meet this challenge, it is necessary that there is a good qualitative understanding of river ice processes. However, there still remain serious gaps, such as research into anchor ice, break-up, ice jamming, and the influence of climate change on the river-ice process. The main opportunities lie in how to use new technologies to understand the river ice processes. These new technologies include instrumentation, numerical modelling, mitigation and prediction of climate impact on river ice processes.

Research into ice cover has developed from static ice cover research, such as Beltaos and Wong (1986), into dynamic ice cover research due to the fact that the original research did not take ice dynamics into consideration. Dynamic ice cover research has been further divided into one-dimensional dynamic ice cover research (Shen et al., 1990) and two-dimensional dynamic ice cover research, such as the DynaRICE model (Shen et al., 2006), in order to take into account the frictional resistance of riverbanks and channel bottoms, irregular cross sections of river channels, and unsteady flow state in reality.

Research into ice transportation under the ice covers and ice jams started from a critical velocity criterion or Froude number criterion (Kivislid, 1959; Tesaker, 1975). However, this method could not provide a way to compute an accurate value of the critical velocity and Froude number. Hence, Shen and Wang (1995) developed a concept of ice transport capacity to simulate ice deposition under the ice cover, which was demonstrated by field observation data. Nevertheless Beltaos (2008) still thought that ice transportation under the ice cover was only partially understood, and further research should be continued in the future.

The process of the thermal growth and erosion of ice cover is relatively clearly understood. Firstly, research focused on the ice cover without layers of snow ice, snow slush or black ice (Shen and Chiang, 1984). Secondly, the layers of snow ice, snow slush and black ice were taken into consideration (Calkins, 1979; Shen and Lal, 1986). Finally, based on the previous research, a lot of models were produced to simulate the processes. The most classical one was the degree-day method proposed by Stefan (1889) that has been used to simulate ice cover growth for a long time. Shen and Yapa (1985) refined the classical degree-day method which they then developed into a modified degree-day method, which was demonstrated on the St. Lawrence River.

The break-up of ice cover can be classified as mechanical break-up or thermal break-up. A mechanical break-up can exert a negative influence on the hydraulic facilities and the safety of people along the river. However, the ability to simulate a mechanical break-up is still limited (Shen, 2006).

Only a few researchers have tried to simulate the propagation of ice-jam release waves. Most of them have used a one-dimensional model without considering the effect of ice (e.g. Blackburn and Hicks, 2003). Field investigations were conducted by Jasek (2003), who found out that release wave celerity seemed to change with different ice conditions. Based on this research output, Liu and Shen (2004) started to take into consideration the effect of ice on wave propagation. She and Hicks (2006) built a model, named River-1D, to look into the effect of ice on ice-jam release waves, which was tested by the release event on the Saint John River in 1993, on the Saint John River in 2002, and on the Athabasca River in 2002. At the same time, statistical methods and ANN are also used to forecast the break-up of river ice. The potential for using fuzzy expert systems to forecast the potential risk of ice jams was discussed (Mahabir et al., 2002), and the systems identified five years when high water levels could occur, including four years when the predicted high water level did occur. Based on this potential, more research has been carried out (Mahabir et al., 2006; Wang et al., 2008) which took more related parameters into consideration.

2.2 Ice flood modelling

Among all the river ice research, mathematical modelling is an essential part of the progress. Ice flooding models have developed from 1-D steady state to 1-D unsteady state and then into 2-D models, and data-driven models are also being applied to forecast ice floods.

The 1-D steady-state models are based on the static ice-jam theory, namely that the formulas to determine the final thickness of water surface ice can be deduced according to the static balance of internal and external forces on a floating ice block. The models based on the static ice-jam theory are HEC-2 (US Army, 1990), ICETHK (Tuthill et al., 1998), and HEC-RAS (US Army, 1998). The basic assumption is that the flow is steady, gradually varied, and one-dimensional, and that river channels have small slopes (less than 1 : 10). The basic equation is the force and energy balance equations; a standard step method is used to solve the equation. However, the limitation of 1-D steady-state models is that they ignore ice dynamic conditions.

Due to the limitation of 1-D steady-state models, the development of 1-D unsteady-state models, such as RICE (Lal and Shen, 1991), RICEN (Shen et al., 1995), and the Comprehensive River Ice Simulation System (CRISSP1D) model (Chen et al., 2006), was promoted. These models are based on the assumption of unsteady-state flow; the governing equations are 1-D Saint Venant equations (i.e. mass and momentum conservation equations with floating ice) and they can be used to simulate the entire ice process in rivers during the winter season. The basic assumptions of the RICE, RICEN, and CRISSP1D models are that they ignore the following: (1) the effect on the water body mass balance due to the change in ice phase; (2) the river has a floating ice cover; (3) two-layer ice transportation theory; (4) suspension ice is full of the suspension ice layer; and (5) the thickness of the ice layer on the water surface is equal to the thickness of the ice block floating on the water surface. However, the limitation of these models is that they lack detailed consideration of complex flow patterns and river geometry (Shen, 2010).

The RICE and RICEN models have been used to simulate the ice regime for a few rivers; for example, the RICE model was used for the Upper St. Lawrence River near New York and the Peace River in northern British Columbia (Li et al., 2002) and the RICEN model was used for ice-jam simulation in 1995 on the Niagara River.

Based on the limitations of 1-D unsteady-state flow, 2-D models have also been developed, such as DynaRICE (Shen et al., 2000) and CRISSP2D (Liu and Shen, 2006) that can be used to simulate the ice regime. These models solve the 2-D depth-integrated hydrodynamic equations for shallow water flow. The two basic assumptions are that the movement of the surface ice layer is continuous, and ice is a kind of continuous medium. A finite element method is used to solve the equations. However, the limitation is that they lack detailed consideration of the third dimension.

DynaRICE has been applied to understand the ice-jam evolution, ice boom design and navigation structure design, on several rivers. DynaRICE was applied in the Niagara Power Project to study both the ice control and ice-period operation, on the Missouri–Mississippi River to study ice-jam formation, and on the Shokotsu River in Hokkaido, Japan, to study break-up jams. CRISSP2D has also been applied to several rivers and lakes to simulate the ice regime. It has been applied to the Nelson River to simulate freeze-up ice conditions (Malenchak et al., 2008); on the Red River near Netley Cut (Haresign and Clark, 2011) to model ice formation; and to test the potential for anchor ice growth after the Conawapa Power Generation Station was constructed and operational (Morris et al., 2008).

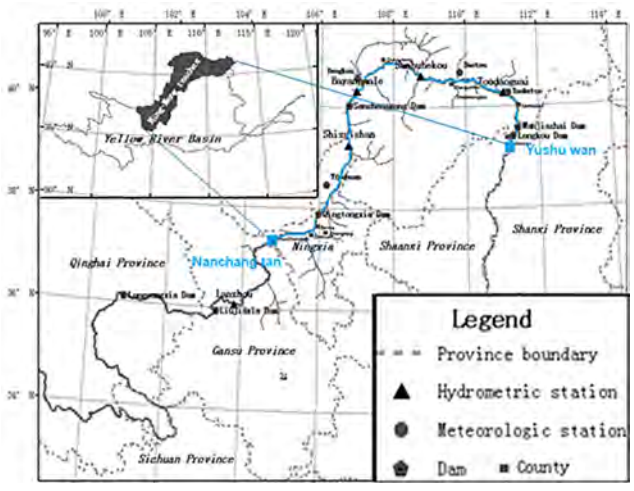

Fig. 1. Location of the Ning–Meng reach.

The potential for using fuzzy expert systems to forecast the potential risk of ice jam was discussed by Mahabir et al. (2002). Based on this research, more research has since been carried out (Mahabir et al., 2006; Wang et al., 2008). Data-driven modelling has been applied to the ice regime simulation several times and one example is the Yellow River (Chen et al., 2012). However, the limitation of using such methods is that ice processes cannot be understood by data-driven models.

3 Case study description

3.1 The Ning–Meng reach

The Yellow River is the second longest river in China and is also one of the most famous rivers in the world. Due to the fact that it is the cradle of the Chinese civilization, the river is also called the "Mother River of China". The Yellow River originates from the Bayanhar Mountain in Qinghai Province and flows through the nine provinces of Qinghai, Sichuan, Gansu, Ningxia, Inner Mongolia, Shaanxi, Shanxi, Henan, and Shandong, and finally flows into the Bohai Sea in Shandong Province. It has a total length of 5464 km and a basin area of 752 443 km^2.

The Ning–Meng reach of the Yellow River (Fig. 1) is where the ice flood mainly happens. The lenght of the Ning–Meng reach is 1237 km and it consists of two consecutive parts, the Ningxia and Inner Mongolia reaches. The starting point for it is Nanchangtan, Zhongwei County in the Ningxia Hui autonomous region, and the end point is Yushuwan, Mazha Town, Zhungeerqi County in the Inner Mongolia autonomous region.

The Qingtongxia and Sanshenggong reservoirs are located on this reach. The total length of the Ningxia reach is 397 km and it flows from southwest to northeast. On this reach one

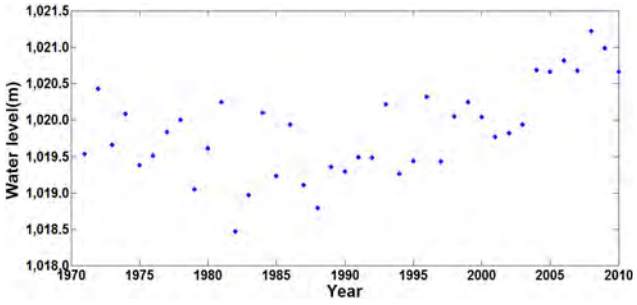

Fig. 2. Hystorical maximum recorded water levels.

part, from Nanchangtan to Zaoyuan, does not usually freeze because the bed slope is steep and the water velocity is large, and it only freezes in a very cold winter. The rest of the reach, from Zaoyuan to Mahuanggou, flows from south to north and it often freezes because the slope is gentle and the water velocity is low. The Inner Mongolia reach is located on the north of the Yellow River basin between 106°10′ E–112°50′ E, 37°35′ N–41°50′ N. The total length of the main stream is 830 km, and the total height difference is 162.5 m. The reach is wide with a gentle slope and meandering twists and turns. Although it is located in the middle and lower reaches, the slope of the reach that is under study is close to that at the Yellow River estuary.

3.2 Past ice floods

According to historical data for the Yellow River, ice disasters appeared frequently in the past and every year between 1855 and 1949, when dams were destroyed by ice floods 27 times. In the period between 1951 and 1955, the dams on the Lijing breach were also destroyed by ice flood. In addition, there were 28 ice flood seasons with ice disasters between 1951 and 2005 (Rao et al., 2012).

Especially on the Ning–Meng reach, ice jams, frazil jams, and other ice flood disasters have happened frequently. According to the statistics for ice flood disasters on the Ning–Meng reach, there were 13 ice flood events from 1901 to 1949, almost one event every 4 years. Even after the Liujiaxia Reservoir and the Qingtongxia Reservoir started operation in 1968 and 1960 respectively, ice flood disasters still occurred in 1993, 1996, 2003, and 2008 (Gao et al., 2012).

During the winter season 2007/2008, the water storage in the Inner Mongolia reach reached the largest value, 1835 million, 730 million more than during the normal flooding condition (1105 million). As a result, the water level at Sanhuhekou hydrological station reached a recorded level of 1021.2 m, 0.41 m higher than any recorded level at the station, which led to dike break at two sites, causing serious economic losses. During the break-up period in the winter 2005/2006, the water level at Sanhuhekou station reached 1020.81 m, for a maximum discharge of 772. Figure 2 show

the maximum recorded water levels, every decade between 1970 and 2009, at Sanhuhekou hydrological station.

The Yellow River basin is under the influence of cold air from the vicinity of Siberia and Mongolia; the mean air temperature during the whole winter is below 0°, and this lasts for 4 to 5 months. Due to the special geographical location and difference in latitude of the Ning–Meng reach on the Yellow River, together with the special river flow direction, the downstream part both freezes and breaks up earlier than the upstream part. Hence it is very easy for ice jams or backwater to be formed, which can result in ice flood disasters such as dams being destroyed and dike breaks. Based on the literature review on ice flood disasters, one can conclude that on the Yellow River the main problems regarding ice floods are caused by ice-dam floods and ice-jam floods which can result in dike breaks and overtopping on the embankment. This is the key problem to be solved on the Ning–Meng reach of the Yellow River.

4 Modelling application

4.1 Model set-up

Present research uses the YRIDM, which is a 1-D unsteady-state model. The YRIDM is based on the RICE and RICEN models, and designed taking river ice processes into account.

The YRIDM model is simulating the ice dynamics, by describing the skim ice formation and frazil ice as it is in the RICE model. Skim ice formation is based on the empirical equation, developed by Matousek (1984); frazil ice along the river channel is described by the mathematical model defined by Shen and Chiang (1984) and the ice dynamics is described by the static border of the ice formation, which is defined by Svensson et al. (1989) in form of a critical value. The other elements of the ice dynamics are the growth of the border ice, the undercover accumulation, erosion of the ice, and erosion of the ice cover due to thermal growth. Though YRIDM is based on the RICE model, there is a major difference in how the break-up of the ice is checked in the two models. In the case of RICE, the critical value of discharge for which the break-up appears is a constant value that has to be determined before using the model, by experiments on site, while in the YRIDM model the critical discharge is determined by an empirical equation.

The YRIDM model was designed to simulate the ice flood for unsteady-state flow, which uses the governing Eqs. (1)–(5) below:

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0, \tag{1}$$

$$\frac{\partial Q}{\partial t} + \frac{\partial \left(Q^2/A\right)}{\partial x} = -gA\left(\frac{\partial Z}{\partial x} + S_\mathrm{f}\right), \tag{2}$$

$$S_\mathrm{f} = n_\mathrm{c}^2 \frac{|u|u}{R^{4/3}}, \tag{3}$$

$$n_\mathrm{c} = \left[\left(n_\mathrm{i}^{3/2} + n_\mathrm{b}^{3/2}\right)/2\right]^{2/3}, \tag{4}$$

$$n_\mathrm{i} = n_\mathrm{i,i} + (n_\mathrm{i,i} - n_\mathrm{i,e})e^{-\alpha_\mathrm{n} T}, \tag{5}$$

where Q = discharge ($\mathrm{m^3\,s^{-1}}$); A = net flow cross-sectional area ($\mathrm{m^2}$); x = distance (m); t = time (s); g = gravitational acceleration ($\mathrm{m\,s^{-2}}$); S_f = friction slope; n_c = equivalent roughness; n_i = roughness of ice cover; n_b = roughness of river bed; u = average velocity ($\mathrm{m\,s^{-1}}$); R = river hydraulic radius (m); $n_\mathrm{i,i}$ = initial roughness of ice cover; $n_\mathrm{i,e}$ = final roughness of ice cover at the end of an ice-covered period ($0.008 \sim 0.012$); T = freeze-up time (day); and α_n = decay coefficient.

The YRIDM equations for ice transportation model are different from the ones in RICE, that is, YRIDM uses convection–diffusion equation between ice run and floating ice, while RICE uses a two-layer mode for the surface ice transportation. The critical value of discharge for which the break-up appears is determined by the empirical Eq. (6):

$$Q \geq \frac{9.5 * a * h_\mathrm{i}}{\sum T_{\mathrm{a}+} + 1}, \tag{6}$$

where Q is discharge ($\mathrm{m^3\,s^{-1}}$); h_i is the ice cover thickness when air temperature varies from positive values to negative ones (m); a is the break-up coefficient ($a = 4$ for mechanical break-up and $a = 22$ for thermal break-up); and $T_{\mathrm{a}+}$ is the average daily accumulated positive air temperature, measured from the day when air temperature turns positive.

The roughness of ice cover is usually computed based on the end roughness of ice cover and a decay constant; however, the YRIDM model is developed with a constant initial roughness of ice cover, which has been determined by YRCC through several measuring campaigns.

The logical flowchart of the YRIDM model is presented in Fig. 3. It consists of three main components, the river hydrodynamics, thermodynamics, and ice dynamic modules. The advantage of this model is that it can be subdivided even further into the following modules: river hydraulics, heat exchange, water temperature and ice concentration distributions, ice cover formation, ice transport and cover progression, undercover deposition and erosion, and thermal growth and decay of ice covers.

Based on the available data analysis for the Ning–Meng reach of the Yellow River, and literature review on ice models, the YRIDM has been selected to conduct the modelling for ice flood middle-term prediction (10–15 days ahead). The ice regime on the Ning–Meng reach is mainly determined by thermal factors, dynamic factors, channel course conditions and human activities. During the whole research period between 2001 and 2011, the most serious ice flood occurred during two winters, namely the winter from 2007 to 2008 and the winter from 2008 and 2009. The winter from 2008 to 2009 was chosen as the simulation winter to calibrate the

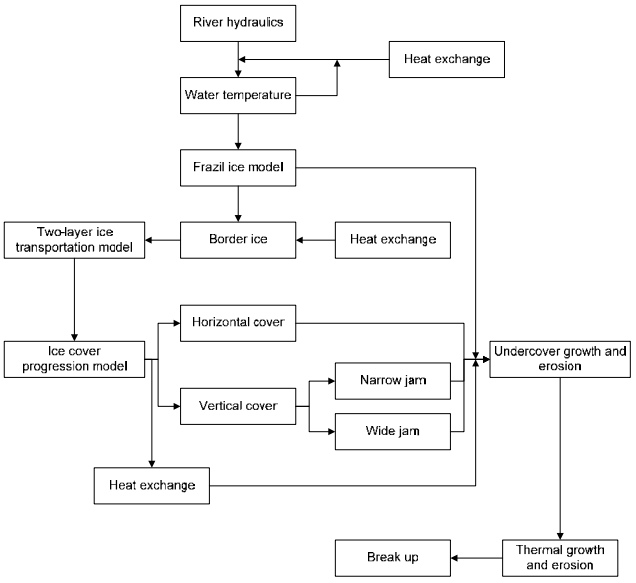

Fig. 3. YRCC river ice dynamic model flowchart.

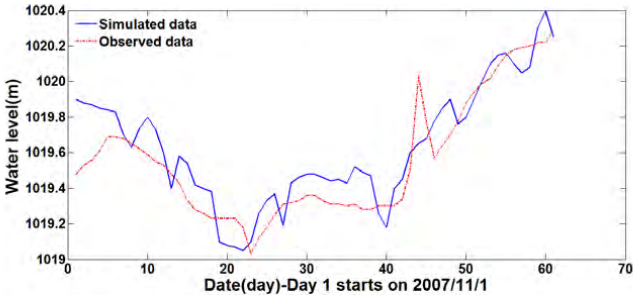

Fig. 4. Verification of the model: water levels.

model, and the winter from 2007 to 2008 was used to verify the calibrated model.

The reach between Bayangaole station and Toudaoguai station was chosen to be modelled due to the fact that ice flood has happened frequently here and data was available on the reach. The simulation reach has a total length of 475.6 km. A hydrological station named Sanhuhekou is positioned in the middle of the reach (i.e. 205.6 km away from Bayangaole station).

The available data includes river hydraulics, meteorological, ice regime, and cross sectional data at the four hydrometric stations, namely Shizuishan, Bayangaole, Sanhuhekou, Toudaoguai stations, and it covers the period of ten winters, from 2001 to 2011. River hydraulic data includes water level and discharge with daily measuring frequency. Meteorological data includes air and water temperature with a daily measurement of the water temperature and twice a day measurement of air temperature (the daily highest and lowest temperature per day). Ice regime data includes ice run data, freeze-up date, break-up date, and ice cover thickness. The measured frequency is per winter except that the measured

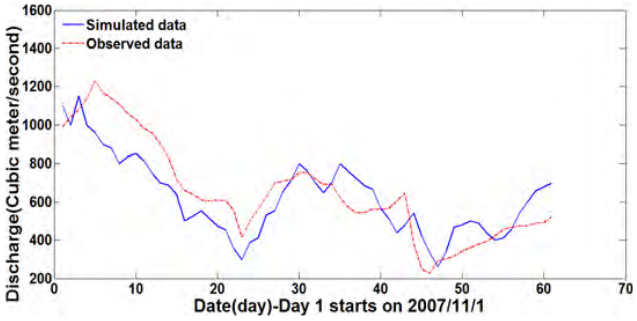

Fig. 5. Verification of the model: discharge.

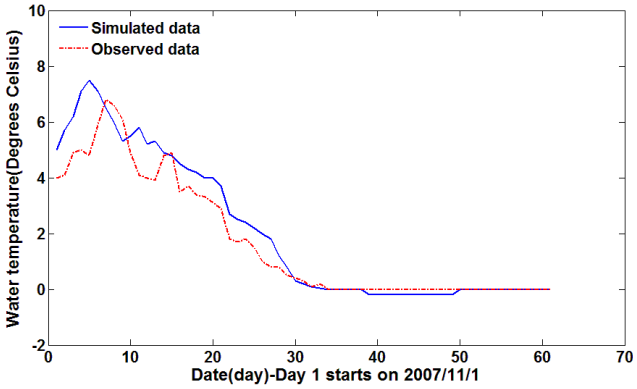

Fig. 6. Verification of the model: water temperature.

frequency of ice cover thickness is every 5 days. Data for four cross sections for 2009 is available.

During the winter of 2008/2009 the discharge and water temperature at Bayangaole station, the water level at Toudaoguai station, air temperature at Sanhuhekou and Toudaoguai stations, and cross sections and bed elevation were used to build the model.

4.2 Model calibration and verification

During the winter of 2008/2009, daily mean discharge at the Bayangaole station is taken as the upstream boundary condition to the model, and the daily mean water level at the Toudaoguai station was set as the downstream boundary condition. The air temperatures at the Sanhuhekou and Toudaoguai stations were input into the model, and the air temperature at the Sanhuhekou station was used to cover the reach from the Bayangaole station to the Sanhuhekou station, and at the same time, the air temperature at the Toudaoguai station was used to cover the reach from the Sanhuhekou station to the Toudaoguai station. The daily lowest and highest temperatures were interpolated by sine function. Daily mean water temperature at the Bayangaole station was input into the model as the upstream boundary condition of water temperature. The initial conditions were water temperature, water level, and discharge.

Table 1. Results of the sensitivity analysis of the parameters of the YRCC river ice dynamic model.

Sensitivity parameters	Physical meaning	Reference range	Unit	Sensitive object
Nb	Bed roughness	0.019 ∼ 0.045	–	Water level
$n_{i,e}$	End-ice roughness	0.008 ∼ 0.035	–	Water level
α_n	Decay constant	0.005 ∼ 0.4	1/day	Water level
Coe_Cw	Heat exchange coefficient between water and ice	15 ∼ 18	–	–
Coe_Hia	Heat exchange coefficient between ice and air	6 ∼ 12	–	Ice cover thickness
Coe_Co	Heat exchange coefficient between water and air	15 ∼ 25	–	Ice cover thickness

Table 2. Calibrated parameters values.

Parameters	Physical meaning	Value	Unit
N1	Bed roughness (cross1–cross26)	0.011	–
N2	Bed roughness (cross27–cross51)	0.017	–
N3	Bed roughness (cross52–cross87)	0.045	–
N4	Bed roughness (cross88–cross121)	0.004	–
α_n	Decay constant	0.001	1/day
$n_{i,e}$	End-ice roughness	0.01	–
Coe_Cw	Heat exchange coefficient between water and ice	16.62	–
Coe_Hia	Heat exchange coefficient between ice and air	12	–
Coe_Co	Heat exchange coefficient between water and air	19	–

Based on the literature review and the reference book of the numerical model from YRCC, the sensitivity parameters are listed in Table 1. The one-at-a-time sensitivity measure method was used to conduct the sensitivity analysis, which means the value of one parameter is changed from the minimum value to the maximum value while at the same time other parameters are kept constant at the mean value, and then the variation of the model output is checked.

After the sensitivity analysis, scenarios of the model calibration can be designed based on the sensitivity analysis results (Table 1). If the model is sensitive to a certain parameter, then the parameter needs to be calibrated carefully, otherwise the default value is used. In this case, bed roughness, end-ice roughness, and decay constant are sensitive to the water level at the Sanhuhekou station; heat exchange coefficient between ice and air, and heat exchange coefficient between water and air are sensitive to ice cover thickness at the Sanhuhekou station. These five parameters should be calibrated carefully; the heat exchange coefficient between water and ice is not sensitive to the output, hence the default value can be set. The model can be calibrated under different ice regime conditions, which means that some parameters can be calibrated under open water conditions, such as the Manning coefficient of river beds and the heat exchange coefficient between water and air. Some parameters can be calibrated under ice conditions, such as the Manning coefficient under ice cover, end-ice roughness, and decay constant. Hence the model calibration procedure is divided into a model calibration under open water conditions and a model calibration under ice conditions. In the calibration procedure, RMSE (root mean square error) is used as the criterion to check the model performance.

The calibrated bed roughness for different cross sections was varied for a large range of values (i.e from 0.004 to 0.017). These values are provided by YRCC and reflects the four different types of channel beds of the modelled river reach. The Ningxia River reach, which is 397 km long, starts at Nanchangtan and ends at Mahuanggou, in Shizuishan city. From Nanchangtan to Zaoyuan (ca. 135 km) there is an uncommon freezing state of the river reach, because the bed slope is steep and water velocity is high. This state allows the river reach to freeze up only during very cold winters. On the reach from Zaoyuan to Mahuanggou (ca. 262 km), on the other hand, there is a common freezing state of the river reach because the slope is gentle and water velocity is small. The Inner Mongolia reach is located on the north of Yellow River basin and has a total length of 830 km. The reach is wide with gentle slopes, meandering twists and turns. Although it is located in the middle and lower Yellow River, its slope is close to that of the Yellow River estuary. A summary of the slopes, width and roughness of these four types of considered reaches is given in the Table 3.

After calibration of YRIDM during the winter of 2008/2009, the discharge and water temperature at the Bayangaole station, the water level at the Toudaoguai station, air temperature at the Sanhuhekou and Toudaoguai stations, and cross sections and bed elevation during the winter of 2007/2008 were used to build the model. The simulated

Table 3. Channel characteristics of Ning–Meng reach.

Autonomous region	Section	Channel type	Channel length (km)	Average channel width (m)	Main channel width (m)	Channel slope (‰)	Roughness
Ningxia	Nanchangtan–Zaoyuan	valley type	135	200–300		0.8–1.0	0.005–0.014
	Zaoyuan–Mahuanggou	transition type	262	500–1000		0.1–0.2	0.005–0.014
Inner Mongolia	Mahuanggou–Wuda bridge	valley type	69.0	400	400	0.56	0.011–0.020
	Wuda bridge–Sanshenggong	transition type	106.6	1800	600	0.15	0.011–0.020
	Sanshenggong–Sanhuhekou	wandering type	205.6	3500	750	0.17	0.009–0.018
	Sanhuhekou–Zhaojunfen	transition type	126.2	4000	710	0.12	0.009–0.018
	Zhaojunfen–Lamawan	bend type	214.1	3000–2000	600	0.10	0.002–0.010
	Lamawan–Yushuwan	valley type	118.5				0.002–0.010
Total	Nanchangtan–Yushuwan		1237				

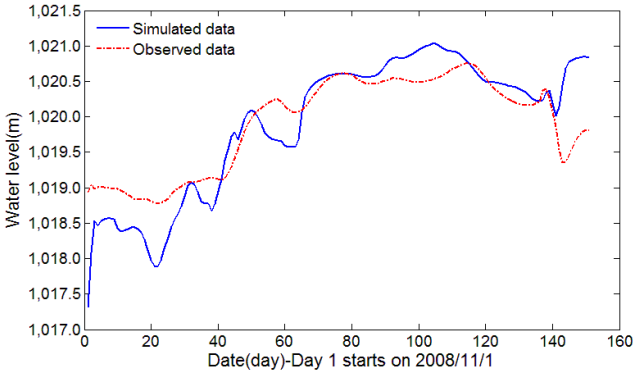

Fig. 7. Water levels at Sanhuhekou station.

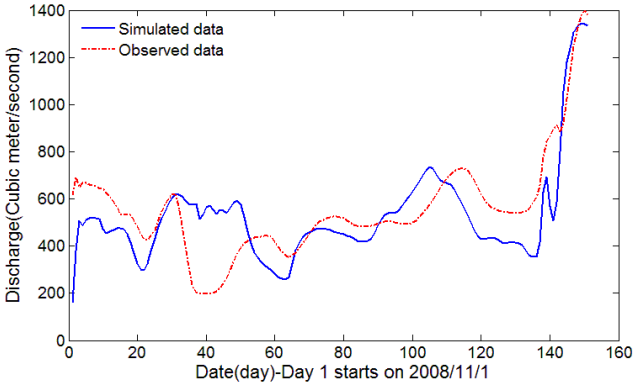

Fig. 8. Discharge at Sanhuhekou station.

water level, discharge, water temperature, and ice cover thickness at the Sanhuhekou station during the winter of 2007/2008 were used to compare with the observed values to verify the calibrated model.

The simulation period used for verification of the model was done for the available data from 1 November 2007 to 31 December 2007. Results of the verification are presented in Figs. 4, 5 and 6 for water level, discharge and temperature, respectively. Though there was just little data measured and available for verification, results shows that the model could capture the trend of water level, discharge, and water temperature. It is worth to be noted that in 2008/2009, the water level did not exceed the embankments.

4.3 Model uncertainty analysis

Based on the sensitivity analysis results, the water level at the Sanhuhekou station is sensitive to the Manning coefficient of river bed, decay constant, and end-ice roughness; and the ice cover thickness at the Sanhuhekou station is sensitive to the heat exchange coefficient between ice and air, and heat exchange coefficient between water and air. Hence, the uncertainty analysis is divided into an uncertainty analysis of the water level at the Sanhuhekou station and the uncertainty analysis of ice cover thickness at the Sanhuhekou station, and the Monte Carlo simulation is used to conduct the parametric uncertainty analysis (Moya Gomez et al., 2013).

Due to the limitation of time, it was impossible to run the model several times. When conducting the uncertainty analysis of the water level at the Sanhuhekou station, the related four parameters were the Manning coefficients of the river bed at the upstream and downstream of the Sanhuhekou station, decay constant, and end-ice roughness. The scenarios of the uncertainty analysis of the water level at the Sanhuhekou station were designed based on the calibrated parameters, the range was calculated by increasing and decreasing the calibrated value by 20 %, the sample generation was uniformly random, and the number of simulation was 500. When conducting the uncertainty analysis of the ice cover thickness at the Sanhuhekou station, the two related parameters were the heat exchange coefficient between ice and air, and the heat exchange coefficient between water and air. The scenarios of the uncertainty analysis about ice cover thickness at the Sanhuhekou station were designed based on the calibrated parameters, the range was calculated through increasing and decreasing the calibrated value by 20 %, the sample generation was uniformly random, and the number of simulations was 400.

Based on the above case designs, the parameters were input into the model, the model was run and the result were

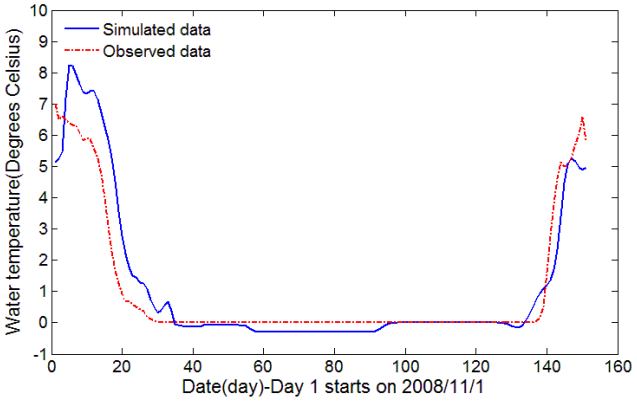

Fig. 9. Water temperature at Sanhuhekou station.

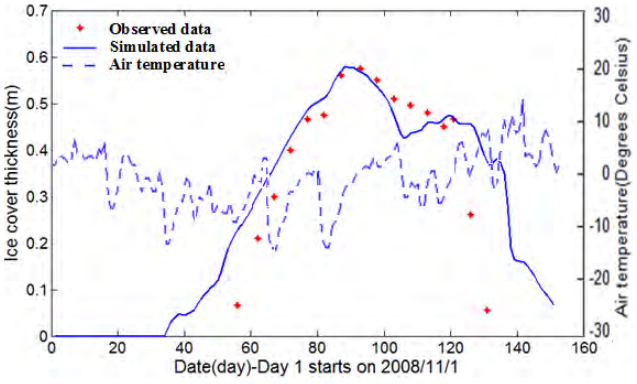

Fig. 10. Ice cover thickness at Sanhuhekou station.

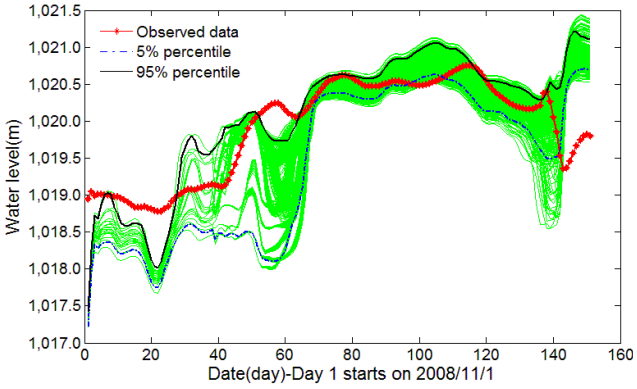

Fig. 11. Water levels at Sanhuhekou station.

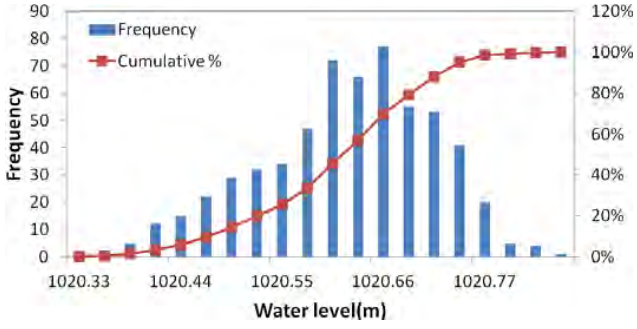

Fig. 12. Probability distribution of the water level on Day 90.

stored, the distribution and quartile of the output were analysed, namely the pdf (probability density function) at one time step and two bounds (5 and 95 %).

5 Results and discussion

The results of the sensitivity analysis show that bed roughness, end-ice roughness, and the decay constant are sensitive to the water level at the Sanhuhekou station, and that the heat exchange coefficient between ice and air, and the heat exchange coefficient between water and air are sensitive to ice cover thickness at the Sanhuhekou station.

Based on the results of the sensitivity analysis, the model was calibrated, and the values of calibrated parameters can be seen in Table 2.

Figure 7 shows the water level comparison at the Sanhuhekou station during the winter of 2008/2009 between the simulated data and observed data. The resulting RMSE is 0.464. The simulated result is acceptable and reasonable. From day 1 to day 40 the simulated result is not good, but the trend in the simulated results is consistent with the observed results, which is because in the beginning, the simulated water level is affected by the initial conditions. From day 125 to the end, the simulated result is not accurate and the trend is even adverse, which is because the model cannot simulate an ice-jam break-up during the break-up period. The data observed during the period clearly shows that the water level decreases suddenly, which means that the ice jam at the downstream of the Sanhuhekou station collapses, but the model cannot simulate this phenomenon, which is why the simulated results are not so accurate. During the rest of the period, the results are acceptable.

Figure 8 shows a comparison between the simulated discharge data and observed discharge data at the Sanhuhekou station during the winter of 2008/2009. The RMSE is 160.5 compared with the peak discharge of 1400 $m^3 s^{-1}$, so the error rate is 11.2 %, and hence the simulated result is acceptable and reasonable. From day 1 to day 25, the simulated results are not accurate, but the trend in the simulated results is consistent with the observed trend, which is because in the beginning, the simulated water level is affected by the initial conditions. From day 25 to day 50, the simulated results are also not accurate. The discharge decreases suddenly; this is because during the freeze-up period, the channel storage capacity increases when the water has been changed into ice, then the amount of water decreases, which can result in a sudden decrease in discharge, but the discharge can increase again when the channel storage capacity of the river continues to be stable. However, there is an assumption that the

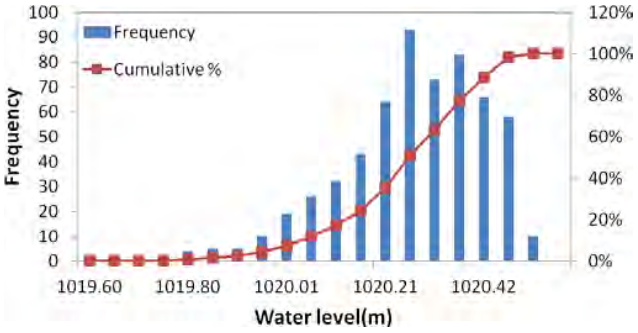

Fig. 13. Probability distribution of water level on Day 130.

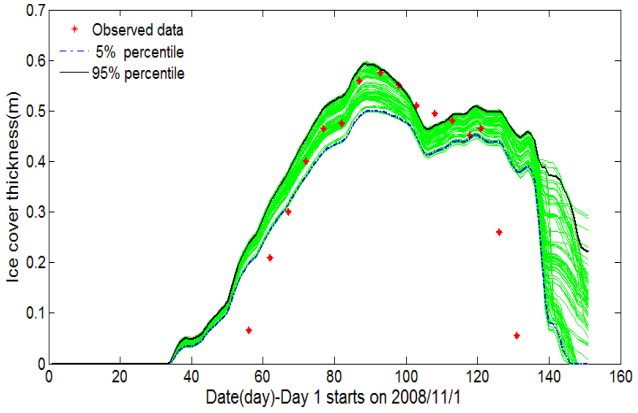

Fig. 14. Ice cover thickness at Sanhuhekou station.

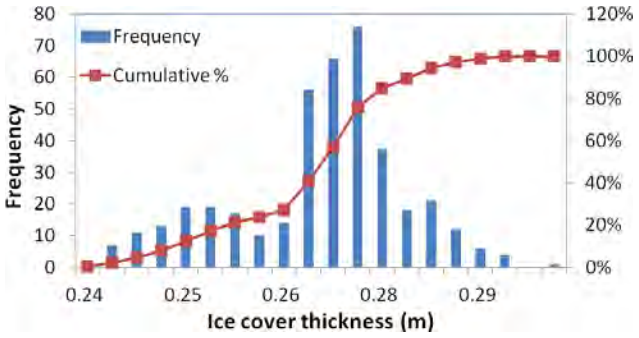

Fig. 15. Probability distribution of ice cover thickness on Day 60.

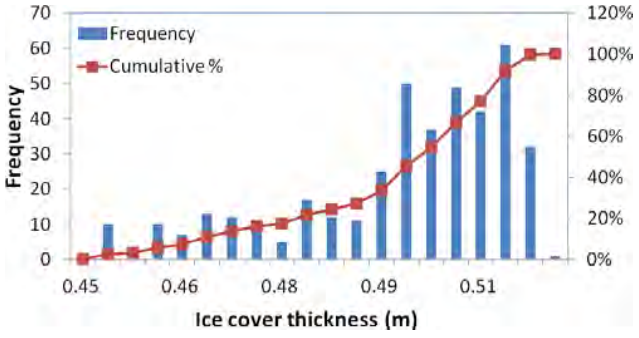

Fig. 16. Probability distribution of ice cover thickness on Day 120.

model ignores the effect on the water body mass balance due to changes in the ice phase; this is the reason why the results are not good enough. During the other period, the result is acceptable.

Figure 9 shows the comparison between the simulated data and observed data for the water temperature at the Sanhuhekou station during the winter of 2008/2009; the RMSE is 0.854. The simulated results are acceptable and reasonable. From day 1 to day 15, the simulated results are not good, which is because in the beginning, the simulated water temperature is affected by the initial water temperature. From day 55 to day 95, the water temperature is below 0°; this is the super cooling phenomenon, which means before the formation of ice cover the water body loses heat very quickly, which can result in the negative value for the water temperature. However, once the ice cover forms, it can prevent heat exchange between the water body and the air; this is the reason why the water temperature keeps steady at 0° between day 100 and day 130. During the rest of the period, the results are acceptable.

Figure 10 shows the comparison between the simulated data and observed data for ice cover thickness at the Sanhuhekou station during the winter of 2008/2009. Because the ice cover thickness data was measured per five days, the observed ice cover thickness data is not continuous and the RMSE between the simulated data and observed data is 0.109 m. The simulated results can describe the variation trend for ice cover thickness, and the simulated maximum value of the ice cover thickness, 0.58 m on 28 January 2009, has the same value as the measured one

The prepared data during the winter of 2007/2008 were input into the model to verify the calibrated model. In the winter of 2007/2008, the water level exceeded the height of embankment at Duguitalakuisu county, and resulted in a dike breach, which could not be captured by the model because as per the reference guide for the model, it cannot work when the water level exceeds the height of the embankment. In this respect YRIDM should be improved so that it has the ability to deal with such kind of problems.

After running the model, the uncertainty analysis results can be summarized in Fig. 11. Figure 11 shows the observed data, 5 % percentile bound, 95 % percentile bound, and the results of 500 cases. Day 90 (Fig. 12) and Day 130 (Fig. 13) are chosen to show the probability distribution. When it comes to the probability distribution of the uncertainty analysis results, if the distribution looks like a normal distribution, the uncertainty analysis results are good. According to Figs. 12 and 13, the probability distribution of the water level on day 90 looks good, which is because it looks like the normal distribution. Although the probability distribution of the water level on day 130 is a skewed normal distribution, at

least it shows the basic shape of normal distribution, and hence the uncertainty analysis results are reasonable.

After running the model, the uncertainty analysis results can be summarized in Fig. 14. Figure 10 shows the observed data, 5 % percentile bound, 95 % percentile bound, and the results of 400 cases. Day 60 (Fig. 15) and Day 120 (Fig. 16) are chosen to show the probability distribution. According to Figs. 15 and 16, the uncertainty analysis results are not good, due to the fact that the probability distribution of the ice cover thickness on these two days is not a normal distribution. This result is due to the fact that the number of cases designed for the uncertainty analysis might not be sufficient to show the characteristics of a model uncertainty.

Results show the sensitivity of bed roughness, end-ice roughness, and decay constant to water level at the Sanhuhekou station; and that both the heat exchange coefficient between ice and air and the heat exchange coefficient between water and air are sensitive to ice cover thickness at the same station. The model however cannot work when the water level exceeds the height of embankments, nevertheless reasonable results about uncertainty analysis can be achieved. Based on the obtained results it can be concluded that the model can be applied to the Ning–Meng reach to simulate its ice regime, and once calibrated, it can be used to forecast the ice regime to support decision making, such as on artificial ice-breaking and reservoir regulation

6 Conclusions

Tracking ice formation from observations and combining with numerical model predictions for advanced warning requires proper understanding of all scientific issues that play a role. In the case of the Yellow River, ice floods impose a threat every year, which is why the Yellow River Commission is putting considerable effort in verifying theoretical formulations with actual field measurements in order to better understand the scientific mechanisms that play a role. Transforming this knowledge into an early warning system that can help save lives is a scientific issue that requires attention (Almoradie et al., 2014; Jonoski et al., 2014; Popescu et al., 2012).

Based on the obtained results with the YRIDM model it can be concluded that it is applicable to the Ning–Meng reach for simulating ice regimes, and once calibrated, it can be used to forecast the ice regime to support decision making, such as on artificial ice-breaking and reservoir regulation. There is however a limitation of the YRIDM model that it cannot simulate a mechanical break-up during the break-up period because the effect of ice phase change on the water body mass balance is ignored.

The used model has the advantage that for 10–15 days forecasts of the meteorological data that are used as hydrological and hydraulic input into the model, the ice regime can be predicted and support decision making.

As the data collection continues the base to determine the possible times of ice formation and consequently flooding is enlarged and improved (Debolskaya, 2009). Based on ice-formation predictions decision makers can take appropriate measures to reduce the risk of flooding. Flooding during cold season is very important, therefore determination of the moments of ice formation that could possibly eliminate flooding, due to the decisions taken is also an important task in modelling. For ice formation and based on data availability a 1-D model is sufficient to be used; however, for the determination of the flood extent and time of flood occurrence, a more complex model, such as a 1D-2D, needs to be made available (Gichamo et al., 2013; Shen et al., 2008).

Though the present research focussed on ice formation rather than floods it can be generally concluded that the measured elements and frequency should be increased, and as recommendation if floods need to be captured and simulated then the one-dimensional models should be extended to two-dimensional models.

Acknowledgements. The authors acknowledge the support of the Hydrology Bureau of the Yellow River Conservancy Commission for the data used for the present research as well as for the YRIDM The work presented here was partially supported by the Ministry of Water Resources, China through grant no. 201301062-03.

Edited by: A. Gelfan

References

Almoradie, A., Jonoski, A., Stoica, F., Solomatine, D., and Popescu, I.: Web-based flood information system: case study of Somesul-Mare, Romania, J. Environ. Eng. Manage., 12, 1065–1070, 2013.

Beltaos, S.: Challenges and opportunities in the study of river ice processes, Proceedings of 19th International Association of Hydraulic Research Symposium on ice, Vancouver, British Columbia, Canada, 1, 29–47, 2008.

Beltaos, S. and Wong, J.: Downstream transition of river ice jams, J. Hydraul. Eng., ASCE, 112, 91–110, 1986.

Blackburn, J. and Hicks, F.: Suitability of dynamic modeling for flood forecasting during ice jam release surge events, J. Cold Reg. Eng., 17, 18–36, 2003.

Calkins, D. J.: Accelerated Ice Growth in Rivers. CRREL Report 79–14, Hanover, NH, US Army, pp. 4, 1979.

Chen, D. L., Liu, J. F., and Zhang, L. N.: Application of Statistical Forecast Models on Ice Conditions in the Ningxia-Inner Mongolia Reach of the Yellow River, Proceedings of 21th International Association of Hydraulic Research Symposium on ice, Dalian, China, 443–454, 2012.

Chen, F., Shen, H. T., and Jayasundara, N.: A One-Dimensional Comprehensive River Ice Model. Proceedings of 18th International Association of Hydraulic Research Symposium on ice, Sapporo, Japan, 2006.

Daly, S. F.: Frazil ice dynamics. CRREL Monograph 84-1, US Army CRREL, Hanover, N. H., 1984.

Dahlke, H. E., Lyon, S. W., Stedinger, J. R., Rosqvist, G., and Jansson, P.: Contrasting trends in floods for two sub-arctic catchments in northern Sweden – does glacier presence matter?, Hydrol. Earth Syst. Sci., 16, 2123–2141, doi:10.5194/hess-16-2123-2012, 2012.

Debolskaya, E. I.: "Numerical Modeling of Ice Regime in Rivers", Hydrologic Systems Modeling, Vol. 1, Encyclopedia of Life Support Systems, Eolss Publishers, Oxford (UK), 137–165, 2009.

Gao, G. M., Yu, G. Q., Wang, Z. L., and Li, S. X.: Advances in Break-up Date Forecasting Model Research in the Ningxia-Inner Mongolia Reach of the Yellow River, Proceedings of 21th International Association of Hydraulic Research Symposium on ice, Dalian, China, 475–482, 2012.

Gichamo, T., Jonoski, A., Popescu, I., Morris, M., and Hassan, M.: Embankment failure modeling using HR Breach Model, J. Environ. Eng. Manage., 12, 865–874, 2013.

Hammar, L., Shen, H. T., Evers, K.-U., Kolerski, T., Yuan, Y., and Sobaczak, L: A laboratory study on freeze up ice runs in river channels. Ice in the Environments: Proceedings of 16th International Association of Hydraulic Research Symposium on ice, 3, Dunedin, 36–39, 2002

Haresign, M. and Clark, S.: Modeling Ice Formation on the Red River near Netley Cut. CGU HS Committee on River Ice Processes and the Environment, 16th Workshop on River Ice, Winnipeg, Manitoba, 147–160, 2011

Jasek, M.: Ice jam release surges, ice runs, and breaking fronts: field measurements, physical descriptions, and research needs, Can. J. Civil Eng., 30, 113–127, 2003.

Jonoski, A., Almoradie, A., Khan, K., Popescu, I., and Andel, S. J.: Google Android Mobile Phone Applications for Water Quality Information Management, J. Hydroinform., 15, 1137–1149, 2013.

Kivisild, H. R.: Hydrodynamic analysis of ice floods, Proceedings of 8th International Association of Hydraulic Research Congress, Montreal, 1959.

Lal, A. M. W. and Shen, H. T.: A Mathematical Model for River Ice Processes, J. Hydraul. Eng., ASCE, 177, 851–867, also USA Cold Regions Research and Engineering Laboratory, CRREL Report 93-4, 1993.

Li, H., Jasek, M., and Shen, H. T.: Numerical Simulation Of Peace River Ice Conditions, Proceedings of 16th International Association of Hydraulic Research Symposium on ice, Dunedin, New Zealand, Vol. 1, 134–141, 2002

Liu, C. J., Yu, H., and Ma, X. M.:Ice flood in Northeast region. 15th International Symposium on Ice, Gdansk, Poland, 303–312, 2000.

Liu, L. and Shen, H. T.: A two-Dimensional Comprehensive River Ice Model. Proceedings of 18th International Association of Hydraulic Research Symposium on ice, Sapporo, Japan, 69–76, 2006.

Liu, L. W. and Shen, H. T.: Dynamics of ice jam release surges, Proceedings of 17th International Association of Hydraulic Research Symposium on ice, Saint Petersburg, Russia, pp. 8, 2004.

Mahabir, C., Hicks, F. E., and Robinson Fayek, A.: Forecasting Ice Jam Risk At Fort Mcmurray, AB, Using Fuzzy Logic, Proceedings of 16th International Association of Hydraulic Research Symposium on ice, Dunedin, New Zealand, Vol. 1, 91–98, 2002.

Mahabir, C. F., Hicks, F. E., and Robinson Fayek, A.: Neuro-fuzzy river ice breakup forecasting system, J. Cold Reg. Sci. Technol., 46, 100–112, 2006.

Malenchak, J., Doering, J., Shen, H. T., and Morris, M.,: Numerical Simulation of Ice Conditions on the Nelson River, Proceedings of 19th International Association of Hydraulic Research Symposium on ice, Vancouver, British Columbia, Canada, Vol. 1, 251–262, 2008.

Moya-Gomez, V., Popescu, I., Solomatine, D., and Bociort, L.: Cloud and cluster computing in uncertainty analysis of integrated flood models, J. Hydroinform., 15, 55–69, 2013.

Morris, M., Malenchak, J., and Groeneveld, J.: Thermodynamic Modeling to Test the potential for Anchor Ice Growth in post-construction conditions on the Nelson River, Proceedings of 19th International Association of Hydraulic Research Symposium on ice, Vancouver, British Columbia, Canada, Vol. 1, 263–272, 2008.

Osterkemp, T. E.: Frazil ice formation : A review, J.Hydraul. Div., ASCE, 104, 1239–1255, 1978.

Petrow, Th., Merz, B., Lindenschmidt, K.-E., and Thieken, A. H.: Aspects of seasonality and flood generating circulation patterns in a mountainous catchment in south-eastern Germany, Hydrol. Earth Syst. Sci., 11, 1455–1468, doi:10.5194/hess-11-1455-2007, 2007.

Popescu, I., Jonoski, A., and Bociort, L.: Decision Support Systems for flood management in the Timis-Bega catchement, J. Environ. Eng. Manage., 11, 847–953, 2012.

Rao, S. Q., Yang, T. Q., Liu, J. F., and Chen, D. L.: Characteristics of Ice Regime in the Upper Yellow River in the last Ten Years, Proceedings of 21th International Association of Hydraulic Research Symposium on ice, Dalian, China, 390–396, 2012.

Rojas, R., Feyen, L., Dosio, A., and Bavera, D.: Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations, Hydrol. Earth Syst. Sci., 15, 2599–2620, doi:10.5194/hess-15-2599-2011, 2011.

She, Y. and Hicks, F.: Modeling ice jam release waves with consideration for ice effects, J. Cold Reg. Sci. Technol., 45, 137–147, 2006.

Shen, H. T. and Chiang, L. A.: Simulation of growth and decay of river ice cover, J. Hydraul. Div., Am. Soc. Civ. Eng., 110, 958–971, 1984.

Shen, H. T. and Lal, A. M. W.: Growth and decay of river ice covers. Proc., Cold Regions Hydrology Symp., AWRA, Fairbanks, 583–591, 1986.

Shen, H. T., Su, J., and Liu, L.: SPH Simulation of River Ice Dynamics, J. Comput. Phys., 165, 752–771, 2006.

Shen, H. T.: A trip through the life of river ice – research progress and needs, Keynote lecture, Proceedings of 18th International Association of Hydraulic Research Symposium on ice, Sapporo, Japan, 2006.

Shen, H. T., Gao, L., Kolerski, T., and Liu, L.: Dynamics of Ice Jam Formation and Release, J. Coast. Res., SI 52, 25–32, 2008.

Shen, H. T.: Mathematical modeling of river ice processes, Cold Reg. Sci. Technol., 62, 3–13, 2010.

Shen, H. T., Shen, H., and Tsai, S. M.: Dynamic transport of river ice, J. Hydraul. Res., 28, 659–671, 1990.

Shen, H. T. and Wang, D. S.: Under cover transport and accumulation of frazil granules, J. Hydraul. Eng., ASCE, 120, 184–194, 1995.

Shen, H. T., Wang, D. S., and Lal, A. M. W.: Numerical simulation of river ice processes, J. Cold Reg. Eng., ASCE 9, 107–118, 1995.

Shen, H. T. and Yapa, P. D.: A unified degree-day method for river ice cover thickness simulation, Can. J. Civil Eng., 12, 54–62, 1985.

Stefan, J.: Uber die theorien des eisbildung insbesondere uber die eisbildung in polarmure, Wien Sitzungsberichte Akademie der Wissenschaften, Series A, 42, 965–983, 1889.

Tesaker, E: Accumulation of frazil ice in an intake reservoir, Proc., IAHR Symposium on Ice problems, Hanover, NH, 1975.

Tuthill, A. M., Wuebben, J. L., and Gagnon, J. G.: ICETHK Users Manual, Version 1, Special Report 98-11, US Army Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire, 1998.

USACE: HEC-2 Water Surface Profiles, Hydrologic Engineering Center, US Army Corps of Engineers, Davis, California, 1992.

USACE: HEC-RAS River Analysis System: Hydraulic Reference Manual, Hydrologic Engineering Center, US Army Corps of Engineers, Davis, California, June, 1998.

Wang, T., Yang, K., and Guo, Y.: Application of Artificial Neural Networks to forecasting ice conditions of the Yellow River in the Inner Mongolia Reach, J. Hydrol. Eng., ASCE, 13, 811–816, 2008.

Ye, Q. S., Doering, J., and Shen, H. T.: A laboratory study of frazil evolution in a counter rotating flume, Can. J. Civil Eng., 31, 899–914, 2004.

Environmental Modelling & Software 41 (2013) 84–92

Contents lists available at SciVerse ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Parametric and physically based modelling techniques for flood risk and vulnerability assessment: A comparison

S.F. Balica [a,b,*], I. Popescu [a], L. Beevers [c], N.G. Wright [a,b,d]

[a] UNESCO-IHE, Institute for Water Education, P.O. Box 3015, 2601 DA Delft, The Netherlands
[b] Delft University of Technology, Postbus 5, 2600 AA Delft, The Netherlands
[c] School of the Built Environment, Heriot Watt University, Edinburgh EH14 4AS, UK
[d] School of Civil Engineering, University of Leeds, Leeds LS2 9JT, UK

ARTICLE INFO

Article history:
Received 21 June 2012
Received in revised form
29 October 2012
Accepted 6 November 2012
Available online

Keywords:
Floods
Vulnerability
Risk
Physically-based models
Flood Vulnerability Index

ABSTRACT

Floods are one of the most common and widely distributed natural risks to life and property. There is a need to identify the risk in flood-prone areas to support decisions for risk management, from high-level planning proposals to detailed design. There are many methods available to undertake such studies. The most accepted, and therefore commonly used, of which is computer-based inundation mapping. By contrast the parametric approach of vulnerability assessment is increasingly accepted. Each of these approaches has advantages and disadvantages for decision makers and this paper focuses on how the two approaches compare in use. It is concluded that the parametric approach, here the FVI, is the only one which evaluates vulnerability to floods; whilst although the deterministic approach has limited evaluation of vulnerability, it has a better science base.

1. Introduction

Floods are one of the most common and widely distributed natural risks to life and property. Damage caused by floods on a global scale has been significant in recent decades (Jonkman and Vrijling, 2008). In 2011, floods were reported to be the third most common disaster, after earthquake and tsunami, with 5202 deaths, and affecting millions of people (CRED, 2012). River, coastal and flash floods can claim human lives, destroy properties, damage economies, make fertile land unusable and damage the environment. The development of techniques, measures and assessment methodologies to increase understanding of flood risk or vulnerability can assist decision makers greatly in reducing damage and fatalities. Different methods to assess risk and vulnerability of areas to flooding have been developed over the last few decades. This paper aims to investigate two of the more widely used methods: traditional physically-based modelling approaches to risk assessment and parametric approaches for assessing flood vulnerability. The paper aims to present and discuss the benefits of each to decision makers.

1.1. Flood risk as a concept

The term "risk" in relation to flood hazards was introduced by Knight (1921), and is used in diverse different contexts and topics showing how adaptive any definition can be (Sayers et al., 2011). In the area of natural hazard studies, many definitions can be found. It is clear that the many definitions related to risk (Slovic, 1987; Alexander, 1993; IPCC, 2001; Plate Erich, 2002; Barredo et al., 2007) are interrelated and interchangeable and each of them has certain advantages in different applications (e.g. Sayers et al., 2011; Merz et al., 2007).

This study will consider risk as the product of two components, i.e. probability and consequence (Smith, 2004):

$$\text{Risk} = \text{Probability} \times \text{Consequence} \quad (1)$$

This concept of flood risk is strictly related to the probability that a high flow event of a given magnitude occurs, which results in consequences which span environmental, economic and social losses caused by that event. The EU Flood European Floods Directive 2007/60/EC (EC, 2007) and UNEP (2004) uses this definition of risk where "flood risk" means the combination of the probability of a flood event and of the potential adverse consequences for human health, the environment, cultural heritage and economic activity associated with a flood event.

* Corresponding author. UNESCO-IHE, Institute for Water Education, P.O. Box 3015, 2601 DA Delft, The Netherlands. Tel.: +31 (0)33670723396.
E-mail address: s.balica@unesco-ihe.org (S.F. Balica).

1364-8152/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.envsoft.2012.11.002

1.2. Hazard and flood hazard as a concept

"The probability of the occurrence of potentially damaging flood events is called flood hazard" (Schanze, 2006). Potentially damaging means that there are elements exposed to floods which may be harmed. Flood hazards include events with diverse characteristics, e.g. a structure located in the floodplain can be endangered by a 20-year flood and a water level of 0.5 m and by 50-year flood and a water level of 1.2 m. Heavy rainfall, coastal or fluvial waves, or storm surges represent the source of flood hazard. Generally these elements are characterised by the probability of flood event with a certain magnitude and other characteristics.

1.3. Vulnerability and Flood vulnerability as a concept

While the notion of vulnerability is frequently used within catastrophe research, researchers' notion of vulnerability has changed several times lately and consequently there have been several attempts to define and capture the meaning of the term. It is now commonly understood that "vulnerability is the root cause of disasters" (Lewis, 1999) and "vulnerability is the risk context" (Gabor and Griffith, 1980). Many authors discuss, define and add detail to this general definition. Some of them give a definition of vulnerability to certain hazards like climate change (IPCC, 2001), environmental hazards (Blaikie et al., 1994; Klein et al., 1999; ISDR, 2004), or the definition of vulnerability to floods (van der Veen and Logtmeijer, 2005; Connor and Hiroki, 2005; UN, 1982; McCarthy et al., 2001).

This study will use the following definition of vulnerability specifically related to flooding:

> *The extent to which a system is susceptible to floods due to exposure, a perturbation, in conjunction with its ability (or inability) to cope, recover, or basically adapt.*

2. The practice of flood risk and vulnerability assessment

Different methods to assess or determine hazard, risk and vulnerability to flooding have evolved through ongoing research and practice in recent decades (Xia et al., 2011; Hartanto et al., 2011; Gichamo et al., 2012). Two distinct method types can be distinguished and are considered in this paper:

- Deterministic modelling approaches which use physically based modelling approaches to estimate flood hazard/probability of particular event, coupled with damage assessment models which estimate economic consequence to provide an assessment of flood risk in an area.
- Parametric approaches which aim to use readily available data of information to build a picture of the vulnerability of an area.

Each method has developed from different schools of thought; the first approach mentioned is the traditional method which is routinely used in practice and academia alike. The second approach has evolved from several concerns such us: the internal characteristics of the system, global climate change and the political and institutional characteristics of the system. However, it takes a long time to develop the structural and non-structural measures required to prepare for flooding. In order to help guide such policy decisions, the development of a practical method for assessing flood vulnerability was needed. Among this need, this parametric approach points on vulnerability assessments to minimize the impacts of flooding and also to increase the resilience of the affected system.

2.1. The physically based modelling approach

Floods are primarily the result of extreme weather events. The magnitude of such an extreme event has an inverse relationship with the frequency of its occurrence i.e. floods with high magnitude occur less frequently than more moderate events. The relationship between the frequencies of occurrence and the magnitude of the extreme event is traditionally established by performing a frequency analysis of historical hydrological data using different probability distributions.

Once the frequency, magnitude and shape of the hydrograph are established, computer models which discretise the topographical river and land form are used to estimate flood depth, flood elevation and velocity (Hansson et al., 2008). Calculation of flood inundation depth and inundation extent is done using computational models based on solutions of the full or approximate forms of the shallow water equations. These types of models are one (1D) or two-dimensional (2D). 1D modelling is the common approach for simulating flow in a river channel, where water flow in the river is assumed to flow in one dominant direction which is aligned with the centre line of the main river channel. A 1D model can solve flood flows in open channels, if the shallow water assumptions that vertical acceleration is not significant and that water level in the channel cross-section is approximately horizontal are valid. However problems arise when the channel is embanked and water levels are different in the floodplain than in the channel and 2D models are needed in this situation. The hydraulic results from a computer model, such as inundation depth, velocity and extent can be used for loss estimation due to a particular design flood event. These parameters can then be linked to estimates of economic damage and loss in the affected area. Different models of damage and loss are available and are based on established economic relationships (ref).

This method relies on a significant amount of detailed topographic, hydrographic and economic information in the area studied. If the information is available, fairly accurate estimates of the potential risk to an area, as a result of economic losses, can be calculated. This type of flood hazard and associated economic loss information is reasonably easily communicated to the public. With the case of economic loss the public is used to hearing information provided in this manner. However, if the information for the model construction is not available, the method is likely to incur significant anomalies, which can call into question the validity of the assessment. These types of knowledge gaps and uncertainties are difficult to communicate effectively and can confuse decision makers and the public alike.

The scientific community therefore has researched methods that will overcome these problems. In this context it becomes important to evaluate the hazard, risk and vulnerability to flooding also from a different perspective: the parametric approach.

2.2. The parametric approach

The parametric approach, introduced in 80's by Little and Robin (1983), starts from the perspective of limited data, and has developed further since. The parametric approach aims to estimate the complete vulnerability value of a system by using only a few readily available parameters relating to that system, though the implementation of the approach is not simple.

Four types of parametric approaches have been developed by the scientific communities: i) estimating the complete vulnerability value of a system by using only few parameters relating to that system, ii) estimation of "the imputation of non-observable values" (Glynn et al., 1993), in which the observed parameters are used to

model the non-observed ones. (This assumption can be wrong), iii) the "parametric modelisation via maximum likelihood" (Little and Rubin, 1987), which is not a direct approach and is based on large number of assumptions; and iv) the "semi-parametric approach" (Newey, 1990) which allows modelling only of what is strictly necessary.

This study considers the first type of parametric approach, where the indicators and results rely on assumptions that cannot be validated from the observed data. This parametric approach tries to design a methodology that would allow the experts to assess the vulnerability results depend on the system characteristics and also to show the drawbacks, the practical and the philosophical in the specifications of the likelihood function (Serrat and Gomez, 2001).

In a general context, vulnerability is constructed like an instrumental value or taxonomy, measuring and classifying social, economic and environmental systems, from low vulnerability to high vulnerability. The vulnerability notion has come from different disciplines, from economics and anthropology to psychology and engineering (Adger, 2006); the notion has been evolving giving strong justifications for differences in the extent of damage occurred from natural hazards.

Whatever the exact measure of vulnerability one chooses to work with, the starting point is to estimate the right parameters of the process under the specification of the datasets. Vulnerability assessments have to be explicitly forward-looking. No matter how rich the data, the vulnerability of various systems is never directly obvious.

At spatial and temporal scales, several methodologies such as parametric-based approaches are applied to a vast diversity of systems: Environmental Vulnerability Index (EVI), Pratt et al., 2004; The Composite Vulnerability Index for Small Island States (CVISIS), Briguglio, 2003; Global Risk and Vulnerability Index (GRVI), Peduzzi et al., 2001; Climate Vulnerability Index (CVI), Sullivan and Meigh, 2003, etc.

This study uses a parametric approach proposed by Balica et al. (2009) and Balica and Wright (2010) to determine and index flood vulnerability for four system components (social, economic environmental and physical).

The parametric approach has some drawbacks, such as: an inevitable level of assumptions, the need for a sensitivity analyses, reliable sources and the subjective manner of interpreting the results.

2.3. Comparison of approaches

Physically based modelling and parametric approaches offer two different techniques for assessing flood risk and vulnerability. In light of these two distinct approaches, a clear question arises: what are the different advantages and disadvantages for decision makers using these techniques and "how do the two approaches compare in use?"

In order to answer this research question it is important to assess what decision makers require from these techniques in order to reach decisions. For the purposes of this study the following key components are identified:

- Information on the mechanism and cause of flooding (flood hazard) in the area studied.
- Information on the health and safety implications for the affected population of the flood hazard posed in the area, and the relative areas or population who are particularly vulnerable (and why).
- Information on the economic damage and losses expected in the area given a particular event.

In addition to these key components a fourth criteria was identified:

- How easily is this information communicated, both
  - From the expert undertaking the study to the decision-maker and
  - From the decision-maker to the public

This study will use the above identified criteria to compare the application of the two techniques (physically based modelling and the parametric approach) to a case study area in Budalangi, on the Nzoia River in Western Kenya. The paper aims to investigate the benefits and drawbacks of each approach, with the purpose of informing decision makers of the use.

3. Methodology

The scope of the present paper is to compare a parametric approach (Flood Vulnerability Index (FVI)) with traditional physically-based hydraulic modelling for flood risk analysis in order to determine what are the advantages of using one or the other in design and decision-making when flood hazard is involved. The general framework for the methodology is set out in Fig. 1.

3.1. Case study area

The Nzoia River originates in the South Eastern part of Mt. Elgon and the Western slopes of Cherangani Hills at an elevation of about 2300 m.a.s.l and it is one of the major rivers flowing into Lake Victoria. Nzoia River Basin covers an area of 12,709 km^2 in Western Kenya (Fig. 2). The Nzoia River discharges into Lake Victoria in Budalangi, Busia district. The river is of international importance, as it is one of the major rivers in Nile basin contributing to the shared water of Lake Victoria (NRBMI, nd).

The Nzoia River Basin is divided into three sub-catchments: the Lower Nzoia, characterised as flat and swampy; the Middle Nzoia and the Upper Nzoia, characterised with hills and steep slopes. The major tributaries of the Nzoia River are: Koitogos (Sabwani), Moiben, Little Nzoia, Ewaso Rongai, Kibisi, Kipkaren and Kuywa. The climate is tropical-humid and the area experiences four distinct seasons. Nzoia catchment has two rainy periods per year, one from March to May, with long rains and a second one from October to December, with short rains associated with ITCZ (the Inter Tropical Convergence Zone). The mean annual rainfall varies from a minimum of 1076 mm in the lowland to a maximum of 2235 mm in the highlands. Average annual volume of precipitation of the catchment is about 1740×10^6 m^3. The average temperature of the area varies from 16 °C in the upper catchment (highlands) to 28 °C in the lower catchment (lower semi-arid areas).

The dominant land use in the river basin is agriculture and the main agriculture production of the area are corn, sorghum, millet, bananas, groundnuts, beans, potatoes, and cassava and cash crops are coffee, sugarcane, tea, wheat, rice, sunflower and horticultural crops (Githui et al., 2008). The river basin plays a large

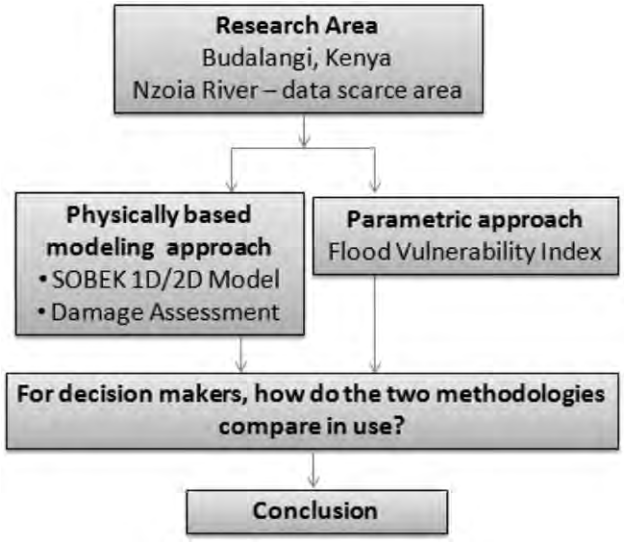

Fig. 1. Proposed methodology.

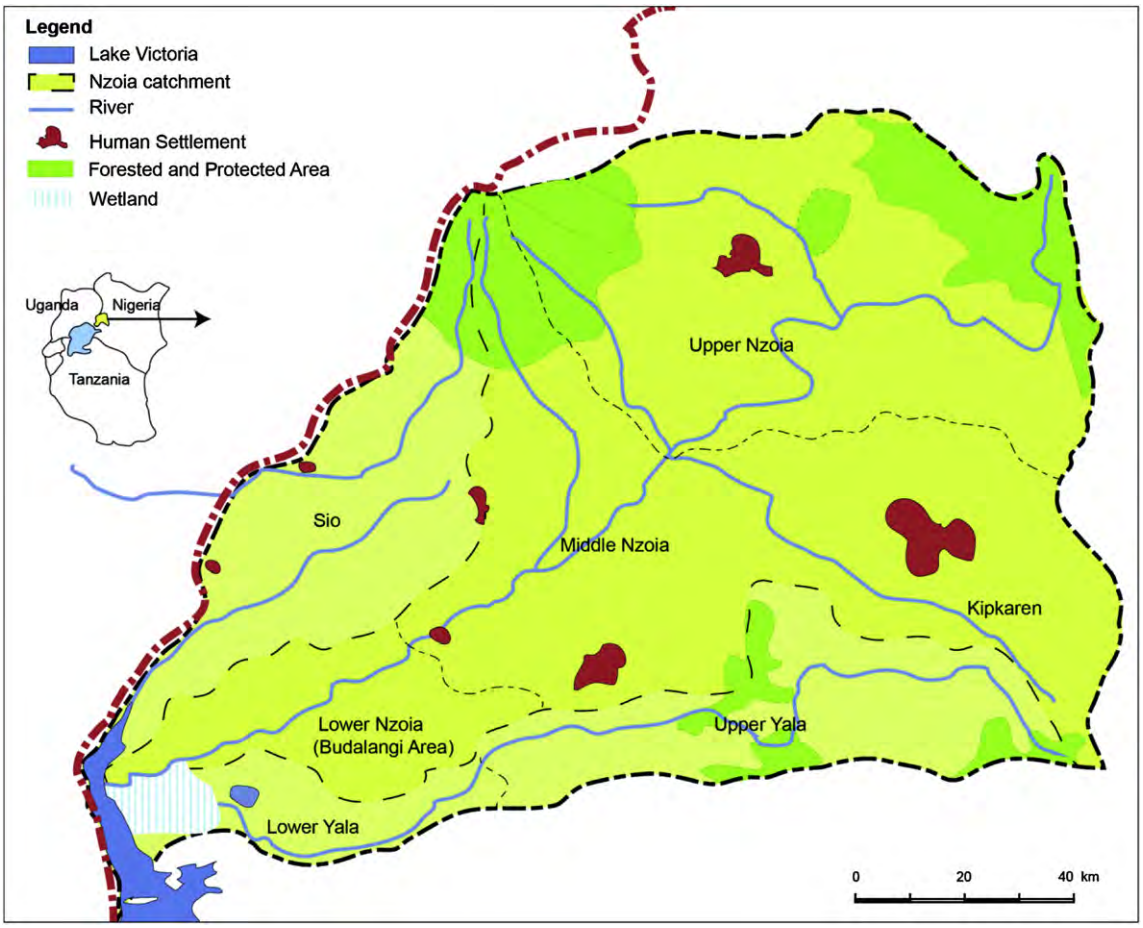

Fig. 2. Nzoia River Basin.

role in economic development at local and also at national level. Major problems and challenges in the basin are soil erosion and sedimentation, deforestation, flooding, and wetland degradation. The area located at the most downstream end of the catchment is, as previously mentioned the Budalangi area, which is the focus of the present study. Floods are frequent in the Budalangi area (WMO/MWRMD/APFM, 2004) and their impact is felt through loss of life, damage to property and agricultural/crop destruction.

This case study is data scarce area. The lower the accuracy in the data, the lesser the accuracy in the predictions, therefore in data scarce areas this can result in bad or poor vulnerability predictions. Consequently, the results of the two approaches chosen may prove which one is a more appropriate approach to be used by the decision makers in such cases.

3.2. *Assessing the flood risk of Budalangi region using physically based modelling*

There are many simulation models available for solving problems of unsteady or steady flow. In this present study, an unsteady flow analysis was carried out using the SOBEK 1D/2D tool, developed by Deltares. SOBEK 1D/2D couples one-dimensional (1D) hydraulic modelling of the river channel to a two-dimensional (2D) representation of the floodplains. The hydrodynamic 1D/2D simulation engine is based upon the optimum combination of a minimum connection search direct solver and the conjugate gradient method. It also uses a selector for the time step, which limits the computational time wherever this is feasible. Detailed numerical implementation of the solution of the Saint Venant flow equations in SOBEK 1D/2D is given in the technical user manual of Verwey (2006).

Generally the damages by flooding are classified as damages which can be quantified as monetary losses (tangible) and the damages which cannot be evaluated quantitatively in economic terms (intangible). These damages may be direct or indirect depending upon the contact to the flooding.

Flood damage estimation methodologies are applied worldwide (Dutta et al., 2003). For example, in the United Kingdom a standard approach to flood damage assessment is used (developed in the mid 1970s). Since then continually refined, this approach is mandatory for local authorities and agencies wanting central government assistance with flood mitigation measures. In United States, U.S. Army Corps of Engineers (USACE) has developed its own guidelines for urban flood damage measurement (USACE, 1988). The method is based on the US Water Resources Council's 1983 publication on 'Principles and Guidelines for Water and Related Land Resources Implementation Studies'. The approach adopted in the method is very comprehensive for estimation of damage to urban buildings and to agriculture. In Australia, authorities considered that is no standard approach and it is a little attempt to achieve standard approach. Flood damage estimation methodologies are applied as well in many countries in Europe (Forster et al., 2008). These approaches are useful in conducting cost-benefit analyses of the economic feasibility of flood control measures.

This paper uses the Forster et al., 2008, approach where the expected damage (ED) on agriculture was calculated using the following equation, which is modified from Forster et al. (2008).

ED = MV*Y*A*DI, where ED – estimation damage; MV – market value; Y – yield per unit area; A – area of cultivation; DI – damage impact factor.

The number of houses in the inundated area was calculated using the information on population density and average number of family member per household.

NH = IA*PD/FM; where NH – number of houses in inundated area; IA – inundated area; PD – population density; FM – average number of family per household.

In order to estimate the flood damage, the estimation of some flood parameters are needed: flow velocity, depth and duration at any given point, proper classification of damage categories considering nature of damage, establishment of relationship between flood parameters and damage for different damage categories.

3.2.1. *Flood inundation modelling*

In order to build the 1D/2D hydrodynamic model of the Budalangi River, in SOBEK, available topographical information from the Shuttle Radar Topography Mission (SRTM) at a resolution of 90 m by 90 m and sparse cross-section data were used. Hydrograph variations at the upstream boundaries of the model were provided by a calibrated hydrological SWAT model of the Nzoia catchment. Recorded water levels for Lake Victoria were used as downstream boundary conditions. The SWAT model used to provide the upstream boundary condition was the one originally built and described by Githui et al. (2008) and recalibrated. The 1:200 years design flood determined by SWAT was routed downstream by the hydrodynamic

SOBEK model and inundation extents were drawn. A 1 in 200 year return flood was recorded on Nzoia River on November 2008, and therefore the inundation extent produced by the model was compared with available aerial information captured by to the Advanced Land Imager (ALI) on NASA's earth observing-1 satellite on the 13th November 2008.

The results of the model, at the moment of the largest flood extent, for the 1:200 return flood period are represented in Fig. 3.

3.2.2. Flood damage evaluation

Many flood damage assessment methods have been developed since 1945 (White, 1945).

However, quantifying the expected flood damage is very difficult because the impact of a flood is a function of many physical and behavioural factors. For the purposes of this paper, flood damage was assumed to be related only to the flood depth.

The Budalangi region is a poorly developed rural area whose main industry is agriculture. Consequently the main expected damages were anticipated to be on the agricultural sector and were calculated based on a formula developed by Forster et al. (2008). The main cash crops in the area are known to be sugarcane, maize and rice. These crops were used, with readily available yield and expected local market values, to calculate the potential losses due to floods as a result of the 200 year return period event. In addition, loss of property and the affected population were included in the damage estimation, however it is recognised that in excluding the calculation of damage in relation to velocity this estimation is significantly simplified.

3.3. Assessing flood vulnerability of Budalangi using a parametric method

As mentioned above the parametric method used in this study is the one developed by Balica et al. (2009), which consists in determining a Flood Vulnerability Index (FVI), based on four components of flood vulnerability: social, economic, environmental and physical and their interactions, which can affect the possible short term and long term damages.

The four components of the flood vulnerability have been linked with the factors of vulnerability: exposure, susceptibility and resilience (Bosher et al., 2007; Penning-Rowsell and Chatterton, 1977).

The conceptual FVI equation is:

$$\mathrm{FVI} = \frac{E^{*}S}{R}, \tag{2}$$

where E – exposure, S – susceptibility and R – resilience.

The indicators belonging to exposure and susceptibility increase the Flood Vulnerability Index therefore they are placed in the nominator; however the indicators belonging to resilience decrease the FVI, this is why they are placed in the denominator (Dinh et al., 2012).

The application of this formula for each component leads to four distinct FVI indices; $\mathrm{FVI}_{\mathrm{Social}}$, $\mathrm{FVI}_{\mathrm{Economic}}$, $\mathrm{FVI}_{\mathrm{Environmental}}$ and $\mathrm{FVI}_{\mathrm{physical}}$, which aggregates into:

$$\text{Total FVI} \frac{\left(\frac{E^{*}S}{R}\right)_{\mathrm{Social}} + \left(\frac{E^{*}S}{R}\right)_{\mathrm{Economic}} + \left(\frac{E^{*}S}{R}\right)_{\mathrm{Environmental}} + \left(\frac{E^{*}S}{R}\right)_{\mathrm{Physical}}}{4} \tag{3}$$

The exposure can be understood as the intangible and material goods that are present at the location where floods can occur, such as: loss of photographs and negatives, loss of life, delays in formal education (Penning-Rowsell et al., 2005). The susceptibility relates to system characteristics, including the social context of flood damage formation (Begum et al., 2007) and can be i.e. poverty, people with special needs, education, level of trust. Susceptibility is defined as the extent to which elements at risk (Messner and Meyer, 2006) within the system are exposed, which influences the chance of being harmed at times of hazardous floods. Resilience to flood damages can be considered only in places with past events, since the main focus is on the experiences encountered during and after floods (Cutter, 1996; Cutter et al., 2003; Pelling, 2003; Walker et al., 2004; Turner II et al., 2003). Resilience describes the ability of a system to preserve its basic roles and structures in a time of distress and disturbance. Indicators showing resilience are flood insurance, amount of investment, dikes and levees, storage capacities, etc.

There are in total 29 indicators identified to contribute to Eq (3), each with their own unit of measure. Some indicators are not always used while evaluating the FVI of a region. They are evaluated in each case and the most representative are used for the FVI. A comprehensive description of such indicators in case of floods in the Mekong delta can be found in Dinh et al. (2012).

After identifying the indicators, in order to use them in Eq. (3) they need to be normalised using a predefined minimum and maximum. In general classical proportional normalization is used, which keeps the relative ratios in the normalized values of the indicators as they were before normalization. The indicators become dimensionless, but still keep their proportion.

The FVI of each of the social, economic, environmental and physical component is computed using Eq. (1). The results of each FVI component (social, economic, environmental and physical) are summed up in Eq. (3).

The FVI methodology does not require researchers to judge the relative importance of different components, i.e. they do not need to develop arbitrary weights for the indicators.

The Eq. (1) links the values of all indicators to flood vulnerability components and factors (exposure, susceptibility and resilience), without weighting, as suggested by Cendrero and Fischer (1997). This is done because of different number of rating judgements which "lie behind combined weights", or interpolating. The same approach of assigning no weights was used by Peduzzi et al. (2001), the Global Risk

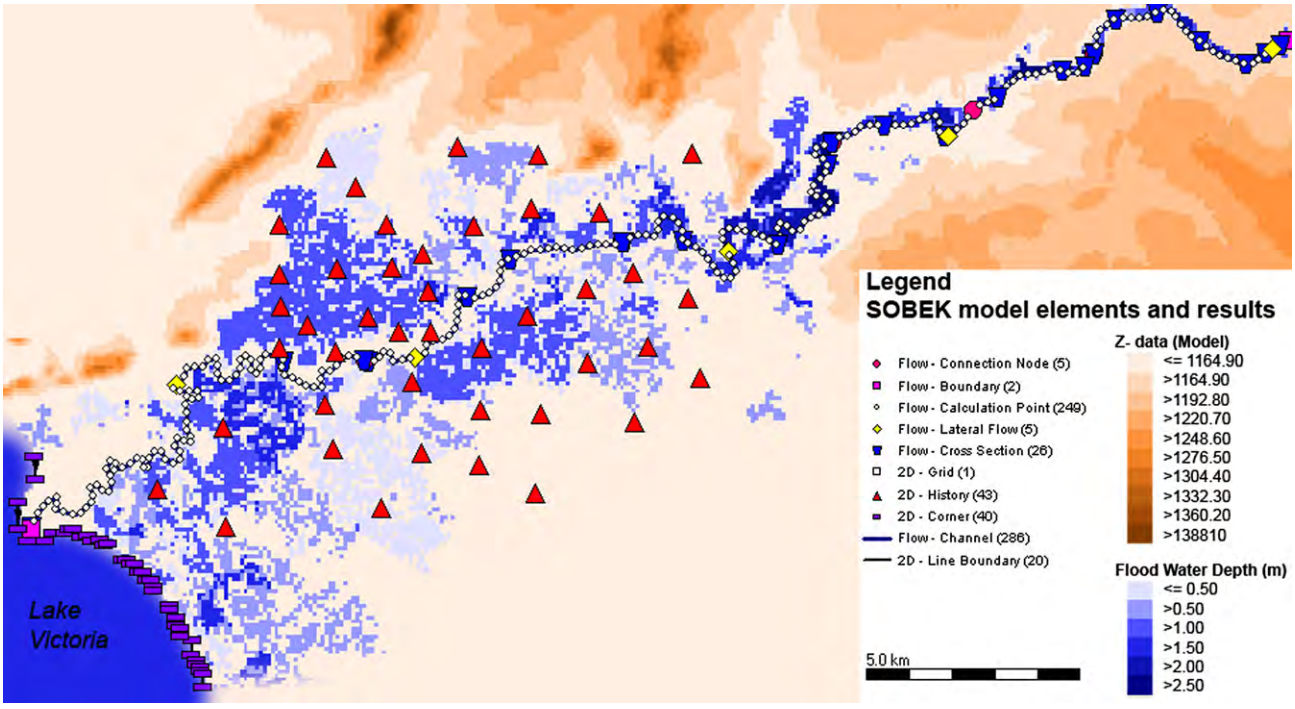

Fig. 3. Lower Nzoia flood inundation extent 1:200 year prediction.

and Vulnerability Index – Trends per Year, GRAVITY, by Briguglio (2003) in the Economic Vulnerability Index and Rygel et al. (2006).

The main issue while computing the FVI is actually to determine these indicators. There are different sources for determining the values of the indicators, and these are in general statistical data stored by environmental agencies, water boards, UN overviews and annual data from city halls.

4. Results obtained when applying the two approaches

4.1. Physically based modelling approach

The SOBEK simulation of the 1:200 year event results were water depths and inundation extents, as can be seen in Fig. 3. The model is able to produce velocities of flow during an inundation event as well; however these velocities were not considered in the estimation of the damages and therefore not reported herein.

The maximum inundation extent was checked with an available satellite image on 13 November 2008. The obtained maximum inundation extent from the model was of 12.61 km^2, which represents 97% of the inundation extent of the satellite image. Due to lack of data in the area, it is considered that this is good for the calibration of the model.

In order to determine the impact of flood and to evaluate the damages water depths obtained from the model were analysed. The obtained water depths were overall less than 2 m (95% of the inundated area), and only 5% bigger than 2 m in the upstream of the river. The main water depth is less than 0.5 m for 30% of the inundated area; 0.5 m for 20% of the inundated area, between 1 m and 1.5 m for 35% of the inundated area; and 1.5–2 m for 10% of the inundated area.

Based on the results from the hydrodynamic model, damage in the Budalangi area was computed using Forster et al. (2008) method and damage functions (Duggal and Soni, 2005).

In the Budalangi area the expected potential damages of 1.54M Euros (±80,000 Euros were calculated for the event of 1:200 year return.

4.2. Parametric approach

The FVI methodology was applied to Budalangi Settlement, the results can be seen in Table 2. Budalangi vulnerability in the social and economic components is higher than the environmental and physical component (1.00 means the highest vulnerability, see Table 1 for Flood Vulnerability Index designations).

The incorporation of flood vulnerability designations is probably the most difficult of all variables to include in the vulnerability index. There are problems involved in deciding how to rank vulnerability zones; but since the purpose of the FVI is to assess vulnerability in relation to flood vulnerability components and indicators, it was decided to rank the designation zones on the basis of standardised vulnerability indices results, between 0 and 1.

Flood vulnerability designations are assigned based on vulnerability potential in the event of flooding. A very high vulnerability designation is assigned if there is very high potential for loss of life and/or extreme economic loss based on vulnerability indicators, i.e. low amount of investment in counter measures or very slow recovery. A high vulnerability designation is assigned if there is a high potential for loss of life but still high economic loss. A medium vulnerability designation is assigned if there is a medium potential for loss of life but an appreciable economic loss, the area can recover in months and the amount of investment in counter measures is enough to maintain the existing structural measures. A low flood vulnerability designation is assigned if there is a small but still existing potential for loss of life and the economic loss is minor. Lastly, a very low flood vulnerability designation is assigned if there is a vanishingly small potential for loss of life and the economic loss can be minor or even if flood insurances apply.

Table 1
Flood vulnerability designations.

Designation	Index value	Description
Very small vulnerability to floods	<0.01	Very small Vulnerability to floods, the area recover fast, flood insurances exist, Amount of investment in the area is high
Small vulnerability to floods	0.01–0.25	Social, economic, environmental and physical the area can once in a while experience floods, the area is vulnerable to floods and the recovery process is fast due to the high resilience measures, high budget, on the other hand if the area is less developed economic, even if a flood occurs the damages are not high, so small vulnerability to floods
Vulnerable to floods	0.25–0.50	Social, economic, environmental and physical the area is vulnerable to floods, the area can recover in months average resilience process, amount of investments is enough
High Vulnerability to floods	0.50–0.75	Social, economic, environmental and physical the area is vulnerable to floods, recovery process is very slow, low resilience, no institutional organizations
Very high vulnerability to floods	0.75–1	Social, economic, environmental and physical the area is very vulnerable to floods, the recovery process very slow. The area would recover in years. Budget is scarce.

The data for the Budalangi area consulted to gather the indicators are: UNDP: United Nations Development Programme (HDI, child mortality, inequality); INTUTE: a web-site which provides social data for education and science research (population density, unemployment, disabled people); the World Fact-Book, a database developed by the CIA with basic information on all the countries in the world (communication penetration rate, past experience); UNEDRA: University Network for Disaster Risk Reduction in Africa; Nzoia River Basin Management Initiative a public private partnership between Water Management Resource Authority and Mumia Sugar, Pan Paper and Nzoia Sugar Company (land use, flood insurance, shelters, closeness to river); DEFRA – Department for Environment, Food and Rural Affairs economic and statistical database at no cost charge (urban growth, population growth, amount of investment, dikes-levees, storage capacity); WKCDD & FMP, Western Kenya Community Driven Development & Flood Mitigation Project (river discharge, rainfall, evaporation); Western Water Board, Kenya (drainage, topography, industries, evacuation roads).

Socially, the Budalangi area has very high vulnerability to floods, since has high population density, high child mortality rate, and a large affected population due to floods. The study also shows that the region has few shelters (0.6/km^2), no warning systems, no evacuation roads (no asphalted road), and only limited emergency services.

Table 2
Budalangi FVI results.

Budalangi Flood Vulnerability Index		
FVI components	FVI values	FVI designation
FVI_{Social}	0.768	Very high vulnerability to floods
$FVI_{Economic}$	0.521	High vulnerability to floods
$FVI_{Environmental}$	0.314	Vulnerable to floods
$FVI_{Physical}$	0.341	Vulnerable to floods
FVI Total	0.490	Vulnerable to High vulnerability to floods

Economically the region is high vulnerable to floods since the area has low exposure to floods as the main economic activity is agriculture. The Human Development Index is low, and the area is not covered by flood insurance. Budalangi has few industries, the investment levels and a recovery process take long to recover after a flood event.

Environmentally, the Budalangi settlement is vulnerable to floods. The environmental component includes indicators which refer to damage to the environment caused by flood events or manmade interferences which could increase the vulnerability of certain areas. But activities like industrialisation, agriculture, urbanisation, deforestation, can increase the flood vulnerability, which may also create even more environmental damages.

When examining the physical vulnerability, the Budalangi area has very low slope and the settlement area is in contact with the river all along the length of the river so the exposure of Budalangi is high and has low resilience with little or no installed storage capacity.

Overall, the area following the designations of FVI is high vulnerable to floods, the recovery process is slow, the area has low resilience and no institutional organizations.

5. Discussion (comparison – analysis and discussion of the approaches)

5.1. The physically based modelling approach

Physically based models have the advantage that they calculate the solution of a complicated and coupled set of equations that describe the phenomena of river flow and flooding. These models are dependent on physical knowledge that they incorporate into the equations and associated parameters. A key element for a good physically based model is the minimum of historical data that they need to determine the values for the parameters included in the physically based equations. Often, historical data is not available, in particular in areas of weak infrastructure, and this would make physically based models unusable in certain areas.

The advantage in using physically based models is their high capability for prognosis and forecasting, and their disadvantage is the high input data demand. In the past computational demand was a big disadvantage, but nowadays with the development of cloud and cluster computing capabilities over the internet, this disadvantage is reduced. However this is only true in case of larger, better-funded organisations that have good computer power to create cluster of computers, and not yet true for small consultancy companies or water boards who cannot dedicate cluster of computers for a specific modelling task. Due to the high computation resources demands, in case of 2D and 3D models, the calibration of physically or semi-physically based models can still be a tremendous effort.

In the present study the data on flooding was scarce, however the 2D physically based model was able to predict well the extent of flood, which shows that even in an ungauged catchment if the model is properly build, confidence in the construction of such a model does not require calibration (Cunge et al., 1980) and the results are good for design. A model based on the physics of the phenomena can be used to produce synthetic data to be used with a simple forecasting model (Van Steenbergen et al., 2012).

One of the important tasks of the decision makers in flood situations is not only to take management decision but also to properly disseminate knowledge to involved stakeholders, including the general public. The objectives of knowledge dissemination is to offer simple and clear information, which can prepare the public for the future and also can actively involve the stakeholders in flood management planning. The information should be delivered in relevant spatial and temporal scales. A physically based model has the advantage that can offer all types of information on a very fine spatial resolution, at a level of a street, or a house, in a familiar and easily recognisable user interface. It is very important that the decision makers use thoroughly verified results, rather than results characterised by uncertainties, because the stakeholders and the public are taking often quick evacuation measures based on such information.

5.2. The parametric approach

The FVI approach regarding the *information on the mechanism and cause of flooding* has some limitations, what is given from this approach are the indicators values for river discharge, topography, closeness to the river, the amount of rainfall, dikes and levees. Considering these indicators the FVI approach can only evaluate the flood vulnerability, cannot tell the extent of flooding nor the expected inundation area through the physical and environmental component. The application of this approach takes less preparation time than physically-based model construction, calibration and simulation.

The FVI approach regarding the *information on the health and safety implications to the affected population* is well designed; the approach shows through the social vulnerability indicators the exact population exposed to floods, the ones which are susceptible (youngest and eldest), if these people are aware and prepare, if they have and know how to interpret a warning system, which of the roads can act as an evacuation road. The social Flood Vulnerability Index expresses whether the population of that specific area has experienced floods, the number of people working in the emergency service and the number and locations of shelters in the area. The social FVI provides a greater understanding of how people might be affected, which can feed into emergency services and evacuation strategy development.

The FVI approach regarding the *information on the economic damages and losses to the affected areas* gives basic damage estimation. The economic component is related to income or issues which are inherent to economics that are predisposed to be affected (Gallopin, 2006).

Many economic activities can be affected by flooding events, among them are agriculture, fisheries, navigation, power production, industries, etc. The loss of these activities can influence the economic prosperity of a community, region or a country. The FVI can assess the economic vulnerability using a single number, though this number cannot evaluate the exact damage and losses but instead the index shows the number of industries in the area and their closeness to the river and also the amount of investment in counter measures and the number of flood insurances in that specific area.

5.2.1. How easily the information of the FVI approach is communicated?

From experts undertaking the study to the decision makers it can be said that the use of the FVI approach improves the decision-making process by identifying the vulnerability of flood prone areas. The FVI approach will direct decision-makers to a simplified usage and simpler understanding of the vulnerability; the FVI approach can be seen as a tool for decision making to direct investments to the most appropriate sectors and also to help in the decision-making process relating to flood defence, policies, measures and activities. The FVI approach allows, irrespective of uncertainties, relative comparisons to be made between case studies. While a level of uncertainty is inherent in FVI, the use of it in operational flood management is highly relevant for policy and decision makers in terms of starting adaptation plans. It offers

a more transparent means of establishing such priorities, which inevitably are considered as highly political decisions. It may also be considered as a means to steer flood management policy in a more sustainable direction. However, as individual information is lost in the aggregation process, it needs to be retrieved by a more in-depth analysis of each process in order to design policies and their implementation.

5.2.2. From decision maker to the public

Hence it is useful to have an *easy-to-apply* and *communicating* instrument that can help give an overview of the main points by having one single and comparable number, the FVI. The FVI is necessary, but not sufficient, for decision making and therefore should be used in combination with other decision-making tools. This should specifically include participatory methods with the population of areas identified as vulnerable and should also include a team of multidisciplinary thematic specialists and knowledgeable societal representatives and those with expert judgements.

6. Conclusions

The two approaches, modelling and parametric, have been applied to a data-scarce area - the Budalangi settlement. Examining the **approaches** in the context of this study leads to the following conclusions:

1. FVI does not assess flood risk directly, but does contribute to assessing flood risk. Vulnerability takes a step further and covers other aspects, such as: social aspects, environmental damage and infrastructure resilience.
2. The deterministic approach has a better science base, but limited evaluation of vulnerability;
3. FVI gives a wider evaluation, but is less rigorous. Therefore FVI is useful in a larger-scale vulnerability assessment, but a deterministic approach is better for more focused studies. In fact FVI could be used to decide where a deterministic model is necessary.

The Flood Vulnerability Index as analysed in the research provides a quick, reliable evaluation of flood vulnerability and in fact is the only method for assessing the vulnerability to flooding of a particular geographical area. The fact that indicators are used, allows for comparison of flood vulnerability in different areas as well as the identification of which indicators can determine the relative level of flood vulnerability. FVI can measure trends in the changing natural and human environments, helping identify and monitor priorities for action. These features, alongside the ability to identify the root causes of increased vulnerability, provide key information at a strategic level for flood risk planning and management. However the results would provide neither sufficient information nor the required level of detail for input into engineering designs or project level decisions.

FVI can provide an insight into the most vulnerable locations. It can analyse the complex interrelation among a number of varied indicators and their combined effect in reducing or increasing flood vulnerability in a specified location. It is very useful when there is a large level of uncertainty and decision makers are faced with a wide array of possible actions that could be taken in different scenarios, in this case the FVI can present readily understood and readily communicated results that can decision-makers in identifying the most effective measures to be taken. In this way the proposed measures can be prioritised for areas that are at greatest risk. Uncertainty is not removed, but is integrated into the assessment. On the other hand this complexity is also a negative point, since it takes a long time and good knowledge of the area and the system behind the FVI to be able to implement it.

As all with models, this FVI model is a simplification of reality and its application should be compensated for with thorough knowledge and expert-based analysis. The difficulties that the quantification of social indicators poses to the calculation may constitute a considerable weakness of the model. The FVI is a useful tool to identify the most vulnerable elements of the water resource system and safety chain components (Pro-action, prevention, preparation, response and recovery).

Obviously such a parametric model is limited by the accuracy and availability of good datasets. A number of the indicators are very hard to quantify especially when it comes to the social indicators. On the other hand, such a model can give a simplified way of characterising what in reality is a very complex system. Such results will help to give an indication of whether a system is resilient, susceptible or exposed to flooding risks and help identify which measures would reap the best return on investment under a changing climate and population and development expansion. The important point is that such a model is used as one tool among others within the whole process of deciding on a roadmap for flood assessment.

References

Adger, W.N., 2006. Vulnerability. Global Environmental Change 16, 268–281.

Alexander, D., 1993. Natural Disasters. Chapman & Hall, New York.

Balica, S.F., Douben, N., Wright, N.G., 2009. Flood Vulnerability Indices at varying spatial scales. Water Science and Technology Journal 60 (10), 2571–2580.

Balica, S.F., Wright, N.G., 2010. Reducing the complexity of Flood Vulnerability Index. Environmental Hazard Journal 9 (4), 321–339.

Barredo, J.I., de Roo, A., Lavalle, C., 2007. Flood risk mapping at European scale. Water Science and Technology 56 (4), 11–17.

Begum, S., Stive, M.J.F., Hall, J.W., 2007. Flood Risk Management in Europe: Innovation in Policy and Practice. Springer, London, 237 pp.

Blaikie, P., Cannon, T., Davis, I., Wisner, B., 1994. At Risk – Natural Hazards, People's Vulnerability and Disasters. Routledge, London.

Bosher, L., Dainty, A., Carrillo, P., Glass, J., 2007. Built-in resilience to disasters: a pre-emptive approach, Engineering. Construction and Architectural Management 14 (5), 434–446.

Briguglio, L., 2003. Methodological and Practical Considerations for Constructing Socio Economic Indicators to Evaluate Disaster Risk. Institute of Environmental Studies, University of Colombia, Manizales, Colombia. Programme on Information and Indicators for Risk Management, IADB-ECLAC-IDEA.

Cendrero, A., Fischer, D.W., 1997. A procedure for assessing the environmental quality of coastal areas for planning and management. Journal of Coastal Research 13, 732–744.

Connor, R.F., Hiroki, K., 2005. Development of a method for assessing flood vulnerability. Water Science and Technology 51 (5), 61–67.

Cred Crunch, 2012. Disaster Data. A Balanced Perspective, Issue 27, February 2012. Centre for Research on the Epidemiology and Disasters, Universite Catholique de Louvain. As seen on 8th March 2012 on. http://www.cred.be/sites/default/files/CredCrunch27.pdf.

Cunge, J.A., Holly Jr., F.M., Verwey, A., 1980. Practical Aspects of Computational River Hydraulics. Pitman, New York.

Cutter, S.L., 1996. Vulnerability to environmental hazards. Progress in Human Geography 20 (4), 529–539.

Cutter, S.L., Boruff, B.J., Shirley, W.L., 2003. Social vulnerability to environmental hazards. Social Science Quarterly 84 (2), 242–261.

Dinh, Q., Balica, S., Popescu, I., Jonoski, A., 2012. Climate change impact on flood hazard, vulnerability and risk of the Long Xuyen Quadrangle in the Mekong Delta. International Journal of River Basin Management 10 (1), 103–120.

Duggal, K.N., Soni, J.P., 2005. Elements of Water Resources Engineering. New Age International Publishers, New Delhi, 413 pp.

Dutta, D., Srikantha, H., Katumi, M., 2003. A mathematical model for flood loss estimation. Journal of Hydrology 277, 24–49.

European Floods Directive, L 288/27, 2007. Official Journal of the European Union, Brussels.

Forster, S., Khuhlmann, B., Lindenschmidt, K.E., Bronstert, A., 2008. Assessing flood risk for rural detention area. Natural Hazards and Earth System Sciences 8, 311–322.

Gabor, T., Griffith, T.K., 1980. The assessment of community vulnerability to acute hazardous materials incidents. Journal of Hazardous Materials 8, 323–333.

Gallopin, G.C., 2006. Linkages between vulnerability, resilience, and adaptive capacity. Global Environmental Change 16, 293–303.

Gichamo, Z.,G., Popescu, I., Jonoski, A., Solomatine, D.P., 2012. River cross section extraction from ASTER global DEM for flood modeling. Journal of Environmental Modelling & Software 31 (5), 37–46.

Githui, F., Gitau, W., Mutuab, F., Bauwensa, W., 2008. Climate change impact on SWAT simulated streamflow in western Kenya. International Journal of Climatology 29, 1823–1834.

Glynn, P.W., Heidelberger, P., Nicola, V., Shahabuddin, P., 1993. Efficient estimation of the mean time between failures in non-regenerative dependability models. In: Proceedings of the 1993 Winter Simulation Conference, pp. 311–316.

Hansson, K., Danielson, M., Ekenberg, L., 2008. A framework for evaluation of flood management strategies. Journal of Environmental Management 86 (3), 465–480.

Hartanto, I.M., Beevers, L., Popescu, I., Wright, N.G., 2011. Application of a coastal modelling code in fluvial environments. Environmental Modelling and Software 26 (12), 1685–1695.

IPCC, 2001. Climate Change 2001: the Scientific Basis. Cambridge University Press, Cambridge.

ISDR, 2004. International Strategy for Disaster Reduction. Living with Floods: UN Guidelines Offer Decision Makers Hope to Reduce Flood Losses. Available from: www.unisdr.org/archive/5353 (5.04.11.).

Jonkman, S.N., Vrijling, J.K., 2008. Loss of life due to floods. Journal of Flood Risk Management 1, 43–56.

Klein, R.J.T., Nicholls, R.J., Mimura, N., 1999. Coastal adaptation to climate change: can the IPCC Technical Guidelines be applied? Mitigation and Adaptation Strategies for Global Change 4, 51–64.

Knight, F.H., 1921. Risk, Uncertainty, and Profit. Hart, Schaffner & Marx; Houghton Mifflin Company, Boston, MA. Available online: http://www.econlib.org/library/Knight/knRUP1.html. As seen on 14th of December 2011.

Lewis, J., 1999. Development in Disaster-prone Places. Intermediate Technology Publications, London.

Little, R.J.A., Rubin, D.B., 1987. Statistical Analysis with Missing Data. J. Wiley & Sons, New York.

Little, R.J.A., Rubin, D.B., August 1983. On jointly Estimating Parameters and Missing Data by Maximizing the Complete-Data Likelihood. In: The American Statistician, vol. 37 (3). Published by American Statistical Association, pp. 218–220.

McCarthy, M.A., Possingham, H.P., Day, J., Tyre, A.J., 2001. Testing the accuracy of population viability analysis. Conservation Biology 15 (4), 1030–1038.

Merz, B., Thieken, A.H., Gocht, M., 2007. Flood risk mapping at the local scale: concepts and challenges. In: Begum, S., Stive, M.J.F., Hall, J.W. (Eds.), Flood Risk Management in Europe, 25. Advances in Natural and Technological Hazards Research. Springer, Netherlands, pp. 231–251.

Messner, F., Meyer, V., 2006. Flood damages, vulnerability and risk perception—challenges for flood damage research. In: Schanze, J., Zeman, E., Marsalek, J. (Eds.), Flood Risk Management: Hazards, Vulnerability and Mitigation Measures. Springer, pp. 149–167.

Newey, W.K., 1990. Semiparametric efficiency bounds. Journal of Applied Econometrics 5, 99–135.

NRBMI, nd. Nzoia River Basin Management Initiative, a Public Private Partnership between Water Resources Management Authority and Mumimas Sugar Company Limited/Nzoia Sugar Company Limited/PAN PAPER/Civil Society, Learning Institutions and Communities. http://www.unep.org/Training/downloads/PDFs/NRBMI_ small.pdf. Cited December 2009, van Hoey 2008.

Peduzzi, P., Dao, H., Herold, C., Rochette, D., Sanahuja, H., 2001. Feasibility Study Report – on Global Risk and Vulnerability Index –Trends per Year (GRAVITY). United Nations Development Programme Emergency Response Division UNDP/ERD, Geneva.

Pelling, M., 2003. The Vulnerability of Cities; Natural Disaster and Social Resilience. Earth Scan Publications, UK & USA.

Penning-Rowsell, E.C., Chatterton, J.B., 1977. The Benefits of Flood Alleviations: a Manual of Assessment Techniques (The Blue Manual). Gower Technical Press, Aldershot, UK.

Penning-Rowsell, E., Floyd, P., Ramsbottom, D., Surendran, S., 2005. Estimating injury and loss of life in floods: a deterministic framework. Natural Hazards 36, 43–64.

Plate Erich, J., 2002. Flood risk and flood management. Journal of Hydrology 267, 2–11.

Pratt, C., Kaly, U., Mitchell, J., 2004. How to Use the Environmental Vulnerability Index, UNEP/SOPAC South Pacific Applied Geo-science Commission. Technical Report 383. www.vulnerabilityindex.net/Files/EVI%20Manual.pdf (accessed 17.09.11.).

Rygel, L., O'sullivan, D., Yarnal, B., 2006. A method for constructing a social vulnerability index: an application to hurricane storm surges in a developed country. Mitigation and Adaptation Strategies for Global Change 11, 741–764.

Sayers, P., Hall, J., Dawson, R., Rosu, C., Chatterton, J., Deakin, R., 29 September 2011. Risk Assessment of Flood and Coastal Defenses for Strategic Planning (rasp) – a High Level Methodology. Wallingford, as seen on. http://www.raspproject.net/RASP_defra2002_Paper_Final.pdf.

Schanze, J., 2006. Flood risk management – a basic framework, flood risk management: hazards, vulnerability and mitigation measures NATO science Series: IV. Earth and Environmental Sciences 67 (Part 1), 1–20.

Serrat, C., Gomez, G., 2001. Estimating the Stratified Survival with Missing Covariates: I A Semiparametric Approach. DR 2001/12. Departament d'Estadística i Investigació Operativa. UPC, 69 pp. (Chapter 4).

Slovic, P., 1987. Perception of risk. Science 236, 280–285.

Smith, K., 2004. Environmental Hazards: Assessing Risk and Reducing Disaster. Routledge, London.

Sullivan, C.A., Meigh, J., 2003. Using the Climate Vulnerability Index to assess vulnerability to climate variations. Water Policy and Management. CEH Wallingford, as seen on: http://www.ceh.ac.uk/sections/ph/ClimateVulnerabilityIndex.html (12.12.11.).

Turner, B.L., Kasperson, R.E., Matson, P.A., McCarthy, J.J., Corell, R.W., Christensen, L., Eckley, N., Kasperson, J.X., Luers, A., Martello, M.L., Polsky, C., Pulsipher, A., Schiller, A., 2003. A framework for vulnerability analysis in sustainability science. Proceedings of the National Academy of Sciences of the United States of America 100 (14), 8074–8079.

U.S. Army Corps of Engineers, 1988. Master Reservoir Regulation Manual: Lake Red Rock. USACE, Rock Island District.

UNEP, 2004. Manual: How to Use the Environmental Vulnerability Index (EVI). As see on. http://www.vulnerabilityindex.net/EVI_Calculator.htm (27.09.10.).

United Nations, 1982. Proceedings of the Seminars on Flood Vulnerability Analysis and on the Principles of Floodplain Management for Flood Loss Prevention, September. Bangkok.

van der Veen, A., Logtmeijer, C., 2005. Economic hotspots: visualizing vulnerability to flooding. Natural Hazards 36 (1–2), 65–80.

Van Steenbergen, N., Ronsyn, J., Willems, P., 2012. A non-parametric data-based approach for probabilistic flood forecasting in support of uncertainty communication. Environmental Modelling & Software 33, 92–105.

Verwey, A., 2006. Computation Hydraulics, WL Delft Hydraulics, Version 2006 – 1, UNESCO-IHE Lecture Notes.

Walker, B., Holling, C.S., Carpenter, S.R., Kinzig, A., 2004. Resilience, adaptability and transformability in social–ecological systems. Ecology and Society 9 (2), 5 (online) URL: http://www.ecologyandsociety.org/vol9/iss2/art5/.

White, G.F., 1945. Human Adjustments to Floods: a Geographical Approach to the Flood Problem in the United States Doctoral dissertation and Research Paper No. 29. Department of Geography, University of Chicago.

WMO/MWRMD/APFM, 2004. Strategy for Flood Management for Lake Victoria Basin, Kenya, Associate Programme on Flood Management. http://www.apfm.info/pdf/strategy_kenya_e.pdf.

Xia, Junqiang, Falconer, Roger A., Binliang, Lin, Tan, Guangming, 2011. Numerical assessment of flood hazard risk to people and vehicles in flash floods. Environmental Modelling & Software 26 (8), 987–998.

Cloud and cluster computing in uncertainty analysis of integrated flood models

V. Moya Quiroga, I. Popescu, D. P. Solomatine and L. Bociort

ABSTRACT

There is an increased awareness of the importance of flood management aimed at preventing human and material losses. A wide variety of numerical modelling tools have been developed in order to make decision-making more efficient, and to better target management actions. Hydroinformatics assumes the holistic integrated approach to managing the information propagating through models, and analysis of uncertainty propagation through models is an important part of such studies. Many popular approaches to uncertainty analysis typically involve various strategies of Monte Carlo sampling of uncertain variables and/or parameters and running a model a large number of times, so that in the case of complex river systems this procedure becomes very time-consuming. In this study the popular modelling systems HEC-HMS, HEC-RAS and Sobek1D2D were applied to modelling the hydraulics of the Timis–Bega basin in Romania. We considered the problem of studying how the flood inundation is influenced by uncertainties in water levels of the reservoirs in the catchment, and uncertainties in the digital elevation model (DEM) used in the 2D hydraulic model. For this we used cloud computing (Amazon Elastic Compute Cloud platform) and cluster computing on the basis of a number of office desktop computers, and were able to show their efficiency, leading to a considerable reduction of the required computer time for uncertainty analysis of complex models. The conducted experiments allowed us to associate probabilities to various areas prone to flooding. This study allows us to draw a conclusion that cloud and cluster computing offer an effective and efficient technology that makes uncertainty-aware modelling a practical possibility even when using complex models.

Key words | cloud and cluster computing, flood mapping, flood modelling, hydroinformatics, uncertainty analysis

V. Moya Quiroga
I. Popescu (corresponding author)
D. P. Solomatine
Integrated Water Systems and Governance Department,
UNESCO-IHE Institute for Water Education,
PO Box 3015,
2601 DA, Delft,
The Netherlands
E-mail: *i.popescu@unesco-ihe.org*

D. P. Solomatine
Water Resources Section,
Delft University of Technology,
PO Box 5048,
2601 DA, Delft,
The Netherlands

L. Bociort
Romanian National Water Authority,
Banat Branch,
Timisoara,
Romania

INTRODUCTION

Water is one of the most vital elements for human life, but it can also become a devastating force. Although the classical approach towards flood mitigation is to apply structural measures, engineers all over the world realised that such measures introduce new factors to be considered such as probable failure, new geographical interactions or performance. Moreover, new flood problems are appearing faster than structural measures can be implemented. Therefore, the concept of flood prevention is being replaced by the concept of flood management, which gives non-structural measures much higher weight (Schanze 2006; Soldano *et al.* 2007). Flood maps can be seen as the technical base for non-structural measures (Riccardi 1997). Both the European Union (EU) and United States Flood Emergency Management Agency (US FEMA) assessed the importance of developing flood maps (Federal Emergency Management Agency 2003; Aldescu 2008). Such maps can only be developed by modelling flood events using hydraulic and hydrologic models.

There are different approaches towards flood modelling, from simple geographic information system (GIS) techniques combined with Manning's equation (Herold &

doi: 10.2166/hydro.2012.017

Mouton 2006), to advanced numerical models in 1D (Ferreira 2004; Knebl *et al*. 2005; Smemoe *et al*. 2007), quasi-2D (Garcia *et al*. 2007; Lindenschmidt *et al*. 2008; Soumendra *et al*. 2010), integrated 1D2D, 2D (Cobby *et al*. 2003; Tarrant *et al*. 2005; Neal *et al*. 2010) or 3D (Kang & Choi 2005; Wilson *et al*. 2006). Each approach has its own advantages and disadvantages. For instance, while 1D models may not represent the flood pattern well, potentially more accurate 2D and 3D models require more data and computational resources, which will demand a longer running time of such a model. Hence there is a need to overcome the running time barrier that restricts the use of 2D and 3D models and one way to decrease this time is to improve the way computation is done by making use of supercomputers (which is still rarely done in engineering practice), or by distributing computations across computers arranged in clusters, or employing grid and cloud computing.

Apart from the necessity to speed up the models, additional demand on computing power comes from the need to perform uncertainty analysis of models and model chains. Such an analysis is an important part of any modelling study, and there are hydroinformatics tools to support it. Usually, input and parametric uncertainties are considered; certain descriptors of uncertainty are assumed (often by prior probability density functions, or fuzzy descriptors), and then 'propagated' through a model leading to uncertain outputs. Such analysis is a must if probabilistic flood maps are to be built. Uncertainty is typically treated as an aleatoric one (i.e. associated with randomness) and uses probabilistic descriptors. Uncertainty analysis usually involves using various versions of a Monte Carlo approach when parameters or inputs are sampled from the assumed distributions, and a model is run a large number of times. In the case of complex river systems, this procedure becomes time-consuming (Macdonald & Strachan 2001). Therefore, performing uncertainty analysis of a 2D flood model prompts for computational power, so cloud computing and cluster computing might be helpful tools in this respect. The present study shows the potential of using cloud and cluster computing when analysing uncertainty of complex hydraulic systems by running multiple simulations in parallel.

Cluster computing was successfully applied in different fields such as biological sciences (Boukerche *et al*. 2007), task scheduling (Lin *et al*. 2006), computer-aided engineering (Maccyzk & Janusz 2008), optimisation of drainage systems and water supply (Damas *et al*. 2001). Nevertheless, setting up a computer cluster may involve high initial cost; hence new alternatives such as grid computing and cloud computing have arisen. Nowadays these are commercial services deployed on the internet used only when needed, always updated and available at a low price, implementing thus the principle of Software as a Service, SaaS (see, for example, http://aws.amazon.com, https://saas.dhigroup.com).

Both grid computing and cloud computing are distributed computing services based on the possibilities offered by the internet. These are somewhat different paradigms and there is some confusion about the definition and differences between them. The literature identifies 11 technical differences, out of which the most important are type of application, virtualisation and access or control (Weinhardt *et al*. 2009). While cloud computing offers interactivity that makes its use easier, grid computing is usually oriented at experts and needs batches of jobs to be prepared in advance, which makes its usage more difficult compared to cloud computing.

Scientific applications of such services refer to biomedical applications, service-level-agreement-aware resources or high energy physics (Dejun *et al*. 2010; Rosenthal *et al*. 2010; Sevior *et al*. 2010). Application of cloud computing in water-related studies is just beginning with the first publications starting to appear. At UNESCO-IHE we experimented with the Amazon Elastic Compute Cloud platform for uncertainty analysis of a hydrological model (Moya *et al*. 2010) and multi-objective evolutionary optimisation of complex hydrodynamic wastewater models where we achieved seven times speedup when using 12 instances of virtual machines on the cloud (Xu *et al*. 2010). In the groundwater domain the GoGrid cloud computing internet service was tested to do parameter estimation of a model, showing that running a single processor required between 26 and 41 h while four machines on the cloud performed the same task in 9 h (Hunt *et al*. 2010).

The aim of the present paper is to demonstrate the applicability of distributed computing, in particular cloud and cluster, as a tool supporting uncertainty analysis for hydrologic and hydraulic studies, as well as the benefits of saving

computational time while using these tools. This paper is based on the models reported in Moya *et al.* (2010), which were developed further and used not only in cloud computing but also in cluster computing. The cloud and cluster computing component of the study is presented in detail, along with the comparison of the two technologies and their use in assessing the uncertainty of the land topography (digital elevation model (DEM)) in hydraulic modelling.

CASE STUDY

Cloud and cluster computing technology was applied to the Timis–Bega catchment in Romania (Figure 1). The province of Banat, where the catchment is located, is one of the most flood-vulnerable regions of Romania. The rivers Timis and Bega are of transboundary nature, flowing towards the Danube and the Tisa rivers, respectively, in the neighbouring country of Serbia. This catchment has a complex and varied topography starting from the Carpathian Mountains upstream, to a vast floodplain downstream.

The Timis river has a basin area of 5,573 km^2 and 240 km in length. It begins at the Semenic Mountains as a mountainous river with a slope of 20 m/km, and decreases to 20 cm/km at the most downstream section in the Romanian territory, at station Graniceri (Aldescu 2008). The Bega river springs in the Poiana Ruscai Mountains and flows a total length of 256 km. Both rivers are connected by a double connection, consisting of two canals, one of which during normal conditions diverts water into the Bega river in order to maximise the flow through the city of Timisoara, and the second one, which during times of high waters diverts water into the Timis river to prevent flooding of the city of Timisoara (Teodorescu 2008).

In April 2005 the region suffered an extreme flood event, due to low pressure systems from the Adriatic region and the Black Sea combined with an intense convective activity and snowmelt. Such an event flooded hundreds of square kilometres in this area (Popescu *et al.* 2010). This type of event showed the need to analyse different engineering measures to prevent human and material losses from future flood events.

Although there is a flood forecasting and warning system in operation which is based on empirical models, the 2005 event proved that more accurate models are required. This prompted the development of an integrated hydrological–hydraulic model of the Timis–Bega catchment to evaluate different flood scenarios considering flood extent, water depth and velocities. An integrated hydrological–hydraulic model had been developed in 2009 (Popescu *et al.* 2010) in the framework of a Dutch–Romanian project. In that project four mitigation measures, along with their

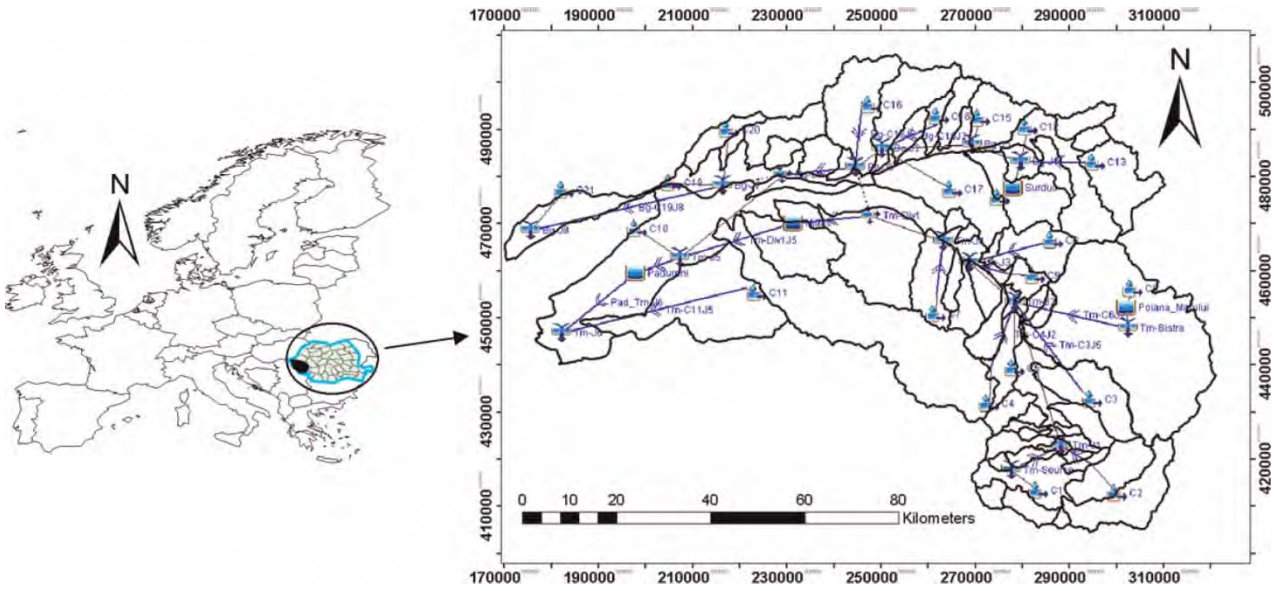

Figure 1 | Timis–Bega catchment location and HEC-HMS model.

flooding extent, have been tested. The results were promising; however, it was concluded that to perform a hydrological and hydraulic simulation over this catchment of 8,085 km^2 using a 2D flood model was a complex and time-consuming task.

Historically, the study presented herein originated from the need to extend the applicability of the hydrological–hydraulic model of the Timis–Bega catchment to generate probabilistic flood maps, and the need to analyse the possible uncertainties of the model. It was immediately concluded that the required Monte Carlo simulations would need serious computer resources; therefore the exploring of the possibilities of using cloud and cluster computing became the main challenge for this study.

EVALUATING UNCERTAINTY OF THE INTEGRATED MODELS

Integrated model of the Timis–Bega basin

As from 2009, the Romanian Water Board of Banat area uses, for simulation of flood routing, an integrated model consisting of three models: one for the rainfall–runoff component, for the most upstream part of the catchment; a second one for the middle part of the catchment consisting of a 1D river flow model; and a third one for the most downstream part of the catchment, a coupled river–floodplain 1D2D hydrodynamic model.

The Hydrologic Modeling System (HEC-HMS) software, developed by the US Army Corps of Engineers (USACE), was used to convert rainfall to runoff and to route the hydrographs from all the sub-catchments of the basin using a kinematic wave approach. The HEC-HMS component of the integrated model simulates precipitation, runoff, routing processes and man-made structures. It computes runoff by taking into account the relation between a number of components, such as runoff volume, direct runoff, base flow and channel flow. After that, the 1D hydrodynamics of the rivers Timis and Bega was simulated in the River Analysis System (HEC-RAS) software, also developed by USACE. HEC-RAS solves the 1D Saint-Venant equations and allows for computing the water levels at all time steps and all cross sections. The most refined part of the model contains an integrated 1D2D model in the downstream part of the catchment, for which the Sobek software developed by Deltares is used. The gauging stations at two locations, Remetea and Brod, are the starting points for a combined 1D2D simulation for the rivers Timis and Bega and the floodplain. Figure 2 schematises the geographic coverage of these models. This integrated model was extended by adding two existing reservoirs, Surduc and Poiana Marului, into HEC-HMS, and the processes of inter-model data transfer and model execution in the model chain was automated by writing special scripts.

The hydrologic model was calibrated for the 2003 flood event by comparing the calculated flood extent with the Moderate Resolution Imaging Spectroradiometer (MODIS) image corresponding to this event. The model was tested on the 2005 flood event.

The flow of information through the model chain is shown in Figure 3. The HEC-HMS resulting hydrographs at specific locations were used as inputs (i.e. boundary conditions) for the HEC-RAS component of the integrated model. In turn, the resulting hydrographs of the HEC-RAS model served as the boundary conditions for the Sobek model. The automation of data transfer and model execution was done by additional routines written in Delphi, VB.NET and Windows command line batch scripts. After the finalisation of the execution of one integrated model run for a particular scenario, the results were converted to text files and automatically uploaded to the Google Docs platform, by using the Java application 'google-docs-upload-1.3.1.jar'

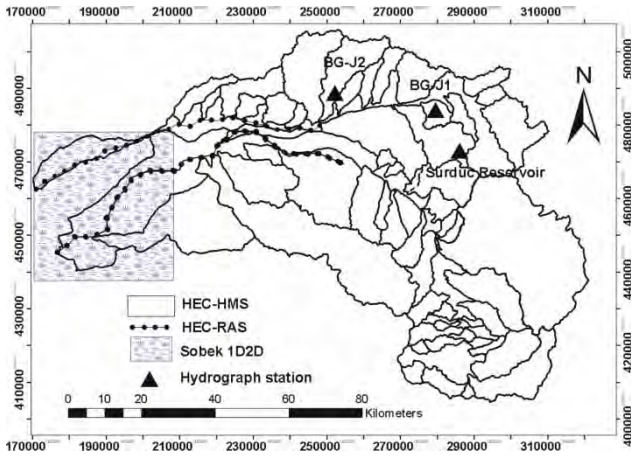

Figure 2 | Geographical coverage by hydrologic/hydraulic models.

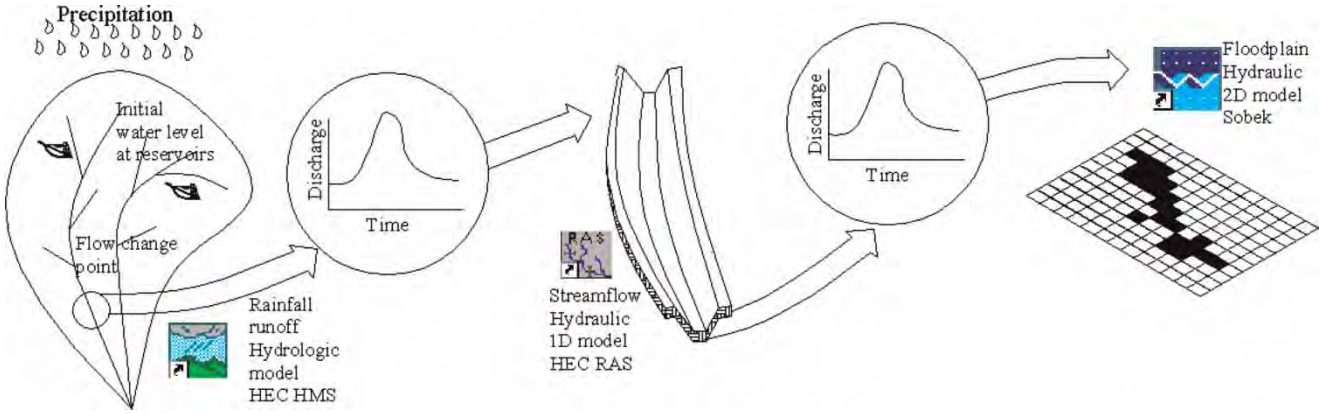

Figure 3 | Flow of information through models.

(http://docs.google.com/p/google-docs-upload/). This action facilitates the access by any program, including those on the cloud, to the results of each run in order to perform their uncertainty analysis. Uncertainty analysis focused on using various boundary conditions and samples of DEM.

Sources of uncertainty and methods for their analysis

Due to the complexity of hydrological and hydraulic models, analytical methods of studying their uncertainty are rarely used, and Monte Carlo analysis is the traditionally employed method. For example, if uncertainty of input data is considered, a sufficient number of samples of values for input variables are generated, a model is run for all of them, and the probabilistic properties of the model output are analysed. Such an approach typically requires hundreds or thousands of model runs and hence is computationally demanding. A widely popular version of Monte Carlo approach is the Generalised Likelihood Uncertainty Estimation (GLUE) (Beven & Freer 1996). With an objective of reducing computational load, several algorithms were designed recently, allowing for a reduced number of Monte Carlo simulations required for uncertainty analysis, such as Shuffled Complex Evolution Adaptive Metropolis (SCEM-UA) (Vrugt *et al.* 2003; Cutore *et al.* 2008), Differential Evolution Adaptive Metropolis (DREAM) (Vrugt *et al.* 2008), Adaptive Cluster Covering (ACCO). Shrestha *et al.* (2009) presented a framework to encapsulate the results of Monte Carlo simulations in a fast machine learning-based model like a neural network. However, the main purpose of the present study is to analyse the applicability of distributed computing and cloud computing as tools, which can easily provide the computational power required for such analysis. In this study it was decided to concentrate on showing the value of parallelisation and its practical implementation, so we deliberately used a standard Monte Carlo simulation experiments, leaving the exploration of other schemes and their relative performance to further studies.

For any flood model there are many possible sources of uncertainty related to input, output, initial conditions, model parameters and model structure, the most important being the influence of uncertainties related to discharge measurements, cross sections and DEM. In the present study, two sources of uncertainty were considered: (1) uncertainty brought by the insufficient knowledge of the initial water levels (storage) in the reservoirs Surduc and Poiana Marului, and (2) uncertainties brought about by inaccuracies in DEM. In this study we are not discussing if it is proper to describe the mentioned manifestations of epistemic uncertainty (lack of knowledge) by the probabilistic measures rather than using fuzzy logic; we just follow the route of many other researchers who used probabilistic variables for this purpose. The two identified sources of uncertainty are detailed below.

Initial water levels

The two reservoirs in the catchment, included in the analysis, have initial water levels which influence flow in the downstream part of the catchment. The Poiana Marului reservoir has a volume of 96,200,000 m^3 and a dam height of 125 m. Its main purpose is energy production. The

other reservoir considered is Surduc (a volume of 50,000,000 m^3 and a height of 36 m) mainly used for water supply. The initial water levels of the reservoirs were considered to be one of the sources of uncertainty. A simple Monte Carlo experiment with a limited number of runs was set up to analyse propagation of these uncertainties to the model output – flow at the Bega station downstream.

Digital elevation model (DEM)

Earth surface data is vital for flood modelling, but unfortunately such data is not always available and survey studies are both costly and also time-consuming. Several studies have been undertaken in order to have more realistic models of the DEM uncertainty on topographic parameters (Wechsler & Kroll 2006; Wechsler 2007). Nowadays, in the absence of accurate locally measured data, the Shuttle Radar Topography Mission (SRTM) data is often used; it provides global DEM covering all the planet between latitudes 60° N and 56° S (Carabajal & Harding 2006), with a resolution of 30 or 90 m. There have been several studies about the SRTM quality (Gorokhovich & Voustianiouk 2006; Sharma *et al.* 2010) and applicability towards morphometry, hydrology or remotely sensed water stages (Ludwig & Schneider 2005; Schumann *et al.* 2008; de Oliveira *et al.* 2010). These studies focused on macroscopic scale and compared results using SRTM DEM with results using other DEMs in particular cases. Although recently some research was carried out in order to extract from SRTM more detailed data such as cross sections (Pramanik *et al.* 2010), yet they do not consider the SRTM influence on the floodplain hydrodynamics. Hengl *et al.* (2010) state that Monte Carlo simulation can be used as an approach to analyse the error of stream networks extracted from DEM. Software tools for assessing uncertainties in environmental data are being developed; an example of such software is the Data Uncertainty Engine (DUE) that handles different types of data such as spatial vectors, spatial raster data or time series (Brown & Heuvelink 2007). We can state, however, that by the time of conducting the research presented herein we could not find references to studies explicitly analysing the uncertainty of integrated flood extent modelling due to uncertainties in DEMs obtained from SRTM data.

For the DEM-related uncertainty analysis, we used simple probabilistic descriptors of DEM uncertainty and employed standard Monte Carlo simulation to analyse the propagation of uncertainty through the model, in this case, a Sobek 1D2D model. Each cell was changed according to a uniform distribution within a range of 2 m, which is a reasonable value for the SRTM in that area (Rodriguez *et al.* 2005). The following sampling scheme was used: changes in the cell altitude were applied systematically, starting with the cells from the centre row. First, all the cells were changed and a new DEM generated. Then, all cells except the first row were changed, then all except the second row and the procedure was repeated (125 times) until the central row was reached. After that the procedure was repeated backwards, obtaining another 125 DEMs. To increase the total number of samples, the procedure was repeated obtaining a total of 500 samples of different DEMs (Figure 4).

One problem of the Monte Carlo analysis is that it is not possible to know *a priori* the needed number of simulations ensuring statistical reliability of the results. In the present study, we used the mean probability interval (MPI); it is defined by the 5 and 95% quantiles of the estimate of pdf of flood depth at some critical point.

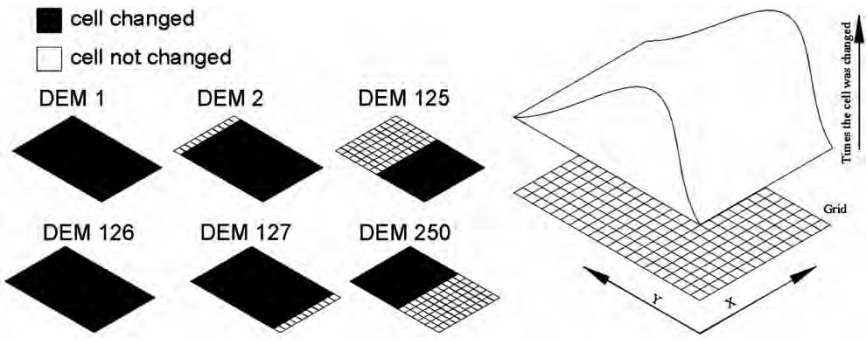

Figure 4 | Sampling DEM cells.

Cloud computing and cluster computing

Nowadays there are several cloud computing providers such as Amazon (http://aws.amazon.com), Rackspace (http://www.rackspace.com/), Verizon (http://www.verizonbusiness.com/), Joyent (http://www.joyent.com/) and GoGrid (http://www.gogrid.com/). In the present study the Amazon Elastic Compute Cloud (Amazon EC2) platform, launched in 2006, was used. The Amazon EC2 has the lowest price, as of 2010, and is the most popular and fast growing service of this kind. It provides a collection of computing services via stand-alone computers (virtual machines) of different capacity called instances. These instances are accessed via remote desktop. The fee for using these instances is around $0.15 per hour for one machine. Other companies develop tools to simplify the use of this service; examples are ElasticFox and S3Fox – extensions for the Firefox browser to enable cloud computing via Amazom EC2.

Amazon EC2 makes it possible to launch and manage server instances in the Amazon data centre using the available tools and application programming interfaces (APIs) from Amazon (Amazon Web Services (AWS), 2009). AWS provides seven types of virtual machines (called instances) that can run under different operating systems. Among these, the so-called small instance was selected. Although this is the cheapest instance, its 1.7 GHz and 160 GB storage make it powerful enough for many modelling applications. Zheng (2010) made an extensive research into the performance of the different instances, and by comparing speedup, performance and prices he found that the choice of the optimal instance can be posed as a multi-objective problem. For instance, if a small instance takes 1 h for some simulations it will charge 1 h of small instance, and if a medium instance takes 0.4 h for the same computation, due to the pricing policy it will charge 1 h of medium instance, which is more expensive than the small one. This can lead to an idea that a small instance is always better, and in this particular case it is. However, if we need two simulations, it will cost 2 h of small instance, while the medium instance will still cost just 1 h, since 0.8 h will be charged as 1 h, so the choice of an instance type to use could be different. Note that launching an instance with more than one core but using just one would mean wasting money.

An alternative to cloud computing is to use a cluster of PCs. A computer cluster can be defined as a group of computers linked through a local area network (LAN), so that they can work together and ensure higher performance. Such clusters may be built on the PCs solely dedicated to this purpose, or based on the standard office PCs, forming thus an 'office cluster'.

In order to demonstrate and test the possibilities of cloud and cluster computing for the considered case study two approaches have been tested. First, cloud computing was tested using the cloud service from Amazon Web Services. Due to licence limitations, we could not deploy Sobek software on remote PCs, so in this part of the study only a part of the integrated Timis–Bega model employing the free software (HEC-HMS and HEC-RAS) was used. Second, cluster computing was tested for the part of the Timis–Bega model, which uses the licensed Sobek software for which we had the LAN licence. A small office cluster of five computers was used which we considered to be enough for the proof of concept.

To create a virtual computational platform on the basis of Amazon Web Services, one has to go through a number of steps (Figure 5). First, an account has to be created to get access to the services provided by AWS. While accessing the EC2 service there is a list of the available instances. Within the available instances, one has to be chosen as the base one. The process of choosing an instance can be accomplished either with the normal AWS user interface, or with ElasticFox, an extension for the Firefox browser. The access to the selected instance is done via the Windows Remote Desktop. Since both HEC-HMS and HEC-RAS are licence-free software there was no problem with the installation and further launching of multiple instances. Then, all the needed software was downloaded and installed in the same way as it would have been done on any other computer. Besides, the entire folder with the respective data was uploaded into that instance. For the task of copying data the Firefox S3Fox application proved to be very useful, since it simplified the process to a simple drag-and-drop operation. A bundle service was used to save the instance with the software and data installed, so that next time repeating some steps is avoided.

Once the instance was created and saved, the task of performing multiple scenarios in parallel was relatively

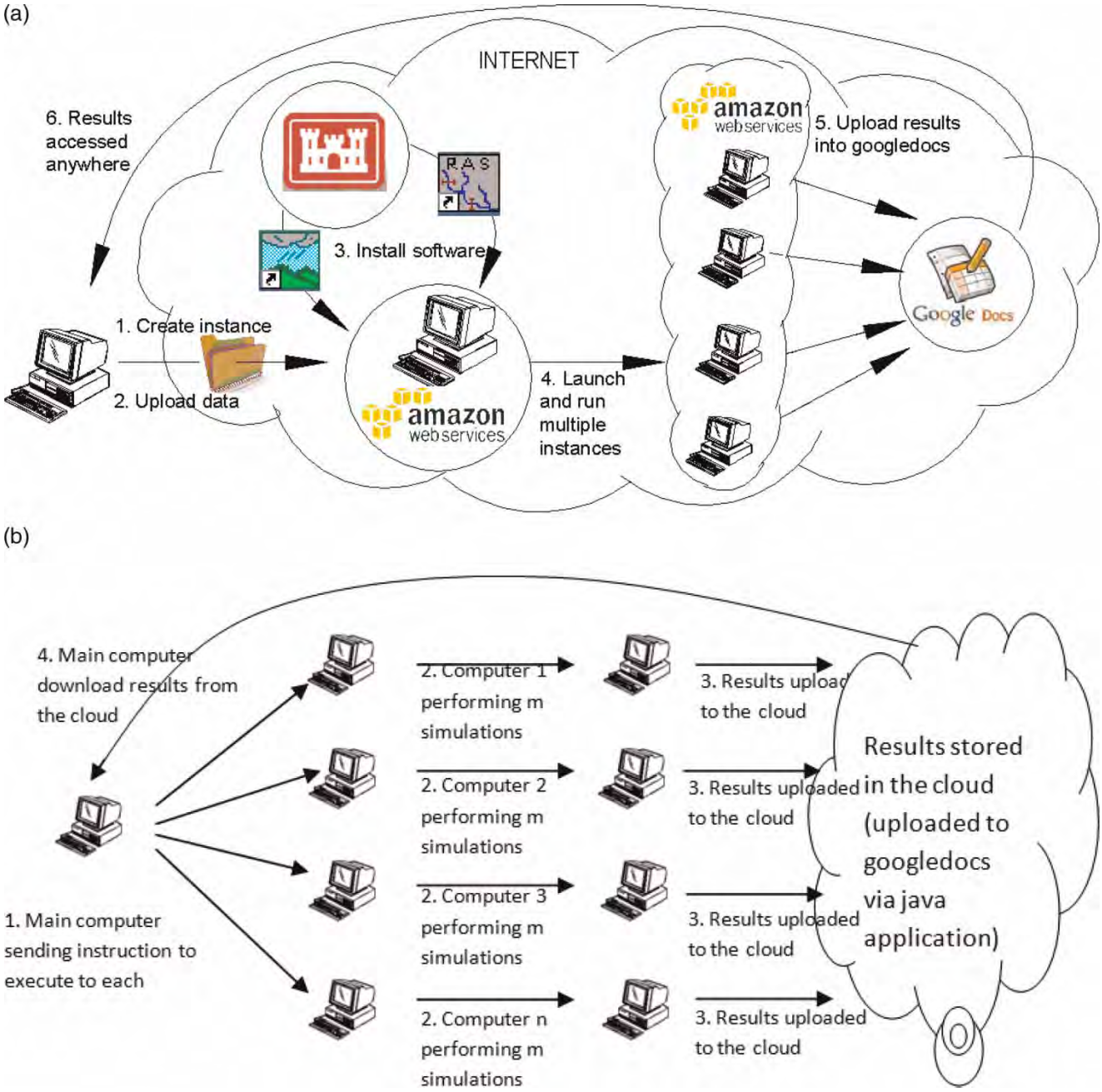

Figure 5 | (a) Cloud computing process. (b) Cluster computing process.

easy. The standard Monte Carlo approach relies upon independent generation of all samples, so no special job manager is needed and using the available tools was enough. However, in more complex sampling strategies groups of samples are generated conditionally, only after previous samples are evaluated, and then the use of a job manager is, of course, vital, and in this case developing additional software tools may be needed (see also experiences of Hunt *et al.* (2010)).

Use of cloud computing may be limited when licensed software is to be used, so in our case it was not possible to use licensed Sobek software on the cloud and a local cluster was employed. The availability of several computers to form a cluster typically is not a problem, but one needs the

software tools to operate such a cluster. Thus, different tools such as remote control, remote management, virtual network and telnet were tested, and it was found that the one with the best performance was the software for remote control *PsExec* available at http://technet.microsoft.com/en-us/sysinternals/default.aspx. Unlike the other tools where software packages need to be installed on each computer, *PsExec* is a very light tool that allows for executing processes on other computers on the local area network, and needs to be installed once on the master computer. While a PC is used as part of a cluster, the other users can log in, log off and use the computer without any problem (experiencing, of course, the lower performance since the processor is shared between several applications).

RESULTS AND DISCUSSION

As mentioned before, two types of uncertainties were investigated: the water levels in the reservoirs in the catchment and the DEM of the catchment.

Initial water level in the reservoirs

The initial water level in the Surduc reservoir has a small influence on the water level downstream of the catchment where the reservoir is located. Figure 6 presents the hydrographs at the Surduc reservoir, at station BG-J1, 10 km downstream of the Surduc reservoir and at station BG-J2, 30 km downstream of the BG-J1 station. Due to the scale of the plots it is difficult to distinguish all the data. In Figure 6, although there seem to be only two lines, actually there are 100 lines, each one representing the hydrograph for a given initial water level condition. It shows that in the reservoir the outflow becomes available as downstream flow once the reservoir is full. As we look at the downstream hydrographs this is less evident because of the contribution from tributaries through the lateral inflows.

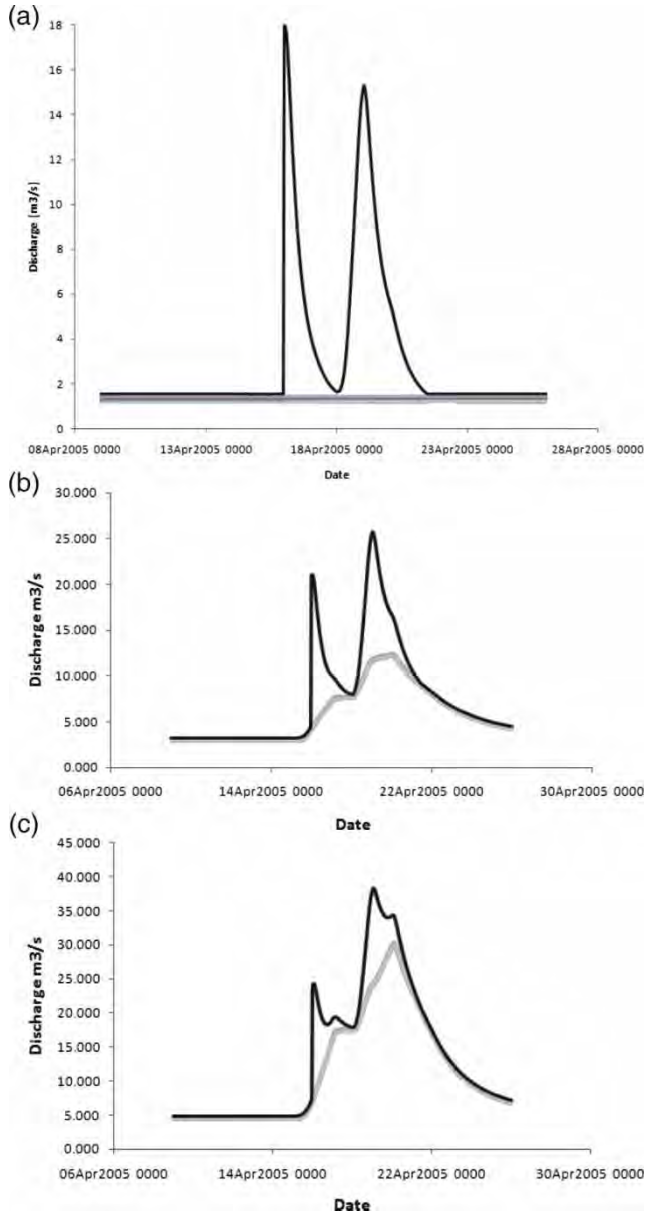

Figure 6 | (a) Hydrographs at the outlet from the Surduc reservoir. (b) Hydrographs at station BG-J1, 10 km downstream of the Surduc reservoir. (c) Hydrographs at station BG-J2, 30 km downstream of station BG-J1.

DEM

The Sobek model of the Timis–Bega catchment is the most complex and time-consuming part of the integrated model. After testing the possibility to use cloud computing and automation of data handling for different running scenarios, we focused on the applicability of distributed computing to the time-consuming flood inundation 1D2D Sobek model. The resulting file of each time step and each parameter is about 1 MB and the total number of time steps of computation for the 2005 event is 2,952, covering 10 d. Saving all the results from all the simulations would have taken a lot

of computer memory resources (taking into account that after each simulation 17.7 GB of data is generated). This was the reason to use just the results of a certain time step critical for a considered event (typically, corresponding to maximum flood depth). Also the flood map was saved after each simulation allowing us to develop flood probability maps representing the percentage of times that each cell get flooded with respect to the total number of simulations.

The Sobek model was run on a cluster of computers set-up on the corporate LAN. The number of computers available in the cluster was smaller than the number of scenarios to be run. According to the list of scenarios a queue of tasks was created to be run on the cluster. As soon as one scenario (task) was finished on one computer, the next scenario from the queue was taken to be performed as a task in the cluster. Based on the results from the scenarios, probability flood maps presenting the cells to be flooded were developed, using the ratio of the number of times a cell is flooded to the total number of simulation scenarios. The cell probability to be flooded (CPF) is formulated as

$$\mathrm{PC}_i = \sum \mathrm{C}_i / N \tag{1}$$

where PC_i = probability of a cell i to be flooded; $C_i = 1$ if cell i is flooded, 0 otherwise; and N = total number of simulation scenarios.

Different CPF maps were developed taking into account different numbers of simulations (Figure 8(b) presents the map generated after 500 simulations). Such maps were compared with the MODIS image registered for the event. There is a high correlation between the MODIS image and the 10% CPF area (i.e. the area for which the probability to be inundated is estimated at 10%). Both on the MODIS image and in the Sobek model, the water from the river Timis reaches the Bega river, and the flow paths at the border between Romania and Serbia are the same (Figure 7). The 90 and 50% CPF areas are also in an area covered by the MODIS image, but with less overlapping between the MODIS image and the CPF maps. By following the path of the CPF maps, it is easy to find the pattern of the event. It begins by flooding the area corresponding to 90% CPF,

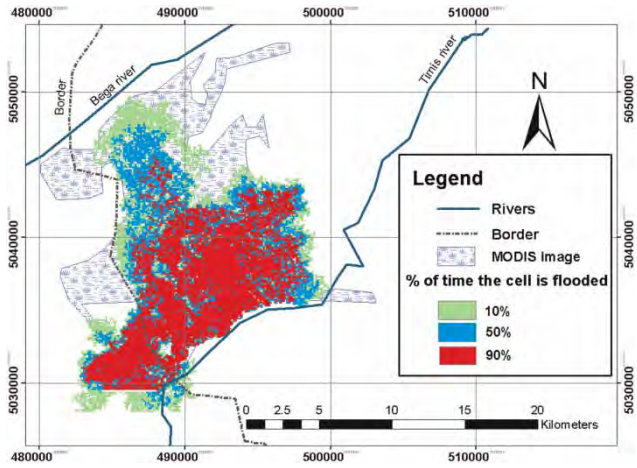

Figure 7 | MODIS flood extent and simulated flood extent.

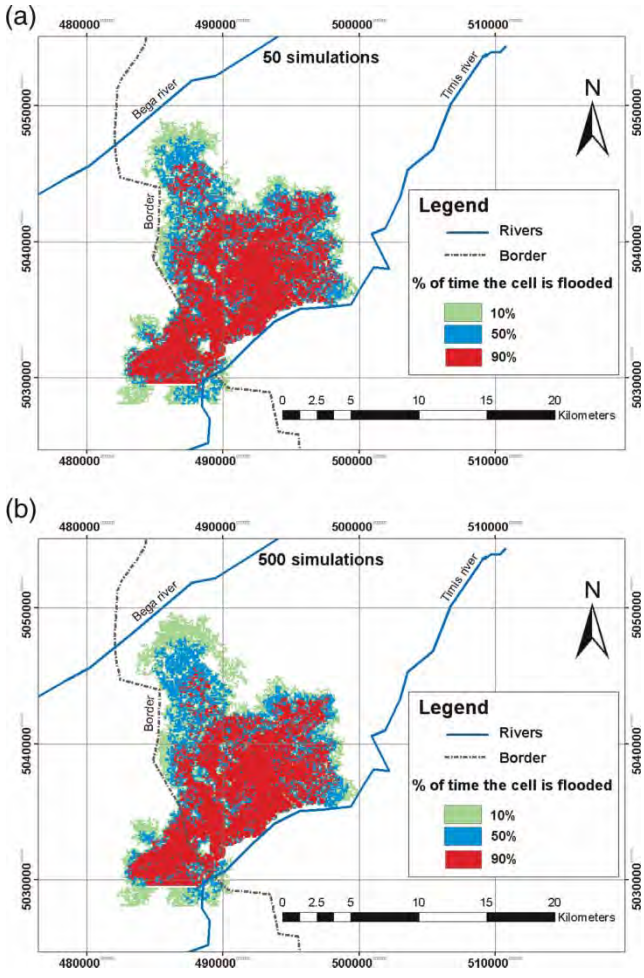

Figure 8 | (a) CPF map for 50 simulations. (b) CPF map for 500 simulations.

then the 50% CPF area and finally the 10% CPF area. Thus, the area of the MODIS image not covered by the 10% CPF area can be defined as cells with a probability to be flooded lower than 10%. The logical question which arises is: how many simulations/scenarios of DEM uncertainty maps are needed in order to have a correct representation of the CPF maps. In order to answer the question, three kinds of maps were compared: mean flood extent maps, CPF maps and standard deviation of water depths.

Analysing the CPF maps (Figure 8), most of the computational cells are flooded in 90% CPF of the simulations, the 50% CPF is located in an area surrounding the 90% CPF, and the outer region is flooded 10% of the time. Although there are cells that are flooded with a given probability and not flooded for a different probability, the overall flood pattern is the same. The cells with lower probability are the furthest ones and the differences increase as the depth increases; all this fits what is expected from flood probability maps in general.

For the analysis and representation purposes, flood water depths were divided into three ranges: below 1 m, between 1 and 2 m, and deeper than 2 m (Figure 9). In all cases they have a similar pattern and can be related to the pattern of flood propagation. The cells with flood water depth deeper than 2 m are the closest to the Bega river and could be considered not only as the most dangerous ones, but also as the first ones to get flooded. The cells with a flood lower than 1 m are located in the outer region of the flooded area. Although the total flood and flood water depth deeper than 1 m are quite similar, there is a notable difference for the flood water depths above 2 m that could be considered as the most critical area. In both cases the affected area is a continuous region, for similar flood water depths, with no isolated cells of different depths. Only in the external regions could some (physically impossible) isolated cells be noticed, but these cells belong to the group of lower probability, i.e. corresponding to a lower degree of hazard, so the influence of such (less reliable) data on decisions concerning hazards of high probability is negligible. The maps for high depths are most useful when evaluating flood hazards since deeper floods are more dangerous than the shallow ones.

Although the results are very similar at the macro-scale, while comparing flood water depths cell by cell, differences

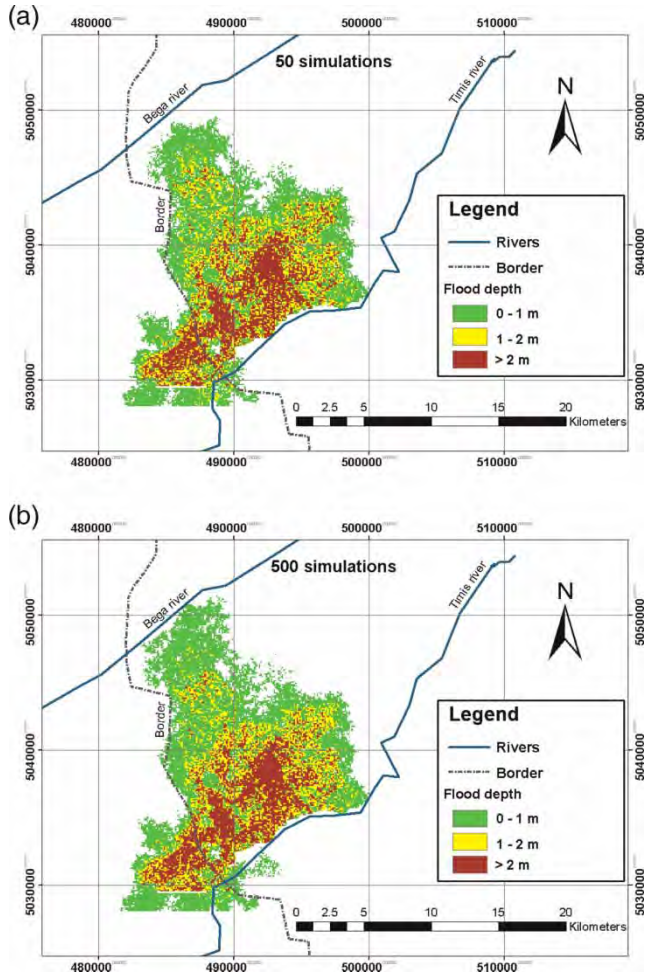

Figure 9 | (a) Mean flood depth for 50 simulations. (b) Mean flood depth for 500 simulations.

can be noted. The analysis of these differences was done using the standard deviation error. Standard deviation of flooded cells from each CPF map was analysed for different number of simulations (Table 1).

Analysing the standard deviation it can be noticed that for a small number of simulations not only the standard deviation is higher, but also it has high variablity across the domain. For instance, Figures 10(a) and (b) shows that for ten simulations there are few neighbouring cells with similar deviation, while for 500 simulations the regions of given deviation are grouped in continuous areas. Grouping the depths into ranges (lower than 1 m, between 1 and 2 m, and deeper than 2 m) also shows the importance and improvement in results when increasing the number of simulations. For instance, in Table 1 it is easy to note that

Table 1 | Mean flood depth standard deviation

No. of simulations	σ = 0–1 m	σ = 1–2 m	σ = deeper than 2 m
1	–	–	–
5	0.2930	0.1360	0.1560
10	0.0663	0.1803	0.2467
25	0.1015	0.0599	0.1597
50	0.0942	0.0101	0.1056
100	0.2286	0.3760	0.6036
200	0.2714	0.1602	0.4316
400	0.0180	0.0446	0.0626

for five simulations and floods lower than 1 m the standard deviation is 0.3 m, so it is almost 30% of deviation, while for 500 simulations and the same depth range the standard deviation decreases to 0.018. Moreover, it was found that the standard deviations of depth get lower as the number of simulations increases. Also the standard deviation becomes more uniformly distributed as the number of simulations increases.

Effectiveness of cloud and cluster computing

Uncertainty analysis involving the HEC-HMS and HEC-RAS models of the Timis–Bega were tested on the cloud computing services. The simulation comprised of running the HEC-HMS part of the model in sequence with the HEC-RAS model. In the HEC-HMS model, several scenarios on the use of the reservoirs were tested. Due to the large area of the catchment the influence in the downstream station used as boundary conditions for the hydraulic model appeared not to be large.

The chosen sampling strategy required about 100 runs of the HEC-HMS model, followed by 100 runs of the HEC-RAS model. One single simulation of HEC-HMS took about 13 s and running 99 simulations on a single computer took a little more than 22 min (Figure 11). On the other hand, running 99 simulations on many computers in parallel reduces the total time. For instance, with five computers 99 simulations can be finished in little more than 4 min. This time economy determines the benefits on running multiple computers in parallel. For a small number of simulations the total time is quite similar, while for many simulations

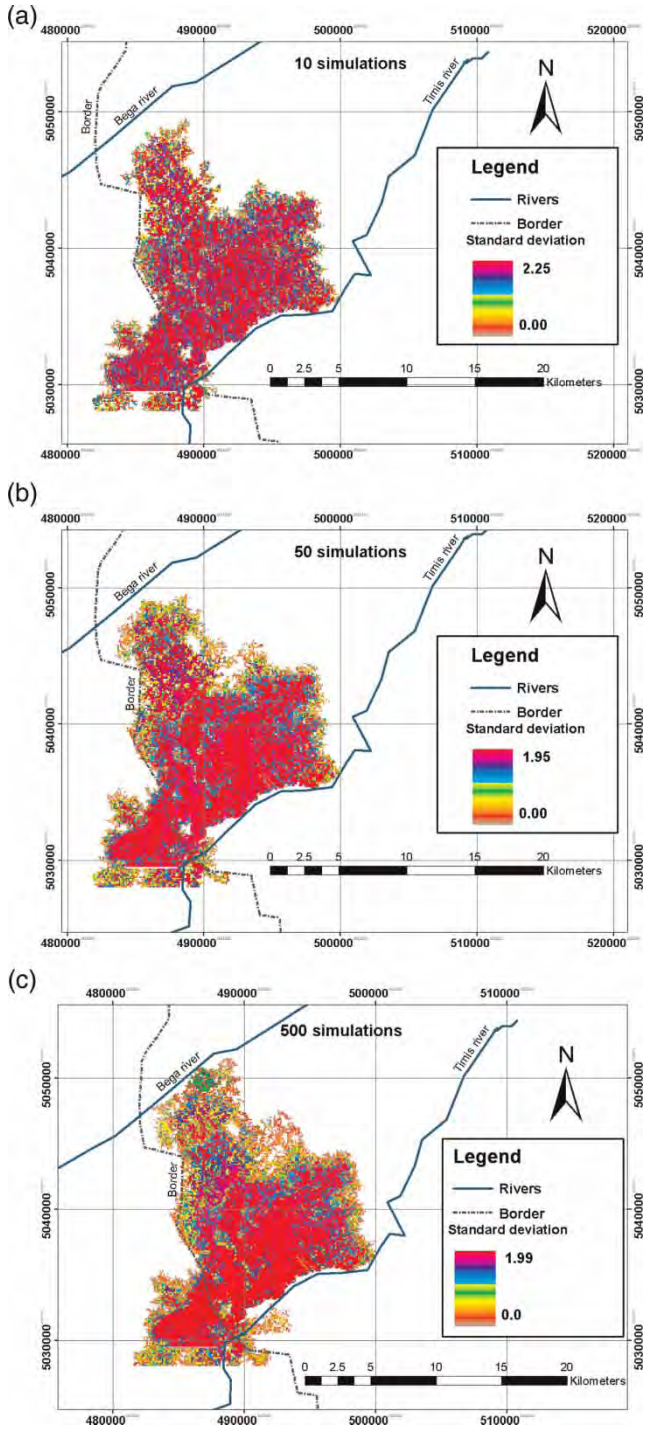

Figure 10 | (a) Standard deviation of depth for 10 simulations. (b) Standard deviation of depth for 50 simulations. (c) Standard deviation of depth for 500 simulations.

the time saved is considerable. In the present case, the time saved for 10 or five simulations is almost the same, but for 99 simulations there is a bigger gain. However,

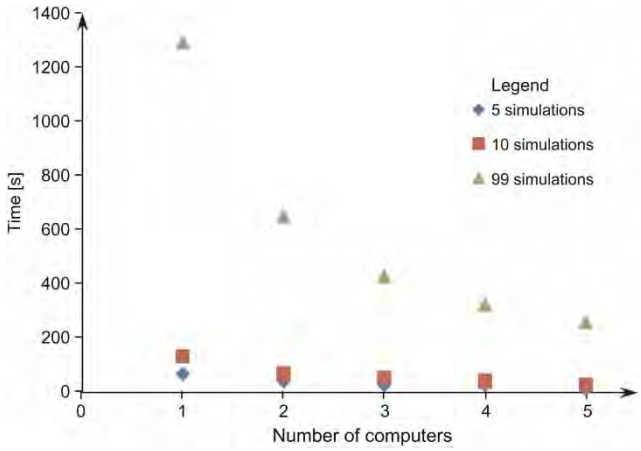

Figure 11 | Time saved in HEC-HMS model using distributed computing.

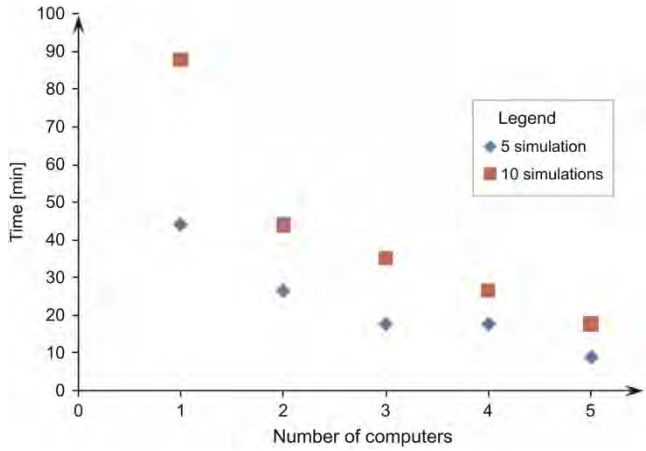

Figure 12 | Time saved in HEC-RAS model using distributed computing.

these times do not consider the launching time of the virtual machine on the cloud, which depends on several factors such as the instance type or the data loaded, and in the worst cases it may take up to 30 min (but typically it is much less). Thus, in the case of running a simple model with quite a low number of simulations, the benefits may be minimal. However, in the case of a more complex model like HEC-RAS where one single simulation takes around 10 min, the benefits of saving time in running many simulations are evident, even considering a relatively high initialisation time (Figure 12). For the long-running models the issue of the observed relatively slow initialisation of virtual machines on the Amazon EC2 platform becomes negligible, so that total running time will be approaching the theoretical limit of T/N, where T is the time needed for all Monte Carlo simulations and N is the number of the activated virtual machines.

The most complex and time-consuming model is the Sobek part of the model. It takes around 1.5 h for each scenario simulation. The benefit of running several scenarios on parallel computers is considerable. If 100 simulations are done with one computer only, it will take around 150 h, while with five computers the time reduces to only 30 h. As for the HEC-RAS model, the logarithmic trend relation between the number of computers versus the time of simulation is shown in Figure 13. Thus, as all the cases show the same trend, it can be concluded that there is a limit in increasing the number of computers running in parallel, beyond which time saving becomes negligible.

Number of simulations used for uncertainty analysis

As mentioned before, the number of simulations in Monte Carlo experiments is typically determined by analysing some statistical properties (mean, standard deviation and/or the distribution quantiles), which are expected to stabilise as the number of runs increases. In the present study, we used the MPI; it is defined by the 5 and 95% quantiles of the estimate of the pdf of flood depth at a location with maximum flooding depth. The stopping condition was set as follows: Monte Carlo runs would be terminated when the changes in MPI would be negligible (Shrestha *et al.* 2009). From Figure 14 it can be seen that the 500 simulations used were enough to ensure convergence of MPI; however, more accurate estimation of the number of simulations

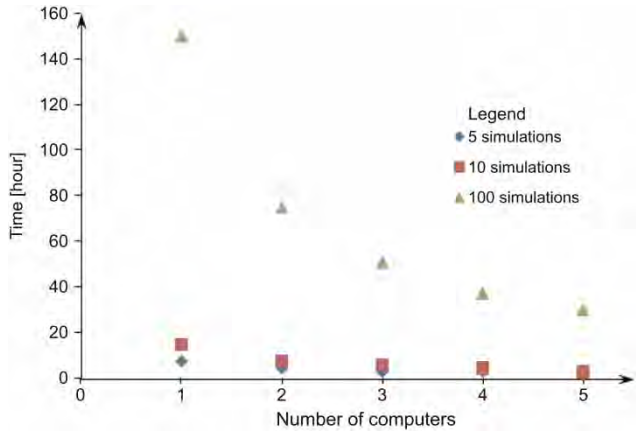

Figure 13 | Time saved in Sobek model using distributed computing.

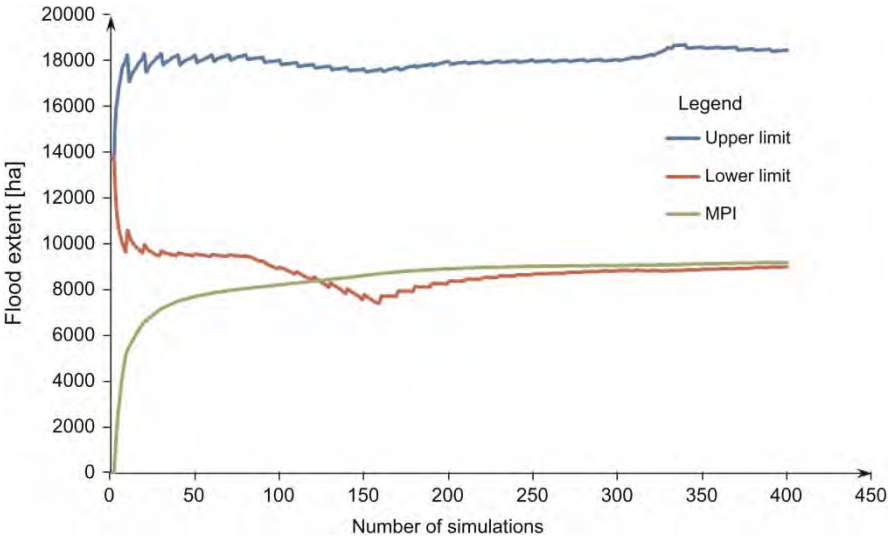

Figure 14 | Mean prediction interval.

needed to ensure statistical reliability of results is still to be done.

CONCLUSIONS

The present study should be seen as a proof of concept, which, however, demonstrates the benefits of using cluster and cloud technology when dealing with uncertainty analysis of complex hydraulic and hydrologic models. For more than four or five computers the economy in computing time approaches a linear trend. Although cloud computing has a strong potential for the application of distributed computing to complex hydraulic systems, at low cost, it still has the barrier of licensing when using commercial software. It is important to note that by the time of the experiment (2010–2011) cloud computing was a new development and was not seen by most hydraulic modellers as a technology to try. One of the practical issues of cloud computing is that licenced software would need a virtual licence and virtual dongle. Recently, however, some of the biggest engineering software companies (e.g. ESRI and AutoDesk) addressed cloud computing as the new trend, and began to offer specific cloud computing services. DHI presented a possibility to run their MIKE software as a service based on a daily fee (https://saas.dhigroup.com). So we expect that in the near future running licensed software on a cloud will not be a problem.

The presented work shows that uncertainties in DEM have a considerable influence on the flood extent. We used a simplified sampling technique of changing the cells' elevations independently, so further research is suggested on developing more elaborate sampling techniques based on the real statistical properties of DEM, with a correlated variation of different cells' elevations correlated.

Uncertainty analysis is important for evaluating floods to analyse the magnitude of the event and for better estimation of the affected area. In the present case we demonstrated that the deterministic simulation resulted in the flooded area of around 14,000 ha, whereas the uncertainty analysis allowed us to show that also considering the cells flooded in 10% of time increases the area up to 18,000 ha.

It is difficult to label one method of distributed computing (cloud or cluster) as the best choice for all cases. If the user already has enough computational power (a dedicated cluster or a supercomputer) with the required software, often there is no need for cloud computing. But if the user does not have enough computing power or does not have all the required software, then using the cloud should be seriously considered. Cloud computing may be the preferred option if one has to run a full-fledged Monte Carlo uncertainty analysis experiment that needs thousands of simulations employing computationally intensive models.

Besides, cloud computing provides additional benefits by virtualisation of the work, e.g. instant access to the data and computer power required regardless of the time or place.

The proposed procedure of model integration and distributed computing is a low cost alternative that can be easily applied to cases were fast evaluation of different scenarios is needed. In the present case different DEMs were evaluated, but the methodology of model integration and cluster computing can also be applied to evaluate different types of uncertainties, flood measures, and for multiple runs in climate change scenario-based studies.

REFERENCES

Aldescu, C. 2008 The necessity of flood risk maps on Timis river. In *Proc. 24th Conf. of the Danubian Countries on Hydrological Forecasting and Hydrological Bases of Water Management, Bled, Slovenia. IOP Conf. Series Earth Environ. Sci.* Vol 4. IOP Publishing, Bristol, pp. 54–60.

Beven, K. & Freer, J. 1996 Bayesian estimation of uncertainty in runoff prediction and the value data: an application of the GLUE approach. *Water Resour. Res.* **32**, 2161–2173.

Boukerche, A., de Melo, A. C. M. A., Ayala-Rincón, M. & Walter, M. E. M. T. 2007 Parallel strategies for the local biological sequence alignment in a cluster of workstations. *J. Parallel Distrib. Comput.* **67** (2), 170–185.

Brown, J. D. & Heuvelink, G. B. M. 2007 The Data Uncertainty Engine (DUE): a software tool for assessing and simulatng uncertain environmental variables. *Comput. Geosci.* **33**, 172–190.

Carabajal, C. C. & Harding, D. J. 2006 SRTM C-band and ICES at laser altimetry elevation comparisons as a function of tree cover and relief – validation of SRTM C-band DEMs using ice, cloud, and land elevation satellite (ICESat) data. *Photogramm. Eng. Remote Sens.* **72** (3), 287–298.

Cobby, D. M., Mason, D. C., Horritt, M. S. & Bates, P. D. 2003 Two-dimensional hydraulic flood modelling using a finite-element mesh decomposed according to vegetation and topographic features derived from airborne scanning laser altimetry. *Hydrol. Process.* **17** (10), 1979–2000.

Cutore, P., Campisano, A., Kapelan, Z., Modica, C. & Savic, D. 2008 Probabilistic prediction of urban water consumption using the SCEM-UA algorithm. *Urban Water J.* **5** (2), 125–132.

Damas, M., Salmeron, M., Ortega, J., Olivares, G. & Pomares, H. 2001 Parallel dynamic water supply scheduling in a cluster of computers. *Concurrency Comput. Pract. Exp.* **13** (15), 1281–1302.

De Oliveira, P. T. S., Sobrinho, T. A., Steffen, J. L. & Rodrigues, D. B. B. 2010 Caracterização morfométrica de bacias hidrográficas através de dados SRTM (Morphometric characterization of hydrographic basins through SRTM data). *Rev. Bras. Eng. Agric. Ambient. Revista Brasileira de Engenharia Agricola e Ambiental* **14** (8), 819–825.

Dejun, J., Pierre, G. & Chi, C.-H. 2010 EC2 performance analysis for resource provisioning of service-oriented applications. In: *Service-Oriented Computing. ICSOC/ServiceWave 2009 Workshops*. Springer, Berlin, pp. 197–207.

Federal Emergency Management Agency 2003 *Guidelines and Specifications for Flood Hazard Mapping Partners*. FEMA Flood Hazard Mapping Program, Washington, DC.

Ferreira, C. 2004 *Metodo para elaboracao de mapas de inundacao: Estudo de caso na bacia do rio Palmital Parana, Setor de Tecnologia* (*Method to Elaborate Flood Maps: Case Study of Palmital Parana River Basin*). Universidade Federal do Paraná, Curitiba.

Garcia, M., Basile, P., Riccardi, G. & Stenta, H. 2007 Modelacion hidrodinamica de sistemas cauce – planicie de inundacion en Grandes Rios Aluviales de LLanura (Hydrodynamic modelling of cannel – floodplain systems in large aluvial river plain). In *Proc. III Simposio Regional sobre Hidraulica de Rios, 7–9 November, Cordoba-Argentina*. Instituto de Recursos Hidricos-Universidad de Santiago del Estero, Instituto Nacional del Agua, pp. 27–42 (Available from: http://irh-fce.unse.edu.ar/Rios2007/index_archivos/A/3.pdf).

Gorokhovich, Y. & Voustianiouk, A. 2006 Accuracy assessment of the processed SRTM-based elevation data by CGIAR using field data from USA and Thailand and its relation to the terrain characteristics. *Remote Sens. Environ.* **104** (4), 409–415.

Hengl, T., Heuvelink, G. B. M. & van Loon, E. E. 2010 On the uncertainty of stream networks deruived from elevation data: the error propagation approach. *Hydrol. Earth Syst. Sci. Discuss.* **7**, 767–799.

Herold, C. & Mouton, F. 2006 *Global Flood Modelling: Statistical Estimation of Peak-flow Magnitude*. World Bank Development Research Group & UNEP/GRID – Europe Early Warning Unit, Geneva, Switzerland.

Hunt, R. J., Luchette, J., Schreuder, W. A., Rumbaugh, J. O., Doherty, J., Tonkin, M. J. & Rumbaugh, D. B. 2010 Using a cloud to replentish parched groundwater modeling efforts. *Ground Water* **3** (48), 360–365.

Kang, H. & Choi, S. U. 2005 3D numerical simulation of compound open-channel flow with vegetated floodplains by reynolds stress model. *KSCE J. Civil Eng.* **9** (1), 7–11.

Knebl, M. R., Yang, Z. L., Hutchison, K. & Maidment, D. R. 2005 Regional scale flood modelling using NEXRAD rainfall, GIS, and HEC-HMS/RAS: a case study for the San Antonio River Basin Summer 2002 storm event. *J. Environ. Manage.* **75** (4), 325–336.

Lin, B., Wicks, J. M., Falconer, R. A. & Adams, K. 2006 Integrating 1D and 2D hydrodynamic models for flood simulation. *Proc. Inst. Civil Engrs. Water Manage.* **159** (1), 19–25.

Lindenschmidt, K.-E., Huang, S. & Baborowski, M. 2008 A quasi-2D flood modelling approach to simulate substance transport in polder systems for environment flood risk assessment. *Sci. Total Environ.* **397** (1–3), 86–102.

Ludwig, R. & Schneider, P. 2005 Validation of digital elevation models from SRTM X-SAR for applications in hydrologic modelling. *J. Photogramm. Remote Sens.* **60** (5), 339–358.

Macdonald, I. & Strachan, P. 2001 Practical application of uncertainty analysis. *Energy Build.* **33** (3), 219–227.

Maccyzk, P. & Janusz, J. 2008 Cluster computing of mechanisms dynamics using recursive formulation. *Multibody Syst. Dyn.* **20** (2), 177–196.

Moya, V., Popescu, I. & Solomatine, D. P. 2010 Monte Carlo uncertainty analysis of hydraulic models using cloud computing. In *Proc. 9th Int. Conf. on Hydroinformatics, Tianjin, China*, Vol 2 (J. Tao, Q. Chen & S.-Y. Liong, eds). Chemical Industry Press, Tianjin, pp. 913–921.

Neal, J. C., Fewtrell, T. J., Bates, P. D. & Wright, N. G. 2010 A comparison of three parallelisation methods for 2D flood inundation models. *Environ. Modell. Softw.* **25** (4), 398–411.

Popescu, I., Jonoski, A., Van Andel, S. J., Onyari, E. & Quiroga, V. G. M. 2010 Integrated modelling for flood risk mitigation in Romania: case study of the Timis–Bega river basin. *Int. J. River Basin Manage.* **8** (3), 269–280.

Pramanik, N., Panda, R. K. & Sen, D. 2010 One dimensional hydrodynamic modelling of river flow using DEM extracted river cross-sections. *Water Res. Manage.* **24** (5), 835–852.

Riccardi, G. A. 1997 Elaboracion de mapas de riesgo de inundacion por medio de la modelacion matematica hidrodinamica (Elaboration of flood risk maps through mathematic hydrodynamic modelling). *Ingenieria del agua* **4** (3), 45–56.

Rodriguez, E., Morris, C. S., Belz, J. E., Chapin, E. C., Martin, J. M. & Hensley, S. 2005 An Assessment of the SRTM Topographic Products. Technical Report JPL D-31639. Jet Propulsion Laboratory, Pasadena, CA.

Rosenthal, A., Mork, P., Li, M. H., Stanford, J., Koester, D. & Reynolds, P. 2010 Cloud computing: a new business paradigm for biomedical information sharing. *J. Biomed. Inform.* **43** (2), 342–353.

Schanze, J. 2006. Flood risk management: hazards, vulnerability and mitigation measures. In *Proc. NATO Advanced Research Workshop on Flood Risk Management – Hazards, Vulnerability and Mitigation Measures*, Ostrov, Czech Republic, 6–10 October 2004. Springer, Berlin, pp. 233–235.

Schumann, G., Matgen, P., Cutler, M. E. J., Black, A., Hoffmann, L. & Pfister, L. 2008 Comparison of remotely sensed water stages from LiDAR, topographic contours and SRTM. *J. Photogramm. Remote Sens.* **63** (3), 283–300.

Sevior, M., Fifield, T. & Katayama, N. 2010 Belle Monte-Carlo production on the Amazon EC2 cloud. In: *Managed Grids and Cloud Systems in the Asia-Pacific Research Community* (S. C. Lin & E. Yen, eds). Springer, New York, pp. 231–249.

Sharma, A., Tiwari, K. N. & Bhadoria, P. B. S. 2010 Vertical accuracy of digital elevation model from Shuttle Radar Topographic Mission – a case study. *Geocarto Int.* **25** (4), 257–267.

Shrestha, D. L., Solomatine, D. P. & Kayastha, N. 2009 A novel approach to parameter uncertainty analysis of hydrological models using neural networks. *Hydrol. Earth Syst. Sci.* **13** (7), 1235–1248.

Smemoe, C. M., Nelson, E. J., Zundel, A. K. & Miller, A. W. 2007 Demonstrating floodplain uncertainty using flood probability maps 1. *J. AWRA* **43** (2), 359–371.

Soldano, A., Giraut, M. & Goniadzki, D. 2007 Mapa de Susceptibilidad Urbana Ante Inundaciones, Caso: Ciudad de Goya, Provincia de Corrientes. *TELEDETECCION – Hacia un mejor entendimiento de la dinamica global y regional*. Martin (ed.), Madrid, Spain, pp. 449–456.

Soumendra, N. K., Sen, D. & Bates, P. D. 2010 Coupled 1D-quasi-2D flood inundation model with unstructured grids. *J. Hydraul. Eng.* **138** (8), 493–506.

Tarrant, O., Todd, M., Ramsbottom, D. & Wicks, J. 2005 2D floodplain modelling in the tidal Thames – addressing the residual risk. *Water Environ. J.* **19** (2), 125–134.

Teodorescu, N. I. 2008. The main characteristics of the high water registered in the river basin Bega in February. In *Proc. 24th Conf. of the Danubian Countries on Hydrological Forecasting and Hydrological Bases of Water Management, Bled, Slovenia. IOP Conf. Series Earth Environ. Sci.* Vol 4. IOP Publishing, Bristol, pp. 147–153.

Vrugt, J. A., Gupta, H. V., Bouten, W. & Sorooshian, S. 2003 A shuffled complex evolution Metropolis algorithm for optimization and uncertainty assesment of hydrologic model parameters. *Water Resour. Res.* **39** (8), 1201–1208.

Vrugt, J. A., ter Brakk, C. J. F., Gupta, H. V. & Robinson, B. A. 2008 Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modelling? *Stoch. Environ. Res. Risk Assess.* **23** (7), 1061–1062.

Wechsler, S. 2007 Uncertainties associated with digital elevation models for hydrologic applications: a review. *Hydrol. Earth Syst. Sci., Spec. Issue: Uncertain. Hydrol. Obs.* **11**, 1481–1500.

Wechsler, S. & Kroll, C. 2006 Quantifying DEM uncertainty and its effects on topographic parameters. *Photogramm. Eng. Remote Sens.* **72** (9), 108–119.

Weinhardt, C., Anandasivan, A., Blau, B., Borissob, N., Meinl, T., Michalk, W. & Stößer, J. 2009 Cloud computing – a classification, business models, and research directions. *Business Inf. Syst. Eng.* **1** (5), 391–399.

Wilson, C. A. M. E., Yagci, O., Rauch, H. P. & Olsen, N. R. B. 2006 3D numerical modelling of a willow vegetated river/floodplain system. *J. Hydrol.* **327** (1–2), 13–21.

Xu, Z., Vélez, C., Solomatine, D. P. & Lobbrecht, A. 2010 Use of cloud computing for optimal design of urban wastewater systems. In *Proc. 9th Int. Conf. on Hydroinformatics, Tianjin, China*, Vol 2 (J. Tao, Q. Chen & S.-Y. Liong, eds). Chemical Industry Press, Tianjin, pp. 913–921.

Zheng, X. 2010 Application of cloud computing and surrogate model approximators in multi-objective optimization of urban wastewater system. MSc Thesis, UNESCO–IHE, Delft, The Netherlands.

First received 9 February 2011; accepted in revised form 11 May 2012. Available online 18 July 2012

Hydrol. Earth Syst. Sci., 16, 4637–4649, 2012
www.hydrol-earth-syst-sci.net/16/4637/2012/
doi:10.5194/hess-16-4637-2012

Hydrology and Earth System Sciences

A study of the climate change impacts on fluvial flood propagation in the Vietnamese Mekong Delta

P. D. T. Van[1,2], I. Popescu[2], A. van Griensven[2,5], D. P. Solomatine[2,3], N. H. Trung[1], and A. Green[4]

[1]College of Environment and Natural Resources, Can Tho University, Vietnam
[2]UNESCO-IHE Institute for Water Education, Delft, The Netherlands
[3]Water Resources Section, Delft University of Technology, The Netherlands
[4]Mekong River Comission, Phnom Penn, Cambodia
[5]Department of Hydrology and Hydraulic Engineering, Vrije University Brussels, Belgium

Correspondence to: I. Popescu (i.popescu@unesco-ihe.org)

Received: 11 May 2012 – Published in Hydrol. Earth Syst. Sci. Discuss.: 8 June 2012
Revised: 7 November 2012 – Accepted: 19 November 2012 – Published: 13 December 2012

Abstract. The present paper investigated the extent of the flood propagation in the Vietnamese Mekong Delta under different projected flood hydrographs, considering the 2000 flood event (the 20-yr return period event, T. V. H. Le et al., 2007) as the basis for computation. The analysis herein was done to demonstrate the particular complexity of the flood dynamics, which was simulated by the 1-D modelling system ISIS used by the Mekong River Commission. The floods of the year 2050 are simulated using a projected sea level rise of +30 cm. The future flood hydrograph changes at Kratie, Cambodia, were also applied for the upstream boundary condition by using an adjusted regional climate model. Two future flood hydrographs were applied at the upstream part of the delta, the first one in a scenario of climate change without considering developments in the Mekong Basin,and the second one in a scenario of climate change taking into account future development of the delta. Analyses were done to identify the areas sensitive to floods, considering the uncertainty of the projection of both the upstream and downstream boundary conditions. In addition, due to the rice-dominated culture in the Vietnamese Mekong Delta, possible impacts of floods on the rice-based farming systems were also analysed.

1 Introduction

Climate change is an on-going process with notable impacts on the eco-hydrological environment (Black and Burns, 2002; Gupta et al., 2002; Prudhomme et al., 2003) leading to significant challenges to the livelihood of local residents in different parts of the world (Lespinas et al., 2009; Muste et al., 2010; Quinn et al., 2010). Vietnam is seen to be one of the most vulnerable countries to global climate change impacts, and the Vietnamese Mekong Delta (VMD) (Fig. 1) is identified as particularly susceptible to the impacts of extreme climate events and climate variability (ADB, 2009; WWF, 2009). Possible changes of the hydrological conditions (including spatial and temporal distribution of floods, modification of wet and dry season precipitation and alteration of salinity intrusion pattern) as a consequence of the global climate change may present significant threats to the socio-economic and environmental systems (Quinn et al., 2010).

Hoanh et al. (2010) conducted a study to explore the possible impacts of both the climate change and the economic development on the flow regimes of the Mekong and its delta. The study concluded that, given the above mentioned conditions, there could be a possible increase of flooding events during wet season, while there will be more periods of water shortage in the dry season. The maximum monthly flow, on the Mekong river is estimated to increase by 35 to 41 %, and by 16 to 19 % in the delta, with lower values estimated for the years 2010 to 2038 and higher values for the 2070 to 2099 (Hoanh et al., 2010; Västilä et al., 2010). The same studies estimate a decline in the minimum monthly flows, by 17 to 24 % in the basin, and by 26 to 29 % in the delta. As a consequence, expansion of areas under severe water stress (flood and drought) would be one of the most pressing environmental problems in the Mekong basin and the delta, as

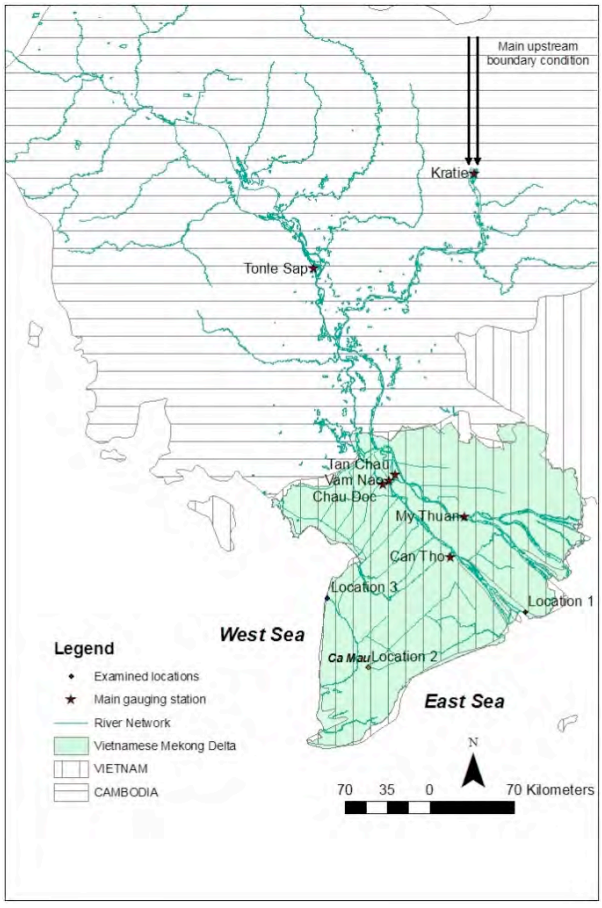

Fig. 1. Mekong river network and the Vietnamese Mekong Delta.

the number of people living in the area is likely to increase substantially.

One of the major impacts of the climate change is expected to be the sea level rise (SLR). The SLR in the West and East Seas were defined by the Vietnamese Ministry of Natural Resources and Environment (MONRE). Looking at the SLR predictions for 2050, as compared with the 2000 situation, this might result in a large inundated area of the VMD mainly along the east and west coast of the delta (WWF, 2009), leading to significant loss of mangrove forest and agricultural lands, while the livelihoods of about 1.9 million local residents will be at risk (Ericson et al., 2006).

Studies of the modifications of the water balance within the Mekong basin and delta, combined with sea level rise, show that this would result in great alteration of the eco-hydrological environment (Lu and Siew, 2006), and consequently, to adverse impacts on the socio-economic system would be felt in the VMD. In fact, the modification of the upstream discharge would be the main concern for the upstream section of the VMD (Västilä et al., 2010; Dinh et al., 2012), while the rise of the sea level resulting in wider (temporal) inundation along the coast, is important for the downstream section (Carew-Reid, 2008). Apart from the significant changes on the livelihood of the local residents (Nuorteva et al., 2010), the sea level rise would contribute to changes of the hydraulic nature along the river network (Dooge and Napiorkowski, 1987).

A number of studies were conducted to project the impacts of changes of the upstream discharge and the sea level rise on the VMD with specific attention to the inundation in the upstream section of the VMD (e.g. Wassmann et al., 2004; Le et al., 2007, 2008; Dinh et al., 2012). It was found that the impacts of projected climate change would be an expansion of the inundation area towards the sea (Wassmann et al., 2004; Dinh et al., 2012), and the average and maximum water levels, as well as the flood duration will increase in the period 2010–2049 (Västilä et al., 2010). Johnson and Kummu (2011) make an extensive review of the available models currently applied in the Mekong Basin. These models are all based on 1-D Saint Venant equations, are using both commercial (e.g. ISIS, Mike 11) and non-commercial developed software tools (i.e. Vietnam River Systems and Plains (VRSAP)) and are used to study the flood situation in the VMD. However, these studies did not consider either the modification of the flood dynamics in the delta after the hydrograph upstream changes in combination with sea level rise or the river triggered floods of short duration, which are affecting the coastal area of Vietnam (i.e. the downstream section of the VMD). In addition, detailed inundation modelling and mapping of the VMD floods of short duration should be addressed as well because of their significant impacts on farming.

The present study is looking into the temporal and spatial dynamics of the fluvial flooding of the whole VMD in cases of short duration events. The study uses the 1-dimensional hydrodynamic model ISIS provided by the Mekong River Commission (MRC). The model is validated for the flood event of the year 2000 (a short flood event), and based on it, we investigate possible impacts of climate change on the flood dynamics (e.g. inundation extent and temporal variation) considering the combined effects of predicted changes of the upstream flow and sea level rise. There are two main drivers for the changes of the upstream flow of the Mekong: the climate change, and the hydropower and irrigation developments that are planned to take place in the near future.

Some analysis of sensitivity of the model results to different boundary conditions is presented in this paper; however, detailed uncertainty analysis of the inundation pattern resulting from the hydrodynamic model is the focus of a separate analysis that is currently in preparation and will be reported in a subsequent publication.

It is important to mention that the study is focused only on the fluvial floods generated by the SLR, and does not consider the floods due to coastal phenomena such as surges, which are occurring on an elevated sea level. In addition, pluvial floods are also not considered because they do occur occasionally and lead to local effects that are less significant than the fluvial ones.

1.1 The study area

According to the Mekong River Commission (MRC) (MRC, 2007), the Mekong Delta begins at Phnom Penh, Cambodia, where the river divides into two main branches, the Mekong and the Bassac (Fig. 1). In Vietnam, the Mekong Delta is relatively flat (generally, the average land surface elevation is not greater than 5.0 m above the mean sea level (m a.s.l.)) and the complex channel network forming the delta has been modified over a long period of time due to agricultural activities and residential developments.

The VMD river network is fluvially unstable (Reichel and Nachtnebel, 1994; Neuhold et al., 2009; MRC, 2010a) and the hydraulic nature of it is highly complex, impacted by both the upstream discharge and sea level along the East and West Seas. The East and West Seas have semidiurnal and diurnal tides, respectively (Nguyen and Savenije, 2006) (Fig. 2a and b). The tidal amplitude from the East Sea is between 3.0 and 3.5 m a.s.l., resulting in significant daily water level variations, especially during the dry season when the river stage is strongly driven by the tidal regimes. For example, at Can Tho (90 km upstream from the sea), the daily variations of the river water level are of about 1.5 to 2.0 m, and at Tan Chau and Chau Doc (190 km upstream from the coast), the variations are of approximately 1.0 m. In Fig. 3a the maximum measured daily discharge in 2000 is presented for two main upstream gauging stations in the VMD (Tan Chau and Chau Doc). The discharge entering the VMD is routed along the Mekong branch and then the flow is divided between the Mekong branch and the Vam Nao canal, which routes part of the water to the Bassac branch of the delta.

As compared with the upstream part of the Mekong River, the flood hydrology of the VMD is typically classified as having relatively low peaks but high volume. This behaviour is due to the effect of the natural Tonle Sap Lake, which is an important reservoir in the Mekong basin. The lake reduces the intensity of the upstream flood hydrograph by storing a part of flood flows and releasing gradually flows to downstream (MRC, 2007). An analysis of the annual floods of the last 50 yr shows that every year flood lasts for several months and is characterised by discharges that are higher than $13\,600\,m^3\,s^{-1}$, which is defined as a critical threshold (MRC, 2007) (Fig. 3b).

Studies on the VMD flood extent, impact and vulnerability need to consider the effects on local economies, which are in direct relation to the land use or the so called "agro-ecological" zones of the VMD (Delgado et al., 2010). The agro-ecological zones and the land use map of 2006 are represented in Fig. 4. It can be noticed that the VMD is mainly dominated by rice-based farming. Fruit gardens are mainly located along the Mekong branch and to some extent along the Bassac. In the Ca Mau Peninsula, the main land cover types are single rice crop, double-crop of single rice crop in combination with a single upland crop, and intensive shrimp farming. In this study the land cover pattern of 2006 (Fig. 4) was available and therefore used as a reference for further discussion about the impacts of future flooding on the agriculture in the VMD, and it could be stated that the analysis of flood inundation dynamics in this area is of the upmost importance (Dinh et al., 2012).

According to the 2009 evaluation of the Institute for Climate Change Research (RICC) of Can Tho University, the rice fields that would be flooded during the second half of the rice season by 20 cm, continuously for 1 to 4 days, would be seriously damaged, resulting in a loss of 80 to 100 % of the yield. The same study analyses what would be the effect on a rice farm, of one day of continuous inundation, by determining the flood inundation extent on the exposed rice areas. A solution to protect these rice farms could be the construction of temporary dikes by farmers. These dikes have to be approximately 20 cm higher than the field surface in order to cope with floods. In cases where the water level is more than 50 cm above the land surface, farmers cannot build sufficient dikes to protect the rice fields, and therefore the mentioned study identified the areas which are potentially seriously harmed by large floods. In case of floods occurring in the early stage of the rice season, even though the rice can stay for a longer time inundated (1 to 14 days) (DRAGON, 2009), the continuous flood would delay the land preparation process for the next farming season. The present study looks at the flood events of short duration, and aims to identify what would be the impact of these events on the agricultural activities in the area under projected climate change conditions. The identified thresholds of 20 and 50 cm are used to map the flood events and to determine the areas exposed to such flood depths.

1.2 Climate change projections

Mekong is highly populated and in continuous development. It has a basin development plan (BDP) that looks at hydropower development and irrigation needs for the next 100 yr. There is a series of dams, constructed or planned to be constructed, on the upstream part of the Mekong River. The operation of dams and their impact on the discharge of the Mekong River, in conjunction with future climate changes, were analysed by MRC in a study in which future flows on the Mekong River and Mekong delta are predicted taking into consideration the combination of the future climate changes and the projected BDP in the basin (Hoanh et al., 2010). The study predicts future flow in six scenario situations by making a combination between the BDP and the SLR at the East and West Seas, as it was predicted by the Vietnam Ministry of Nature and Environment. The operation of dams and their consequent impact on the Mekong flow is part of the MRC study and is not debated in this paper.

The present study considers the MRC predicted flow at the upstream boundary of the VMD in two specific situations:

1. prediction of flow due to climate changes, considering that the actual situation of the development in the

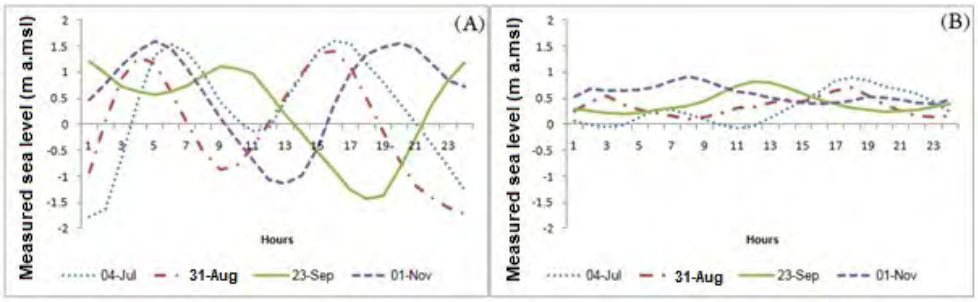

Fig. 2. Measured sea levels in the East **(A)** and West **(B)** Sea.

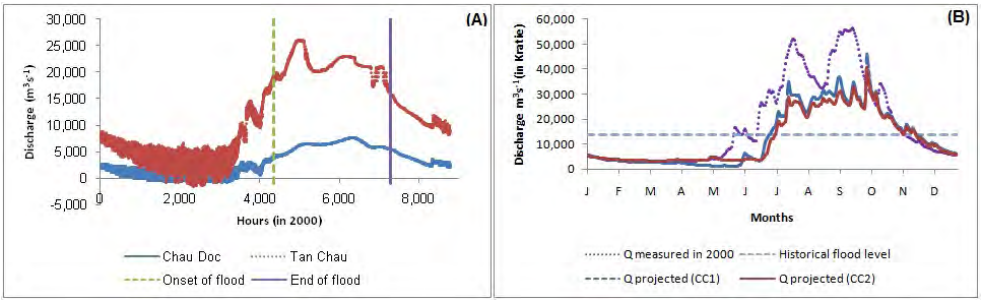

Fig. 3. Measured hourly discharge in 2000 at the Tan Chau and Chau Doc gauging stations **(A)**; The annual hydrograph measured in 2000, historical mean daily discharge (1985–2000) and projected annual hydrograph in 2050 (CC1 and CC2) **(B)**.

basin will remain unchanged for the year 2050. In this scenario the assumption is that any basin development plans (i.e. hydropower development and irrigation) will take place after 2050. This case is denominated in the present paper as climate change situation 1 (CC1);

2. prediction of flow due to climate changes, considering that the BDP will be implemented before the year 2050. This case is denominated in the present paper as climate change situation 2 (CC2);

The discharge hydrographs at the upstream of the VMD at Kratie station are shown in Fig. 3b.

In cases where climate change scenarios CC1 and CC2 are applied, these flows will be affected; however, it is considered that they are still subject to the international treaties, and the projected upstream VMD discharge for the year 2050 is the one predicted by MRC studies.

According to the CC1 climate change projection, the Mekong River discharge is greater than the one of the CC2 projection, because with future developments (hydropower and irrigation) in the upstream Mekong basin, less water would arrive in the VMD. The baseline for the climate change projections was the available data on daily discharge from 1985 to 2000 (Hoanh et al., 2010). Analyses of this data show that, in comparison to the mean daily discharge (1985–2000), the projected flood in 2050 would start earlier in time and last for a longer period of time, e.g. will terminate later than the nowadays flood. If the threshold flood discharge of $Q_{\text{Kratie}} = 13\,600\,\text{m}^3\,\text{s}^{-1}$ (MRC, 2007) is considered as a base case, then the flood in 2050 would start 7 and 14 days earlier for the CC1 and CC2, respectively, and end approximately 14 days later. In general, the projected daily discharge during the flood period in the first projection would remain similar to the historical mean daily discharge, while the projected discharge in the second projection would be lower than the historical mean discharge.

2 Model setup

Many modelling studies were conducted in the VMD by local institutions, international consultants and researchers. A comprehensive overview of these models is given by Hoanh et al. (2010) and Johnson and Kummu (2011). Depending on the problem to be solved, or the objective of the study (e.g. design and planning), different software tools were used to build instantiated models of the VMD. The most used software tools are the Mike 11, Sobek, local Vietnam models, and ISIS 1-D, which are physically based models that solve the 1-D Saint Venant equations for shallow water flows in a river (network) with a mild slope.

The model used in the present study is the ISIS 1-D model of the VMD, which was developed by HR Wallingford and Halcrow, and is maintained by the Mekong River Commission (MRC). The model is used by MRC to study the impacts of climate change on the flow dynamics in the Mekong Delta.

The results of the model are used by MRC to determine the boundary conditions for any detailed study of the Cambodian and Vietnamese Mekong Delta.

The VMD model setup

This study uses the hydrodynamic component of the ISIS 1-D model and not the hydrological component. The hydrograph at the upstream boundaries is determined based on measured discharge and on predicted discharge as per the climate change scenarios CC1 and CC2. The predicted hydrograph of scenario CC2 is in accordance with the international agreements between the countries in which the Mekong basin is located.

The instantiated 1-D ISIS model of the VMD represents the whole Mekong delta (the Cambodian and the Vietnamese part of the delta together) using 3036 cross-sections for a 8619 km network of channels. The upstream boundary has 25 nodes and the downstream one has 19 nodes. The model also represents 193 spills, 528 junctions, 409 reservoirs, 749 floodplain units, and 29 sluices. This model is used by the MRC to determine the annual flood in the VMD. Because of the complexity of the whole VMD river network, the model includes the Cambodian part of the delta as well.

During the peak flood period, apart from the flows entering the VMD from the main rivers (through Tan Chau and Chau Doc), the overbank flows on the Cambodian part are especially important sources of inflows (Fujii, 2003; MRC, 2007). These inflows are represented as distributed lateral flows at the most upstream part of the VMD model. The ISIS model was developed to represent the complex interactions caused by tidal influences (along the East and West Seas), flow reversal in the Tonle Sap River and overbank flow in the flood season, with the varying inflows from upstream. Even though the ISIS model is considered to be able to provide a reasonable representation of the hydrodynamics of the Cambodian floodplain and VMD, MRC suggests that the model should not be used for design purposes (MRC, 2010b), but rather to estimate the trend of changes when the boundary conditions are modified. The reason for such advise is that, for such a large (deltaic) scale model, the lack of details for a specific area will lead to under- and/or over-estimation of the future events, and consequently wrong calculations of the planned construction.

The present study used the ISIS 1-D model in four simulation scenarios as follows:

a. *Scenario 1*: in which the initial set-up of the model was realised based on the flood data from the year 2000 (the 20-yr return period of the annual flood volumes; MRC, 2007). The measured daily discharge at the Kratie gauging station was used as upstream boundary condition, and the hourly measured sea levels at East and West Seas were used as downstream boundary conditions.

Table 1. The setup of the model boundary conditions.

Models setup	Upstream boundary	Downstream boundary
Scenario 1	Discharge hydrograph of the year 2000	Sea water level in year 2000 for both West and East Seas
Scenario 2	Discharge hydrograph of the year 2000	Sea water level in year 2050 (SL2000 + sea level rise)
Scenario 3a	Projected discharge of the year 2050 (CC1)	Sea water level in year 2050 (SL2000 + sea level rise)
Scenario 3b	Projected discharge of the year 2050 (CC2)	Sea water level in year 2050 (SL2000 + sea level rise)

b. *Scenario 2*: the simulation with no changes of the upstream discharge hydrograph, but with changes of the downstream boundary conditions, i.e. the projected sea level rise of up to 30 cm on both East and West Seas.

c. *Scenario 3a*: the simulation considering the projected upstream hydrograph, in accordance with the defined climate change scenarios CC1 and the projected SLR of 2050 (Fig. 3b).

d. *Scenario 3b*: the simulation considering the projected upstream hydrograph, in accordance with the defined climate change scenarios CC2 and the projected SLR of 2050.

The considered scenario descriptions (e.g. the boundary conditions, at both upstream and downstream ends of the hydrodynamic model) are summarised in Table 1.

For each scenario different simulations were carried out with regard to the downstream boundary conditions. The upstream discharge was maintained as the one projected to be valid for the climate change scenarios CC1 and CC2. A change of $\pm 30\,\%$ and $\pm 15\,\%$ of the projected sea level in 2050 has been applied in order to account for uncertainty of the sea level rise at the downstream end. The sensitivity analysis helped in identifying how the flood dynamics will affect the areas under consideration (rice farms, populated areas, etc.) for the year 2050.

The results are presented for the upstream part of the delta, only in cases where the simulated flood depth is of 20 cm or more, because this is the established threshold for protection of rice farms. In the downstream part of the VMD, it was important to determine the areas where the time of inundation changed because of the effect this would have on timing of the two consecutive farming seasons.

3 Results and discussions

For each of the four scenarios, an analysis is done regarding the flood extent over the upstream and downstream section (i.e. the coastal areas) of the VMD. Due to the temporal distribution of the flood in the VMD, the spatial distribution of the flood on four different days is analysed. The days were selected based on historic flood records and are as follows:

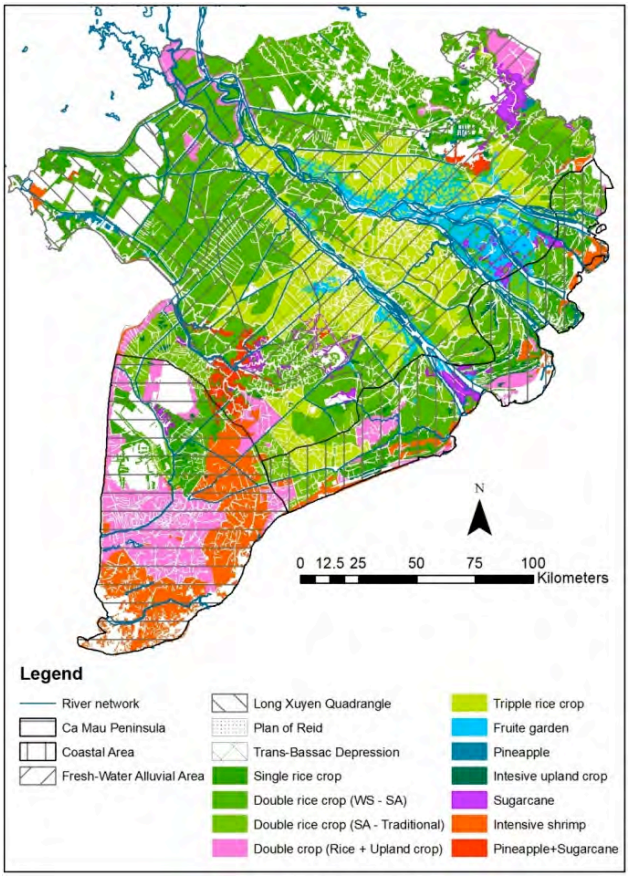

Fig. 4. Agro-ecological zones and land use map in the VMD in 2006.

- 4 July, the day of the early flood season when the flood starts causing damage on rice farming areas;
- 31 August, the day when the rice crop is, in general, harvested;
- 23 September, the day when the year 2000 flood peak was recorded (at Tan Chau and Chau Doc stations); and,
- 1 November, the day when the next rice farming season starts.

3.1 Scenario 1: the calibration of the model with the 2000 yr flood event

As a baseline case, the catastrophic flood of 2000 was considered, and simulated, and was used to calibrate the ISIS 1-D model. Figure 5 presents the maximum measured and simulated stages of one day during the flood period at different gauging stations (Tan Chau, Chau Doc, Vam Nao, My Thuan and Can Tho) in the VMD. The model overestimates the stage at all gauging stations. The maximum difference between the measured and simulated stage is 0.8, 0.5, 0.6, 0.5 and 0.6 m at Tan Chau, Chau Doc, Vam Nao, My Thuan and Can Tho, respectively. The observed model errors are comparable to those observed in other studies: 0.3 m (Le et al., 2008); 1.1 m (Västilä et al., 2010); and, 0.2 m (Dung et al., 2011); and within the limit of errors for hydrodynamic models in cases of severe floods (Hartanto et al., 2011; Moya Quiroga et al., 2012; Quinn et al., 2010).

Spatial distribution of the flood in the upstream section on 4 July, 31 August, 23 September and 1 November in 2000 are presented in Fig. 6 where the water levels greater than 20 and 50 cm are shown. As the upstream discharge is gradually increasing from 4 July to 23 September, so is the flood extent. It can be noticed that the upstream discharge on 23 September is greater than the one on 1 November, however the flood extent is larger on 1 November. This phenomenon is caused by the fact that sea level on the East Sea is higher in November than in September (Fig. 2a and b); therefore, in addition to the flood impact there is a sea backwater effect as well. In addition, high tide on the East Sea is greater than on the West Sea, while low tides at the East Sea are lower than those on the West Sea. The outcome is that the East Sea tidal regime is the main driving force for the flood dynamics in the VMD (both spatially and temporally).

Figure 7 presents the inundated period (in hours) on the four considered dates. It can be seen that the further inland in the VMD, the longer the inundation time. For high values of the discharge at the upstream part of the VMD, the inundation area is extended towards the East Sea.

3.2 Scenario 2: projected flooding due to SLR

The results of the simulation, presented in Fig. 8, show the inundation time differences between scenario 1 and scenario 2. The negative values are illustrating longer flood duration on 2050 as compared with the year 2000. The Ca Mau Peninsula is inundated a significantly longer time, about 4 months in total. This is due to the fact that this region is affected by the tidal regimes of the East and West Seas. The flood duration in the area along the West Sea is changed very little because there is no impact from the East Sea, and the Mekong and Bassac branches are routing the flood from inland to the East Sea.

3.3 Scenario 3: projected flooding of 2050

The results of the simulations under scenario 3a and scenario 3b are presented separately for the upstream and downstream part of the VMD in order to identify the effect of the boundary conditions on the VMD.

Figures 9 and 10 present the inundation extent on the upstream section of the VMD on the four dates selected for analysis. The maps are showing inundation depths of 20 and 50 cm for the CC1 and CC2 scenarios, respectively. Similar to the simulated results of the flood extent in 2000, the flood extent in 2050 would increase from July to November. This is the effect of the backwater (as it is the case for the year

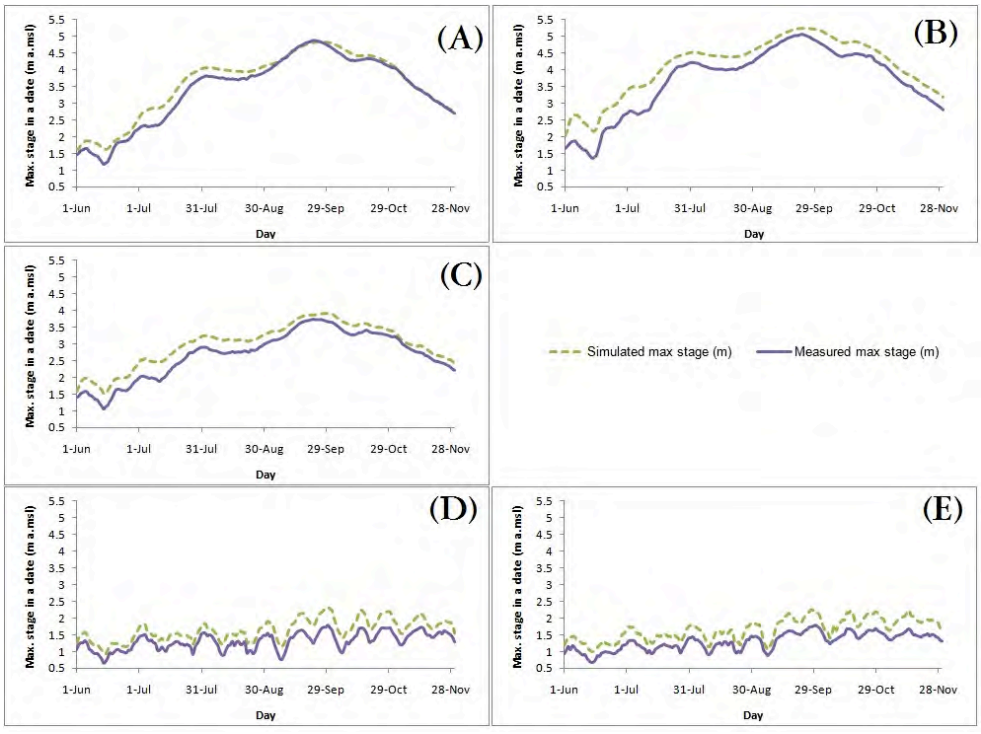

Fig. 5. Daily measured and simulated maximum stages during the flood period in 2000 at different gauging stations, **(A)** Chau Doc; **(B)** Tan Chau; **(C)** Vam Nao; **(D)** Can Tho; and, **(E)** My Thuan.

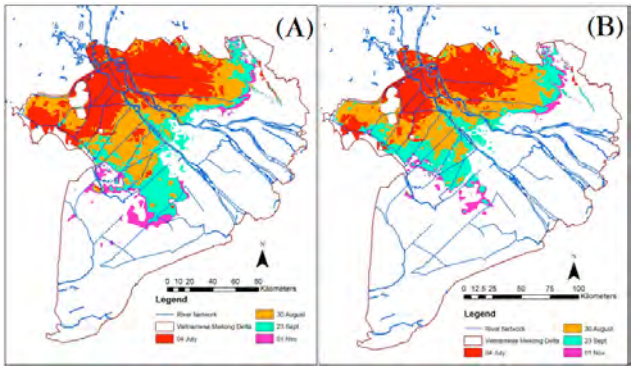

Fig. 6. Flood maps of year 2000, on 4 July, 31 August, 23 September and 1 November, inundation depth of 20 cm **(A)** and 50 cm **(B)**.

2000), combined with the fact that the upstream hydrograph changed its pattern (e.g. the peak of the hydrograph shifted towards the end of October, Fig. 3b). In comparison to the simulated results in 2000, the inundated area in 2050 would extend further to the East Sea.

Figure 11 presents the flood extent for depths greater than 20 cm for the coastal area of the VMD according to scenario 3a. The results of scenario 3b show no change as compared with cenario 3a. The flooding of the VMD remains mainly driven by the sea level rise along the East and West Seas, and the flood near the coastal area would be, as expected, greater in 2050. However, further inland the inundated area in 2050 would be smaller than the one in 2000, as the upstream discharge in 2050 would be lower than the one in 2000.

The flood on 2050 lasts longer along the coast, and shorter on the upstream section of the VMD (Fig. 12). In Fig. 12 the negative and positive values represent the longer and shorter flood durations in the year 2050 as compared with 2000.

The expected effects of longer inundation times in 2050 are the lower inundation depths. Figure 13 presents the differences between the simulated flood duration in 2050 according to the first and second projection; the positive value represents the longer flood duration in the first projection in comparison to the second one. Without development in the upstream section of the Mekong, the inundated period over the whole VMD could be extended from one day to about a month in total, in which the majority of increases would be around four days.

A similar analysis, as that for the 20 cm inundation depths in the year 2050, was done for inundation depths greater than 50 cm. The analysis concluded that the flood in the coastal area was not affected strongly by the relatively small modification of the upstream discharges; in the early flood season (July) the flood that was affected by the tidal regime reached the upstream section of the VMD (up to Tan Chau and Chau Doc). However, with higher floods the inundated area that is impacted by tidal regimes went gradually sea-wards (from

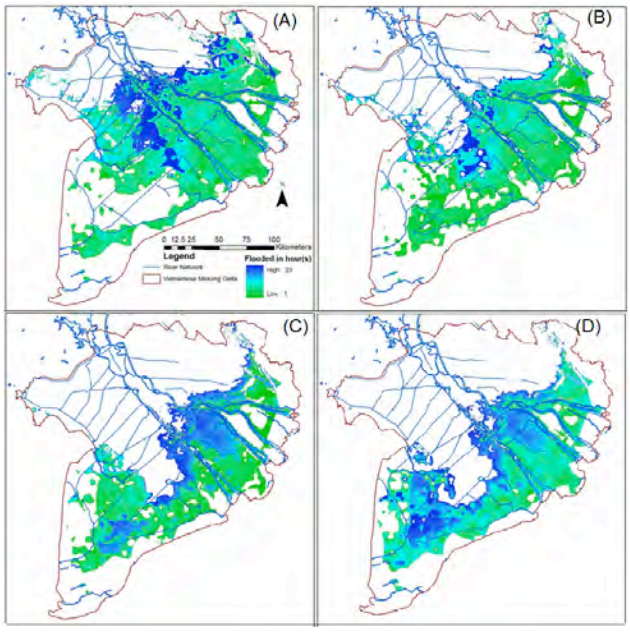

Fig. 7. The period of inundation (in hours) on 4 July **(A)**, 31 August **(B)**, 23 September **(C)** and 1 November **(D)** 2000.

July to November). In addition, due to the relatively poor drainage ability in the Ca Mau Peninsula, later in the flood season more permanent floods are projected; whereas the spatial distribution of flood patterns in the area surrounding the main outlets of the river network remained quite similar over the whole flood period.

4 Discussions

4.1 Climate change and its impacts on the VMD

When dealing with floods, studies on the impacts due to climate change should not consider the effect of a single event but the cummulatve impacts of combined events (e.g. upstream flood, local rainfall and high tide) (Keskinen et al., 2010). The impacts of climate change on the VMD should be addressed in different sections of the delta, such as the upstream vs. downstream (Gichamo et al., 2012; Quinn et al., 2010), the coastal zone vs. Ca Mau Peninsula and the east vs. west coast. The present study shows that in the Ca Mau Peninsula the flood is strongly influenced by both the East and West Seas; therefore, changes of the tidal regime in each sea will significantly change the future flood patterns. One of the main assumptions of the present study is the similarity of the changes of the sea level rise on the East and West Seas; hence, further study on this issue should be undertaken in order to make better projections for the future.

Within a small floodplain, a 1-D hydrodynamic model could result in acceptable accuracy in terms of simulated inundation extents and floodwave travel times (Hartanto et al., 2011). However, for a large river network with extensive

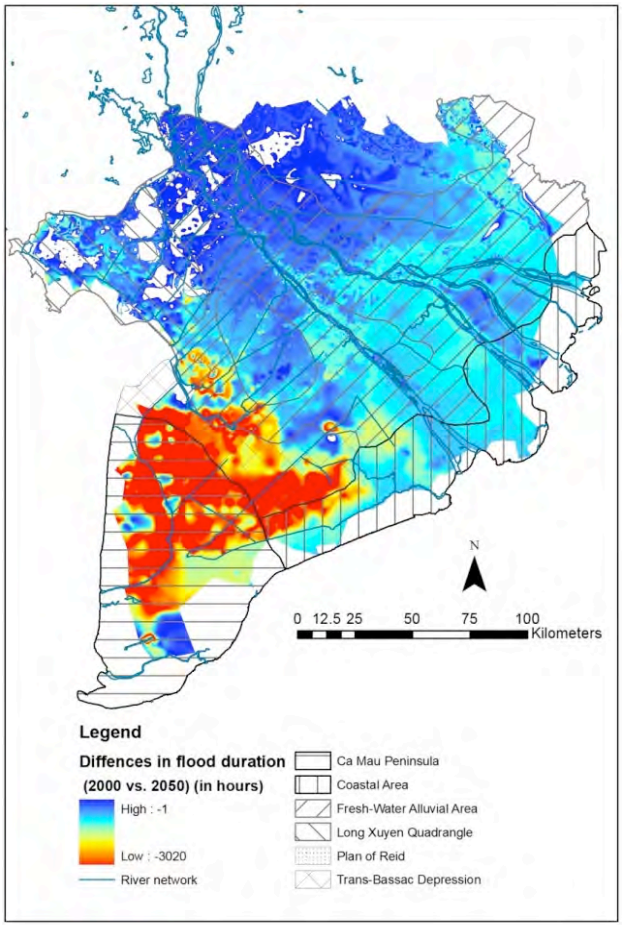

Fig. 8. Differences between the flood duration in 2000 and 2050 (CC2 2).

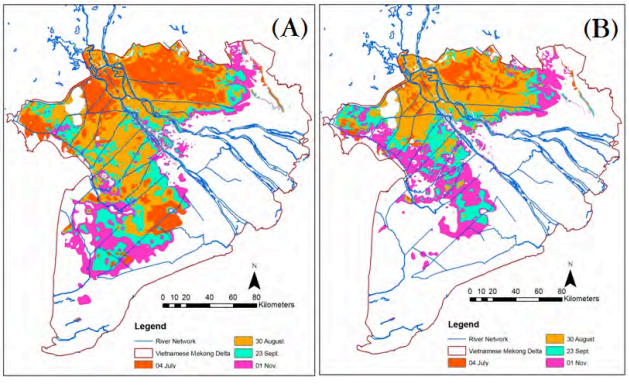

Fig. 9. The year 2050 flood maps, on 4 July, 31 August, 23 September and 1 November corresponding to the CC1 climate change scenario, for inundation depth of 20 cm **(A)** and 50 cm **(B)**.

floodplain, 2-D hydrodynamic models were developed to study the detailed hydraulic nature rather than the mean conditions in a 1-D hydrodynamic model (Gichamo et al., 2012). Even though there were positive aspects of such 2-D

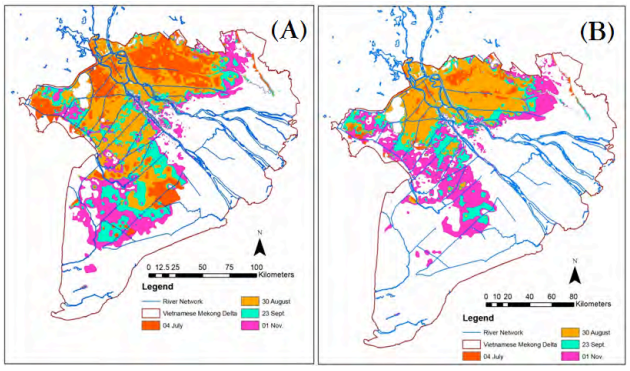

Fig. 10. The year 2050 flood maps, on 4 July, 31 August, 23 September and 1 November, corresponding to the CC2 climate change scenario, for inundation depth of 20 cm **(A)** and 50 cm **(B)**.

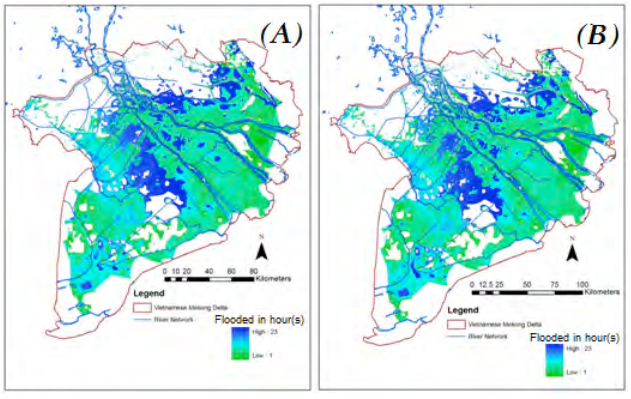

Fig. 11. The period of inundation (in hours) on 4 July, 2050 corresponding to CC1 **(A)** and CC2 **(B)**.

hydrodynamic models, due to the computational requirements as well as the intensive data needs, they were not very popular for large river networks like the one in the VMD area. The complexity of the present ISIS 1-D model was constrained by the availability of data (e.g. detailed bathymetry), and therefore modelling of a complex river network is still a great challenge due to significant differences in the hydraulic nature along the river network (Costelloe et al., 2006; Dinh et al., 2012; Gichamo et al., 2012). It is suggested that a coupled 1-D/2-D or a full 2-D hydrodynamic model should be developed in the future in order to study the flow and its propagation along the floodplain. A detailed 2-D model can be used for specific studies such as flood risk assessments on a small scale (Dutta et al., 2007; Muste et al., 2010; Pender and Neelz, 2007; Balica et al., 2012).

The Mekong River in the VMD is characterized by fluvial-unstable networks and the applied one-dimensional hydrodynamic model did not take into account the changes in river morphology. It is advisable that when the required data (e.g. sediment size) becomes available, a more comprehensive hydrodynamic model should be built in order to account for the changes of river morphology.

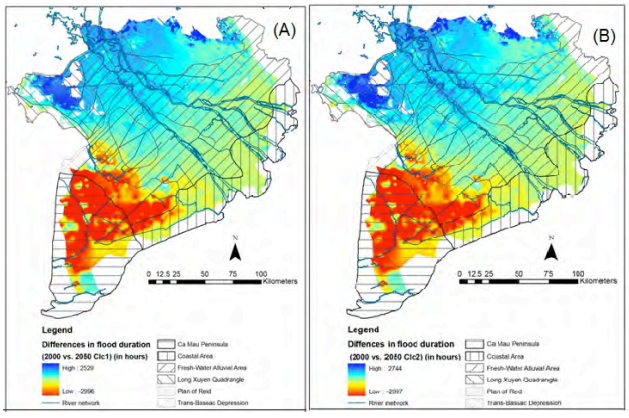

Fig. 12. Differences in flood durations in 2000 and 2050 corresponding to the climate change Scenario CC1 **(A)** and CC2 **(B)**.

4.2 Impacts of the flood on the agriculture in the VMD

It is expected that the future climate changes will significantly impact the agricultural activities but it is difficult to quantitatively assess such influences (Dinh et al., 2012). Västilä et al. (2010) stated that the damage to crops and infrastructure in the VMD will be serious, as the average and maximum water levels and flood duration will increase in the period 2010–2049. The findings of the present research also support this argument. One may expect that with the rise of the sea level only the agricultural system (including the rice-based farming systems and shrimp farming) in the Ca Mau Peninsula will be strongly affected, while the deep flood extent in the upstream section does not change significantly as compared to the 2000 flood.

Because of changes to future annual hydrograph patterns, at the upstream section of the Mekong there will be considerable impacts on the flood extent and water level in the VMD. Consequently, these will lead to significant impacts on the agricultural activities. Moreover, because we could notice a forward in time shift of the flood season, the (rice) cropping activities need to be adjusted in order to avoid the predicted November flood. The timing issue of flooding is critical for agriculture, as can be concluded from the effects of the 2000 extreme flood which caused severe damage not only due to the high discharge and stage, but also due to its early arrival (4 to 6 weeks earlier in comparison to the normal flood) (A. T. Le et al., 2007).

Simulated results show that floods would extend further to the Ca Mau Peninsula with the sea level rise. In fact, with the dense river network in the coastal area and greater differences between the low and high peaks of the tidal regime in the East Sea (in comparison to the West Sea), the floods were routed to the East Sea faster than to the West Sea. In addition, the hydraulic regime in the Ca Mau Peninsula was complicated as it was influenced strongly by both the East and West Seas.

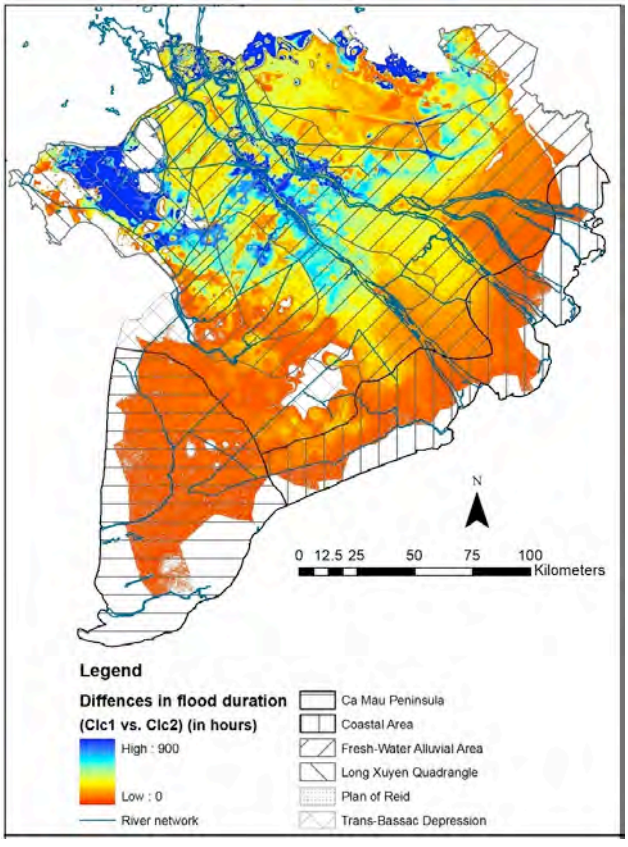

Fig. 13. Differences in flood durations in 2050 corresponding to the climate change scenarios CC1 and CC2.

In specific consideration to agriculture during the flood period in the VMD, agricultural activities in the upstream section of the VMD will benefit as the physical conditions will be more favourable for agriculture due to the "lower flood" conditions. However, along the coastal area, especially in the Ca Mau Peninsula, the flood would be prolonged, which might cause changes in the structure of the current farming systems.

Potential adaptation measures in the VMD, where agriculture is the main economic sector of development, climate change adaptation in agriculture is a point of great interest. With different policies to ensure the rice-based farming systems, permanent dikes along the upstream section of the VMD were raised to effectively protect the rice-based farming systems from the annual floods (Nguyen et al., 2007). However, with the drawbacks of such a permanent dike system (e.g. environmental pollution and degradation of soil fertility) (Nguyen et al., 2007), semi-dike systems (MRC/WUP-FIN, 2006) in combination with water retention (Buijse et al., 2002; Hartanto et al., 2011; Hooijer et al., 2004; Meire et al., 2010; Platteeuw et al., 2010; Popescu et al., 2010) could be considered as a suitable solution to maintain the agricultural production in the VMD. Actually, in the past (decades ago), floods were kept in backswamp areas in order to not only supply freshwater flow in main canals during the early dry season (Dang et al., 2007), but also to reduce the flood during the peak periods. For the irrigated rice-based farming systems the possible adaptation measures might be applied in the VMD, including (i) adjusting the cropping seasons could be a suitable adaptation measure to avoid the peaks of the flood (ADB, 2010; Mainuddin and Kirby, 2009); (ii) shifting from rice intensification systems into livelihood diversification; (iii) alternating wetting and drying irrigation methods (Belder et al., 2004); and (iv) diversifying cropping patterns (Hoanh et al., 2009). Moreover, to study the nature of the future flood in the VMD, climate change aspects (upstream discharge changes and sea level rise) should not be considered separately from other factors (e.g. economic and social environment) (Keskinen et al., 2010). It is important to notice that the physical changes would normally not be happening so suddenly (e.g. sea level would rise gradually over several decades, Wassmann et al., 2004). However, the anthropological factors contribute significantly to dramatic changes, like the dam construction along the upstream section of the Mekong (Yang et al., 2006; Kummu and Varis, 2007; Kummu et al., 2010), and those should be considered carefully and seriously before any action is turned into practice. In fact, "adaptation to climate change should be integrated in social economic planning at all scales" (Quinn et al., 2010; Muste et al., 2010).

Actions in order to inform the stakeholders in the area about the effect of floods and the adaptation measures could include, along with the modelling and prediction of such events, the use of modern technologies such as mobile phones (Jonoski et al., 2012a, b; van der Berg and Poepscu, 2005).

5 Conclusions

Fluvial floods in the VMD can be categorized into two groups: (i) floods in the upstream section mainly affected by the upstream discharge; and (ii) floods in the coastal area strongly influenced by the East and West Sea tidal regimes.

Flood maps (in the upstream section and coastal areas in the VMD) were analysed for the most four critical dates in a flood season. For the highest recent flood in 2000 the affected areas were mainly in the upstream part of the delta. For the year 2050, the flood hazards are expected to be more severe along the coastal area because of the tidal regime (e.g. high-tide induced fluvial floods, low tides allow drainage of the river flow to the sea).

It can be concluded that the future floods would have lower inundation depths but will last longer in the upstream section of the VMD. However, due to the change of the downstream boundary conditions (e.g. SLR), the inundated area along the Eastern and Western coasts, especially for the Ca Mau Peninsula, would be larger than the present one. Due to the dense

river network in the coastal area and low sea level during the low tides, the flood duration in coastal area would not be long, and often will last less than 24 h even if the sea level rises.

Our analysis shows that in the future (year 2050), agriculture in the upstream section of the VMD may be expected to have more favourable conditions, as the fluvial flood would be shifted towards the lower part of the VMD. In the downstream section (along the coastal zone and the Ca Mau Peninsula), longer inundation times might be a good environment for rice cultivation. In fact, the high stage would be higher due to the projected increase of sea level but such floods would not last long due to the impacts of the tidal regime. However, as this study is on floods, we cannot draw definitive conclusions on the impacts of climate change and upstream development on rice cultivation, as changes in the water availability in the dry season have not been studied.

In the end the authors would like to emphasize that all results are based on the available model-based climate change projections, which are highly uncertain, and therefore the flood maps and the conclusions should be considered with care, as indicated throughout.

Acknowledgements. The research presented herein constitutes a part of the Post-doctoral Programme on Climate Change Adaptation in the Mekong River Basin (PRoACC) funded by the Netherlands Ministry of Development Cooperation (DGIS) through the UNESCO-IHE Partnership Research Fund. The study was carried out jointly by UNESCO-IHE Institute for Water Education, the Netherlands, and Can Tho University, Vietnam, with the participation of the Mekong River Commission. It has not been subjected to peer and/or policy review by the mentioned institutions, and, therefore, does not necessarily reflect the views of these institutions.

The authors would like to thank the Mekong River Commission (MRC) for kindly providing and allowing the use of the ISIS hydrodynamic model in order to carry out the study.

Edited by: A. Montanari

References

ADB: Technical Assistance ReportNam: Climate Change Impact and Adaptation Study in the Mekong Delta, 2009.

ADB: Addressing Climate Change in Asia and the Pacific: Priorities for Action, 2010.

Balica, S. F., Popescu, I., Beevers, L., and Wright, N. G.: Parametric and physically based modelling techniques for flood risk and vulnerability assessment: a comparison, Environ. Modell. Softw., doi:10.1016/j.envsoft.2012.11.002, in press, 2012.

Belder, P., Bouman, B. A. M, Cabangon, R., Guan, L., Quilang, E. J. P., Yuanhua, L., Spietz, J. H. J, and Tuong, T. P.: Effect of water-saving irrigation on rice yield and water use in typical lowland conditions in Asia, Agr. Water Manage., 65, 193–210, 2004.

Black, A. R. and Burns, J. C.: Re-assessing the flood risk in Scotland, Sci. Total Environ., 294, 169–184, 2002.

Buijse, A. D., Coops, H., Staras, M., Jans, L. H., Van Geest, G. J., Grift, R. E., Ibelings, B. W., Oosterberg, W., and Roozen, F. C. J. M.: Restoration strategies for river floodplains along large lowland rivers in Europe, Freshwater Biol., 47, 889–907, doi:10.1046/j.1365-2427.2002.00915.x, 2002.

Carew-Reid, J.: Rapid Assessment of the Extent and Impact of Sea Level Rise in Viet Nam, Climate Change Discussion Paper 1. Brisbane, Australia: ICEM – International Centre for Environmental Management, 2008.

Costelloe, J., Grayson, R., and McMahon, T.: Modelling streamflow in a large anastomosing river of the arid zone, Diamantina River, Australia, J. Hydrol., 323, 138–153, 2006.

Dang, K. N., Nguyen, V. B., and Nguyen, H. T.: Water use and competition in the Mekong Delta, Vietnam, Proceedings to Challenges to sustainable development in the Mekong Delta: Resional and national policy issues and research needs, 143–188, 2007.

Delgado, J. M., Apel, H., and Merz, B.: Flood trends and variability in the Mekong river, Hydrol. Earth Syst. Sci., 14, 407–418, doi:10.5194/hess-14-407-2010, 2010.

Dinh, N. Q., Balica, S., Popescu, I., and Jonoski, A.: Climate change impact on flood hazard, vulnerability and risk of the Long Xuyen Quadrangle in the Mekong Delta, International Journal of River Basin Management, 10, 103–120, 2012.

Dooge, J. C. I. and Napiorkowski, J. J.: The effect of the downstream boundary conditions in the linearized St. Venant Equations, Mech. Appl. Math, 40, 245–256, 1987.

DRAGON: Institute Vietnam, Climate change impacts and vulnerabilities assessment for Can Tho City (Technical report), Can Tho, 2009.

Dung, N. V., Merz, B., Bárdossy, A., Thang, T. D., and Apel, H.: Multi-objective automatic calibration of hydrodynamic models utilizing inundation maps and gauge data, Hydrol. Earth Syst. Sci., 15, 1339–1354, doi:10.5194/hess-15-1339-2011, 2011.

Dutta, D., Alam, J., Umeda, K., Hayashi, M., and Hironaka, S.: A two-dimensional hydrodynamic model for flood inundation simulation: a case study in the lower Mekong river basin, Hydrol. Process., 21, 1223–1237, 2007.

Ericson, J. ,Vorosmarty, C., Dingman, S.. Ward, L., and Meybeck, M.: Effective sea-level rise and deltas: Causes of change and human dimension implications, Global Planet. Change, 50, 63–82, 2006.

Fujii, H.: Hydrological roles of the Cambodian floodplain of the Mekong River, J. River Basin Management, 1, 253–266, 2003.

Gichamo, Z. G., Popescu, I., Jonoski, A., and Solomatine, D. P.: River Cross Section Extraction from ASTER Global DEM for Flood Modeling, Environ. Modell. Softw., 31, 37–46, 2012.

Gupta, A., Lim, H., Huang, X., and Chen, P.: Evaluation of part of the Mekong River using satellite imagery, Geomorphology, 44, 221–239, 2002.

Hartanto, I. M., Beevers, L., Popescu, I., and Wright, N. G.: Application of a coastal modelling code in fluvial environments, J. of Environmental Modelling and Software, 26, 1685–1695, 2011.

Hoanh, C. T., Phong, N. D., Gowing, J. W., Tuong, T. P., Ngoc, N. V., and Hien, N. X.: Hydraulic and water quality modeling: a tool for managing land use conflicts in inland coastal zones, Water Policy, 11, 106–120, 2009.

Hoanh, C. T., Jirayoot, K., Lacombe, G., and Srinetr, V.: Impacts of climate change and development on Mekong fow regime. First assessment – 2009, MRC Technical Paper No. 29, Mekong River Commission, Vientiane, Lao PDR, 2010.

Hooijer, A., Klijn, F., Bas , G., Pedroli, M., and Van Os, A.: Towards sustainable flood risk management in the Rhine and Meuse river basins: synopsis of the findings of IRMA-SPONGE, River Res. Appl., 20, 343–357, 2004.

Johnston, R. and Kummu, M.: Water Resource Models in the Mekong Basin: A Review, Water Resour. Manage., 26, 429–455, 2011.

Jonoski, A., Alfonso, L., Almoradie, A., Popescu, I., van Andel, S. J., and Vojinovic, Z.: Mobile phone applications in the water domain, Environ. Eng. Manag. J., 11, 919–930, 2012a.

Jonoski, A., Almoradie, A., Khan, K., Popescu, I., and Andel, S. J.: Google Android Mobile Phone Applications for Water Quality Information Management, J. Hydroinform., doi:10.2166/hydro.2012.147, in press, 2012b.

Keskinen, M., Chinvanno, S., Kummu, M., Nuorteva, P., Snidvongs, A., Varis, O., and Vastila, K.: Climate change and water resources in the Lower Mekong River Basin: putting adaptation into the context, Journal of Water and Climate Change, 1, 103–117, 2010.

Kummu, M. and Varis, O.: Sediment-related impacts due to upstream reservoir trapping, the Lower Mekong River, Geomorphology, 85, 275–293, 2007.

Kummu, M., Lu, X., Wang, J., and Varis, O.: Basin-wide sediment trapping efficiency of emerging reservoirs along the Mekong, Geomorphology, 119, 181–197, 2010.

Le, A. T., Chu, T. H., Miller, F., and Bach, T. S.: Flood and salinity management in the Mekong Delta, Vietnam, in: Challenges to sustainable development in the Mekong Delta: Resional and national policy issues and research needs, edited by: Tran, T. B., Bach, T. S., and Miller, F., 2007.

Le, T. V. H., Nguyen, H. N., Wolanski, E., Tran, T. C., and Haruyama, S.: The combined impact on the flooding in Vietnam's Mekong River delta of local man-made structures, sea level rise, and dams upstream in the river catchment, Estuar. Coast. Shelf S., 71 110–116, 2007.

Le, T. V. H, Shigeko, H., Nguyen, N. H., and Tran, T. C.: Infrastructure effects on floods in the Mekong River Delta in Vietnam Area, Hydrol. Process., 22, 1359–1372, 2008.

Lespinas, F., Ludwig, W., and Heussner, S.: Impact of recent climate change on the hydrology of coastal Mediterranean rivers in Southern France, Climatic Change, 99, 425–456, 2009.

Lu, X. X. and Siew, R. Y.: Water discharge and sediment flux changes over the past decades in the Lower Mekong River: possible impacts of the Chinese dams, Hydrol. Earth Syst. Sci., 10, 181–195, doi:10.5194/hess-10-181-2006, 2006.

Mainuddin, M. and Kirby, M.: Spatial and temporal trends of water productivity in the lower Mekong river, Agr. Water Manage., 96, 1567–1578, 2009.

Meire, D., De Donker, L., Declercq, F., Buis, K., Troch, P., and Verhoeven, R.: Modelling river-floodplain interaction during flood propagation, Nat. Hazards, 55, 111–121, 2010.

Moya Quiroga, V., Popescu, I., Solomatine, D., and Bociort, L.: Cloud and cluster computing in uncertainty analysis of integrated flood models, J. Hydroinform., 15, 55–69, doi:10.2166/hydro.2012.017, 2012.

MRC: Annual Mekong Flood Report 2006, Vientiane, Laos, 2007.

MRC: Annual Mekong Flood Report 2009, Phnom Penh, Cambodia, 2010a.

MRC: Improved flood risk reduction through implementation of structural measures flood proofing, 2010b.

MRC/WUP-FIN: Hydrological, environmental and socio-economic modelling tools for the Lower Mekong Basin impact assessment – Mekong Delta socio-economic analysis, WUP-FIN Phase II – Hydrological, environmental and socio-economic modelling tools for the Lower Mekong Bas. Vientiane, Lao PDR, 2006.

Muste, M., Quinn, P. F., Hewett, C. J. M., Popescu, I., Basu, N. B., Kumar, P., Franz, K., Merwade, V., Arnold, W., and Potter, K.: Initiation of the Upper Mississippi River Basin Observatory, ASCE proceedings, Innovations in Watershed Management under Land Use and Climate Change, 39, 1270–1281, 2010.

Neuhold, C., Stanzel, P., and Nachtnebel, H. P.: Incorporating river morphological changes to flood risk assessment: uncertainties, methodology and application, Nat. Hazards Earth Syst. Sci., 9, 789–799, doi:10.5194/nhess-9-789-2009, 2009.

Nguyen, A. D. and Savenije, H. H. : Salt intrusion in multi-channel estuaries: a case study in the Mekong Delta, Vietnam, Hydrol. Earth Syst. Sci., 10, 743–754, doi:10.5194/hess-10-743-2006, 2006.

Nguyen, D. C., Le, T. D., Nguyen, V. S., and Miller, F.: Chapter 2: Livelihoods and resource use stratefies of farmers in the Mekong Delta, in: Challenges to sustainable development in the Mekong Delta: Resional and national policy issues and research needs, 69–98, 2007.

Nuorteva, P., Keskinen, M., and Varis, O.: Water, livelihoods and climate change adaptation in the Tonle Sap Lake area, Cambodia: learning from the past to understand the future, Journal of Water and Climate Change, 1, 87–101, 2010.

Pender, G. and Neelz, S.: Use of computer models of flood inundation to facilitate communication in flood risk management, Environmental Hazards, 7, 106–114, 2007.

Platteeuw, M., Foppen, R., and van Eerden, M. R.: The need for future wetland bird studies: Scales of habitat use as input for ecological restoration and spatial water management, Ardea, 98, 403–416, 2010.

Popescu, I., Jonoski, A., van Andel, S. J., Onyari, E., and Moya Quiroga, V. G.: Integrated modelling for flood risk mitigation in Romania: case study of the Timis-Bega river basin, International Journal of River Basin Management, 8, 269–280, 2010.

Prudhomme, C., Jakob, D., and Svensson, C.: Uncertainty and climate change impact on the flood regime of small UK catchments, J. Hydrol., 277, 1–23, 2003.

Quinn, P., Hewett, C., Popescu, I., and Muste, M.: Towards New Types of Water-centric Collaboration: Instigating the Upper Mississippi River Basin Observatory Process, Water Management, 163, 39–51, 2010.

Reichel, G. and Nachtnebel, H. P.: Suspended sediment monitoring in a fluvial environment: Advantage and limitations applying an acoustic doppler current, Water Research, 28, 751–761, 1994.

Västilä, K., Kummu, M., Sangmanee, C., and Chinvanno, S.: Modelling climate change impacts on the flood pulse in the Lower Mekong floodplains, Journal of Water and Climate Change, 1, 67–86, 2010.

van der Berg, C. C. and Popescu, I.: An experience in knowledge mapping, Journal of Knowledge Management, 9, 123–128, 2005.

Wassmann, R., Nguyen, X. H., Chu, T. H., and To, P. T.: Sea level rise affecting the Vietnamese Mekong Delta: Water elevation in the flood season and implications for rice production, Climatic Change, 66, 89–107, 2004.

WWF: The Greater Mekong and climate change: Biodiversity, ecosystem services and development at risk, 2009.

Yang, Z., Wang, H., Saito, Y., Milliman, D., Xu, K., Quiao, S., and Shi, G.: Dam impacts on the Changjiang (Yangtze) River sediment discharge to the sea: The past 55 years and after the Three Gorges Dam, Water Resour. Res., 42, W04407, doi:10.1029/2005WR003970, 2006.

Hydrol. Earth Syst. Sci., 16, 3935–3944, 2012
www.hydrol-earth-syst-sci.net/16/3935/2012/
doi:10.5194/hess-16-3935-2012

Hydrology and Earth System Sciences

Experiences from online and classroom education in hydroinformatics

I. Popescu, A. Jonoski, and B. Bhattacharya

UNESCO-IHE Institute for Water Education, Delft, The Netherlands

Correspondence to: I. Popescu (i.popescu@unesco-ihe.org)

Received: 24 December 2011 – Published in Hydrol. Earth Syst. Sci. Discuss.: 24 January 2012
Revised: 28 September 2012 – Accepted: 10 October 2012 – Published: 1 November 2012

Abstract. Universities and other higher education institutions involved in water-related engineering education are facing new challenges in offering lifelong learning services and online educational support. Both the curricula and the form of delivery are changing, as contemporary water problems require interdisciplinary approaches involving diverse and up to date expertise maintained via continuous professional development. Hydroinformatics education faces similar challenges in developing relevant curricula and finding appropriate combinations of course delivery to its target group. This article presents experiences from delivering two hydroinformatics courses in the fields of flood modelling for management (FMM) and decision support systems (DSS) in river basin management that in recent years have been delivered both online and in classroom settings. Comparisons between the two modes of delivery are provided, with the conclusion that online education in this field, although still faced with many challenges, has a promising potential for meeting future educational needs.

1 Introduction

Contemporary water-related engineering projects are characterised by high complexity and clear necessity of interdisciplinary approaches, which in turn requires a broader academic education as well as continuous professional development of modern-day engineers. Consequently, teaching demands are rapidly changing and they are currently significantly different compared to only five or ten years ago. This is a result of the diversification of disciplines, the shortage of skills in key areas and the needs for employability of the graduates.

In order to meet these new requirements and to make the European Higher Education Area attractive to students from all over the world, twenty-nine European countries signed the Bologna agreement in 1999. One of the most important changes introduced with this agreement is the European Credit Transfer System (ECTS), levelling education in Bachelor's and Master's, with 180 to 240 ECTS for Bachelor and 60 to 120 ECTS for Master level. Bachelor's and Master's Degrees are formulated with clearly defined learning outcomes and associated competencies, which can also serve for comparison of higher education among different universities and countries. The Bologna declaration also recognised that, in addition to the necessary changes in higher education, there is a clear need for lifelong learning and professional development, which is especially relevant for engineering education (Gonzalez and Wagenaar, 2003).

Water problems usually cut across boundaries, both geographical and professional, increasingly requiring alliances that link different professionals at many locations. Appropriate education for these new working modes becomes important for enhancing the capacity to manage water-related assets and the aquatic environment.

UNESCO-IHE (Delft, the Netherlands) is an academic, water education institute providing MSc and PhD education, as well as a large number of short courses and online courses (UNESCO-IHP, 1999). Since the institute primarily targets mid-career water professional, there is a clear recognition of the continuous tension between academic educational offering (embedded in larger degree programmes) and the requirements of professional competence development of those who work in a professional water sector environment (frequently for specialised, stand alone learning components). To better address the new demands students are increasingly

being offered more flexibility in their future learning paths by providing some key components of water-related education through online courses. The advantages of online courses are that they are relatively cheaper (smaller tuition fees for the students), they provide flexible learning schedules, travel is not involved, families are not separated, etc. This approach in fact follows the general trend in education of gradually moving more towards asynchronous sharing of knowledge and reduction of direct contact with teachers (Price et al., 2007).

One of the fields where two online courses were developed at UNESCO-IHE is hydroinformatics. Two online courses in this field, which are the primary focus of this article, are flood modelling for management (FMM) and decision support systems (DSS) in river basin management (Jonoski and Popescu, 2012). These same courses are available as regular components of the Master studies in hydroinformatics delivered in a face to face setup. The transformation of these regular components into online courses requires significant efforts in both developing the learning material and supporting the online learning process of diverse groups of learners. Given that these are still relatively new developments in hydroinformatics education, this article presents the experiences from face to face (classroom) and online delivery of these two courses, which is useful for improving the effectiveness of future online learning. It needs to be noted that the face to face and online implementations of these courses run in parallel, however for different types of learners. They are consequently evaluated separately by the enrolled learners/students. So far, for these courses there is no structured pre-designed evaluation for *direct* comparison of the two modes of delivery. Nevertheless, the experiences presented herein provide some important lessons for future online learning in the field of hydroinformatics.

After this first introductory part the second section of the article introduces briefly the hydroinformatics field and its education implementation at UNESCO-IHE. The face to face and online implementations of the two courses are presented in Sects. 3 and 4, followed by discussions and conclusions in Sects. 5 and 6.

2 Hydroinformatics education

2.1 The field of hydroinformatics and current educational approaches

The concepts of Hydroinformatics as a new and distinct academic discipline were conceived and implemented by Professor Michael B. Abbott, about two decades ago (see, e.g. Abbott, 2001). Hydroinformatics has since been widely recognised internationally, attracted a successful series of biennial international conferences and has a peer reviewed journal.

Broadly, hydroinformatics can be defined as "the study of the flow of information and the generation of knowledge related to the dynamics of water in the real world, through the integration of information and communication technologies for data acquisition, modelling and decision support, and to the consequences for the aquatic environment and society and for the management of water based systems" (Abbott, 1991).

This definition includes merging of traditional fields of computational hydraulics with newer developments in numerical analysis, computer science, and communications technology. The fundamentals of hydroinformatics are in the water domain and they integrate knowledge and understanding of both water quantity and quality (Price et al., 2006). Furthermore, the field focuses on the integration of information about various water systems and the aquatic environment obtained from diverse sources, such as data from the field, remotely sensed data, and data from various hydraulic, hydrological and other kinds of numerical models. Aiming at the provision of decision support, typical hydroinformatics applications need to provide further integration with information and data coming from non-engineering fields like ecology, economy and social science. With these established goals, in recent years hydroinformatics has been transforming from a purely technical, into a sociotechnical discipline (Abbott and Jonoski, 1998; Jonoski, 2002).

The diversity of the involved technologies requires structured and goal-oriented development of hydroinformatics applications, which depends on appropriate educational approaches. Most adequate hydroinformatics expertise is commonly provided via educational programmes at MSc level (with duration of 1.5–2 yr) for BSc graduates from civil (hydraulic) or environmental engineering, computer science, or water-related earth sciences. Shorter hydroinformatics-related educational and training courses, often linked to the content of existing MSc programmes, are also provided. These are mostly targeting experts in need of continuous professional development (Kaspersma et al., 2012), dealing with rapidly evolving hydroinformatics technologies.

Hydroinformatics education at a Master level is provided by only few educational institutions in the world. The number of organisations that offer regular short courses or tailor-made courses in this field is higher, but still rather limited. Consequently, attending hydroinformatics courses in the traditional classroom setting may be quite expensive for prospective participants. Especially professionals from developing countries may have significant difficulties in securing funds to follow such a course. New methods of transferring hydroinformatics education are therefore considered, such as group learning via web-based collaborative engineering (Molkenthin et al., 2001) or online courses, such as those presented in this article.

The example MSc specialisation in hydroinformatics from UNESCO-IHE can serve to introduce the common structuring of the content delivered in a typical hydroinformatics programme. The programme starts from the classical approach of developing mathematical models, based on first principles as a means for solving real engineering problems,

and continues by introducing methods from new modelling paradigms such as data-driven or agent-based modelling. The introduction of information technology and software engineering topics is in parallel to these modelling topics. The different modelling approaches, together with their advantages and disadvantages are then demonstrated on different application areas. The structure of the programme together with the targeted application areas and associated tools and techniques are presented in Fig. 1.

This programme has a total duration of 18 months. The first 12 months cover the taught part of the programme structured in 14 educational components, so-called *modules*, each with duration of three weeks. During any particular module students are focusing on one group of thematically interrelated subjects. Many of these modules are also offered as stand-alone short courses for external participants who are not enrolled in the MSc programme. The last 6 months are reserved for individual research resulting in an MSc thesis. The time sequence of these educational components is presented in Fig. 2.

The presented content and structure, with minor modification from one academic year to another, is used in this programme for more than a decade. Other MSc programmes in hydroinformatics cover similar content, although with different structure. All these programmes, however, are faced with the continuous penetration of information and communication technologies in nearly all knowledge domains covered, and with their continuous and rapid changes. This situation requires rethinking and possibly adapting their contents and structure.

2.2 Introducing more flexibility in hydroinformatics education

A major challenge brought about by the evolution of hydroinformatics is that the volume of information that hydroinformaticians are called upon to know is increasing far more rapidly than the ability of engineering curricula to "cover it". The nature of this filed is such that graduates are required to master a broad spectrum of subjects, such as concepts from physics, mathematics, ecology, geography and computer and software engineering. This spectrum is well beyond the range of traditional hydraulic engineering curricula. Modelling and information and communication systems are at the core of hydroinformatics (which are quite specialised), but their adequate implementation for diverse application areas requires the coverage of a broad range of topics (Wagner et al., 2012).

For these reasons, structuring the curriculum that meets the needs of most hydroinformatics students appears to be an elusive goal. One solution is to institute multiple tracks for different areas of specialisation. In the example MSc programme presented in the previous section, this was realised by introducing a number of elective modules, covering a range of different topics. The courses that are subject of this article actually belong to these elective modules. For

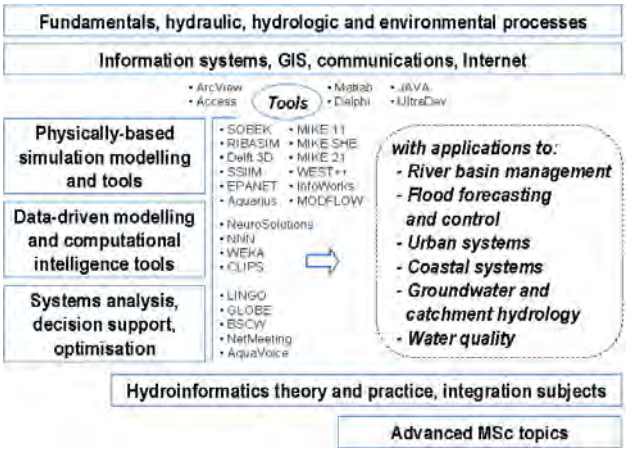

Fig. 1. General thematic structure of the Hydroinformatics MSc programme (taught part).

example, the "flood modelling for management" (FMM) is offered as an elective out of three other modules and the "decision support systems in river basin management" (DSS) is offered as an elective out of five other modules. These elective modules are offered as online courses as well. Once professionally engaged, hydroinformatics graduates can decide to enrol in such an online course (that they did not follow during their formal education) and improve their competences in a particular area. Equally, alumni of the institute or professionals who were not participants in an MSc programme, but they do want to acquire extra competencies can enrol in these online courses. With these arrangements the same content is delivered in a more flexible manner, meeting the needs of different learners interested in hydroinformatics topics.

A comparison of the pre-requisites, target group, learning outcomes and assessments between the two modes of course delivery is presented in Table 1. As can be seen from the table, the target audience in both online and face to face modes of delivery is the same, though in practice we observed that a large number of participants of the online courses were water professionals or PhD researchers who joined the course to improve their skills. In the face to face version of the courses all participants were MSc students, who might have had some professional experiences, but joined the course in the pursuit of obtaining an MSc degree.

3 The FMM and DSS face to face modules

The purpose of the FMM course is to provide professionals with the necessary background to appreciate the role and application of models in flood management, whereas the purpose of the DSS course is to present the general aspects of water resource management on the scale of the whole river basin (Jonoski and Popescu, 2012). These courses are modules in the regular Hydroinformatics MSc programme, but

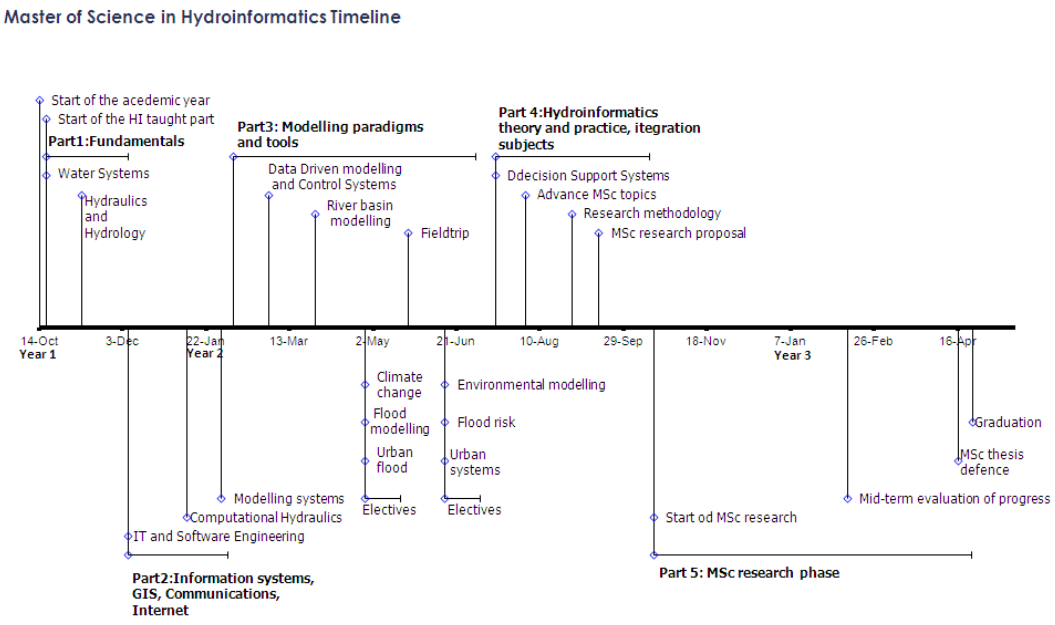

Fig. 2. Time sequence of the modules in the taught part and the research period of the MSc programme in Hydroinformatics.

Table 1. Comparison of key characteristics in the two modes of course delivery.

Component	Face to face delivery of courses	Online delivery of courses
Prerequisites		
Pre-requisites/target group	Same for both modes of delivery	
Time and study load		
Time of course delivery	3 weeks (Study load 6–7 h day^{-1})	10 weeks (2 h day^{-1})
Estimated study load	140 h	140 h
Learning objectives	Same for both courses	
Assessment mode	Assignment and examination	Assignment
European credit transfer load (ECTS)	5	5
Certificate for the course	No	Yes
Advantages/Disadvantages		
Pace of learning	Imposed by the timeframe of the course (disadvantage)	At students own choice (advantage)
Leave from work	Student has to leave the work and go to study (disadvantage)	Student can study while working (advantage)
Direct interaction with instructor	Immediately (advantage)	Not in the same time (disadvantage)
Direct interaction with instructor	Answer given in an oral format (disadvantage of not being able to reconsult the answer)	Instructor answers in a forum (advantage of having the answer for reviewing)

they are also offered to students from other relevant MSc specialisations. FMM and DSS modules are carried out for three weeks in module 8 and 11, respectively, and have curricula as defined in Tables 2 and 3. Due to its learning objectives, the DSS module is taught at a later stage, compared to the FMM module. In Tables 2 and 3, the "contact hours" represent the real contact hours between the lecturer and the students, while the "estimated study load" represents the estimated total number of study hours required for mastering a particular topic. The difference between study load hours and contact hours represents the number of hours that students need to allocate for studying without lecturers' supervision (self-study). The total number of ECTS is related to the number of study load hours of a module. Both DSS and FMM course have a load of five ECTS.

Table 2. The content and assessment of the face to face FMM course.

Subjects	Contact hours [h]			Estimated study load [h]	Assessment
	Lecture	Exercise	Workshop		
Introduction to some application domains of Hydroinformatics: floods, urban systems and the environment	4	0	2	14	
Climate change and its impact on hydrology	4	0	2	14	
Environmental processes and water quality	6	2	2	24	
Introduction to 1-D2-D, 2-D modelling	2		2	8	
Flood analysis, river flood modelling and 1-D flood routing	10	22	0	74	Exercise report (50 % weight)
New data sources to support flood modelling		2	2	6	
Total study load hours				140	Oral exam of all the subjects (50 % weight)

Table 3. The content and assessment of the face to face DSS course.

Subjects	Contact hours [h]			Estimated study load [h]	Assessment
	Lecture	Exercise	Workshop		
System analysis in water resources	8	6	6	42	Exercise report (35 % weight)
Decision support systems	6	4	4	30	Exercise report (30 % weight)
Software technologies for integration	4	8	8	36	Exercise report (20 % weight)
Integration of weather prediction and water models	8	2	4	32	Exercise report (15 % weight)
Total				140	

The learning objectives of the FMM module are to

- understand and explain the main flood management problems;
- understand and explain the governing processes of flood generation and propagation;
- identify the proper modelling methodology for a given problem;
- utilise hands-on experience in the step-by-step modelling procedure (dealing with geometry, bathymetry, boundary conditions, forcing) for carrying out a practical study with river modelling software packages; and
- know how a river flood model may be used for structural and non-structural measures for flood mitigation.

With respect to its target group, the FMM course is designed for current and future water professionals (engineers and scientists), decision makers and others involved in flood modelling and flood management, particularly those who would like to be familiarised with the latest tools and techniques in flood modelling management.

The learning objectives of the DSS module are to

- understand the role of system analysis in water resources planning and management;
- formulate and solve water resources problems as optimisation problems;
- distinguish and properly use different types of decision support methods for water problems;
- build simple software applications that integrate data and models, both as stand-alone and Internet-based; and
- understand the potential of newly available data sources (e.g. remote sensing, web resources, data generated from climate and meteorological models) in advanced integrated modelling and decision support.

The target group of the DSS module are professionals who are developing modelling and information systems support for managing water resources in river basins or are involved in decision making processes in river basins at different levels.

The face to face FMM module is also offered as a short course. A participant to this short course joins the regular MSc programme, just for three weeks during the period of the module. The module is designed in such a way that water professionals, flood managers, or professionals who would like to know more about this topic, can join the MSc programme during this period, given that they meet certain prerequisites. In this case basic knowledge of hydrology and hydraulics is required. Some experience with flood modelling/management is desirable but not a must.

4 The FMM and DSS online courses

4.1 Main challenges in development of the courses

The online modules on FMM and DSS are at postgraduate level, and they are with the same learning objectives and target groups as their face to face versions. The types of students are however different for the two modes of delivery. In the face to face modules, most participants are already enrolled MSc students and some short course participants, whereas in the online courses most students are water professionals from different countries of the world.

The first challenge in development and implementation of any online course is the development of suitable learning material. Traditional learning material, such as lecture notes, tutorials and exercises, already available for the face to face version of the courses, can be easily deployed for online learning. However, this is usually not sufficient for mastering the material by the distant learners. Enrichment by more effective learning material using suitable audio-visual material in electronic format is commonly needed. This requires a serious effort and the time required for developing such material should not be underestimated. For the courses presented here, for example, several man months of work were invested for this purpose.

Once suitable material is developed and deployed on a dedicated educational web-based platform, the advantages of the online courses can be realised. The course participants can download lectures, lecture notes, etc. at their own convenience and they can communicate with fellow participants and teachers using dedicated tools. For the FMM and DSS courses presented here the lectures are provided in the form of video, audio or slides with notes. A video lecture may contain slides with an audio of the teacher's commentary. Occasionally, it can contain a video window as well for displaying explanatory material (such as a movie). Theoretical lectures usually contain slides with notes. Soundtracks of lectures are incorporated in the presentation slides.

When it comes to the field of FMM and DSS, the development and implementation of online courses becomes further challenging, because these topics require not only the introduction of concepts, but also their exemplification by various software tools and systems, such as simulation models, encapsulated optimisation techniques, or MCA tools in case of the DSS course (Calizaya et al., 2010). Therefore, development and deployment of additional online learning material, specific for this purpose, may be required.

In the two implemented courses, modelling tools are introduced with instructional movies. Recent technologies allow such movies to be created easily and cheaply. The instructor presents the modelling tool by running it on his/her computer in a similar manner as in a face to face interaction. Current technologies can record the computer screen and add the instructor's voice to create the movie. Such movies can also be edited easily. Participants can download and follow these movies and become acquainted with the new tools. Because participants can run such a movie a number of times in order to master some specific features of a tool, experience shows that this is preferred to a classroom demonstration of a new tool. Nevertheless the inability to have live question and answer sessions (as in a classroom situation), which are frequently about concepts behind the introduced tools, leads to overall smaller appreciation of online learning when it comes to topics involving learning modelling tools.

This drawback can sometimes be overcome by using educational tools for synchronous discussions. Software tools such as Breeze or even Skype allow for arranging a real-time virtual classroom. During a virtual classroom, many-to-many communications in the form of video, voice and text enable everyone to see, talk and write (chat) to each other. Participants can share documents, computer screens or use a common electronic write board. In the implementation of the two courses presented here these virtual classrooms were not used, but they are being arranged at present times. It needs to be noted, however, that the number of participants when using such tools has to be limited.

Another important consideration related to online learning of modelling topics is the necessity of using public domain modelling tools. These are predominant in both the FMM and the DSS online courses presented here. Commercial modelling software is included for a few topics, but only using free, demo versions, without using any licence. Using full versions of such software can be easily arranged in a classroom setting, but the needed licences are commonly not available to distant learners. One alternative arrangement for overcoming this is the provision of web-based modelling simulations with the licensed software residing on the server side, which is also currently being considered.

Assessing students' performance is also an issue in online settings. In the face to face version of the presented course, the assessment consists of an oral examination, assignment reports on modelling exercises and classroom discussion. In an online version the oral examination cannot be conducted easily because it is difficult to organise a classroom for examination or to conduct an oral examination over the Internet. During initial runs of the courses an oral discussion was attempted using phone and Skype communication, but that turned out to be expensive and not easy to arrange. An alternative of involving other partner institutions in the country (or region) of the distant participant, where examinations can be organised is also currently considered. As a result, the current assessment in the FMM and DSS courses relies heavily on the evaluation of the assignment reports. This in fact necessitates re-adjusting of the learning objectives for the online courses, which are currently being contemplated.

4.2 Implementation of the courses

The FMM and DSS online courses were implemented as from 2006 and 2009, respectively. Since then they are run

regularly once a year. The average time of running an online course is 10 weeks and the content of each course is the same as the corresponding face to face version. The online courses are carrying a load of 5 ECTS and they require the same amount of study hours as in a normal face to face course.

The admitted participants in both courses have always been with a minimum of a BSc degree in fields such as civil, environmental or agricultural engineering, earth sciences, computer science or systems engineering. Acceptance to the course is based on several criteria such as educational background, working experience and motivation expressed in the application. The participants are commonly capable of searching for the needed information, and motivated to develop the respective expertise relevant for their professional occupation. In other respects they are a very heterogeneous group: participants come from all over the world, the number of years of experience in the profession differs, as well as their experience with online learning.

Access to the online courses is via a dedicated educational web-based platform, set up using the open source Moodle software package (http://moodle.org/). The implemented platform is accessible through the UNESCO-IHE's e-learning campus at the following web address: http://ecampus.unesco-ihe.org. Each admitted participant to an online course receives a user name and password for accessing the learning resources on the platform.

The learning resources of an online course are structured into units, where each unit addresses a certain topic. An example implementation of a unit dedicated to multi-criteria analyses in the DSS course is presented in Fig. 3. The advantage of such structuring using the Moodle platform is that the educational material is decomposed in its basic components and allows different people to use the same source material while each participant can have a different focus on individual units depending on their own expertise and insight.

As already mentioned, the lectures are provided in the form of videos, slideshows with audio, or slideshows with text explanations. Additional reading materials are also provided. Data and modelling software required are either provided from within the platform, or, in some cases the participants are directed to download the needed software from another web site.

The participants follow the lectures at their own pace. During the course they communicate with the lecturers and fellow participants using the functionalities of the platform. In the online environment the exchange between the instructors and the students is made via three ways of communication: a discussion forum where students pose questions, which are answered either by the lecturer, or by their fellow colleagues; a special forum called "Ask the teacher", where the lecturer answers the questions and the third option is an e-mail sent directly to the instructor.

All major units of the online courses include an assignment. Both FMM and DSS course have a set of 5–7

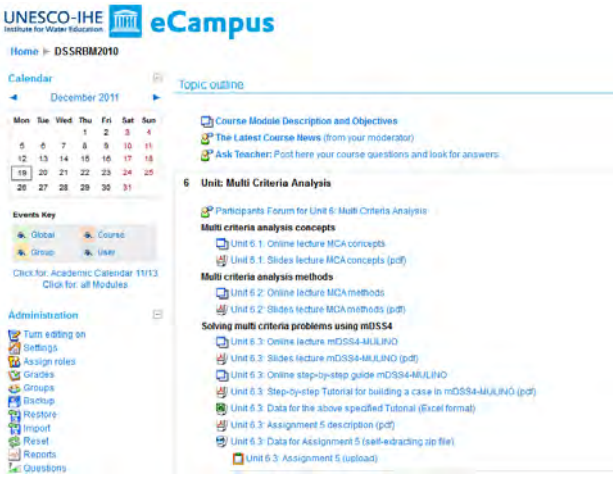

Fig. 3. Moodle implementation of the DSS course – Unit 6 (Multi Criteria Analyses).

assignments that participants have to carry out, and for which they have to submit (upload) reports according to a schedule presented in the course calendar available on the platform. After checking the assignment reports, the instructor provides grades to each participant individually, using the facilities of the platform. Passing grades for the assignment reports lead to successful completion of the course. At the end of an online course, every successful participant receives a certificate, which is a proof that they had followed the course actively and have been assessed through assignments. It needs to be mentioned that till now, the successful participants in FMM and DSS online courses do not automatically receive the ECTS credits when receiving a certificate. This is related to the issue of not having a formal examination in these online courses, as discussed earlier.

5 Discussion on experiences of face to face versus online education

The main educational challenge regarding the two courses described herein was how to structure the content of the course in such a way that the learning objectives are achieved, in both face to face and online mode of delivery. Main topics related to hydrological and hydraulic processes need to be presented to the students in a clear sequence. This is normal in face to face delivery, however, difficult to ensure in an online setting. In this case, once the learning material is deployed, students tend to jump in their learning activities from one topic to another, led by curiosity or confidence that some topics are already familiar to them. This can be controlled by introducing lectures and exercises gradually (hiding some topics for later introduction), but this is normally not appreciated by the students. In both modes of delivery exercises were introduced after each set of lectures

Table 4. Appreciation of collaboration in online courses

What is your opinion on collaborative aspects during the course?	FMM (Completely) agree	DSS (Completely) agree
I had lively and stimulating discussions with other participants in the module.	33.4 %	45.2 %
I learned a lot from other participants in the module.	43.2 %	52.4 %
Other participants in the module were able to answer my questions.	48.6 %	61.9 %
I provided useful help to other participants in the module.	45.9 %	33.4 %
I had feedback that this help to other participants in the module was useful.	48.6 %	35.7 %

to ensure that a particular topic is understood before a new topic was presented. Working with such exercises, to some extent, leads the students to follow a recommended sequence in mastering the topics.

In both the DSS and the FMM course, the approach to structuring of the content was based on the concept of competence-based learning, which ensures that final learning outcomes are achieved (Cheetam and Chivers, 2005). In this concept the whole learning process should be realised in such a way that the learning outcomes of the students lead to attainment of measurable *competencies*. Competencies are defined as final learning outcomes which demonstrate professional ability to perform given *actions* to a sufficient, recognised standard. This attainment can also be at different levels, but, in general, the attainment of a given competence is associated with a required minimum level of demonstrable evidence. It needs to be realised that even when using competence-based learning, prerequisites are important and certain learning sequence is preferred for the kinds of hydroinformatics topics, such as DSS and FMM. In this setting a recommended approach is to include extensive self-evaluation tests for each competence, which was not done so far for the two courses introduced here and remains to be introduced in the future. The design and development of content of the DSS course into competencies, so that understanding is achieved in a short period of time, is described in detail in Jonoski and Popescu (2012). A similar description for the FMM course is available in Popescu et al. (2009).

Structuring of the course content is especially important when it comes to modelling concepts, which represent major parts of both courses. These are first presented in a generic way so that the students can use them with any available software tool, no matter the graphical user interface implementation. In a face to face environment, because of the in-house licenses for commercial software for hydrological and hydrodynamic modelling (such as Mike SHE of DHI or Sobek of Deltares) students can have hands-on exercises and training on all types of tools. The difficulties with this approach for online courses, as already discussed, led to the usage of only open-source or freely available modelling tools. Experience shows that in both face to face and online courses the generic introduction of modelling concepts was adequate for their understanding and students can subsequently easily apply them when requested to solve problems in an exercise.

The key element that made a difference between the classroom education and the online education, in terms of guiding the students learning, is the online discussion forum, which due to its nature of exchanging information becomes, in a certain way, a study guide. Lecturers need to be very active and to follow closely what is posted on the forum because a non-reaction can be understood by the students as an approval. While other fellow students may provide the right guidance, for some questions the interventions from the lecturers are necessary. In a face to face environment this is not the case because students are following the key points emphasised by the lecturer and can ask immediate questions and clarifications.

Due to these differences, the time involvement of the academic staff in the face to face and online courses is also different. While in face to face settings most of the time is spent in a concentrated period of lecturing and exercise sessions, in online courses a prolonged commitment is required because they need to verify on a regular basis if there are questions and if they need to write elaborate answers. From a lecturer's point of view, it is not always clear to which extent certain notions are understood by the students, especially related to modelling concepts and their application.

The time investment of the student in an online course is designed to be the same as for a face to face course. The two courses, in both modes of delivery, presented herein, were designed in a classical teacher-centred approach where most of the learning is reached via information flow between the instructor and the learner. In both cases collaboration between students happens, either directly, or via the forum, and it is well appreciated by the students. The advantage of the online course is that the participants have a higher control on the time they spend for learning, while in a face to face course, due to the fact that the course is carried out for three consecutive weeks, the time for learning is clearly defined and stricter. A thorough evaluation of students' experiences with the two types of courses may provide further insights about the actual time that students spend on learning. Because the same students commonly would not take the same courses both online and face to face, a longer record with many more

online participants is required for such analysis, which remains a task for the future.

As mentioned earlier, topics such as hydraulics, hydrology and decision support systems are better assimilated by the students if analysed and discussed in a group setting, through problem solving. This is done naturally in a face to face environment due to the fact that students share the same classroom, whereas the vehicle for such activities in an online environment is the earlier mentioned discussion forum, or similar tools. One of the major issues for an online course is therefore the level of attainment of the desired collaboration among participants.

In order to assess the collaboration level in the online courses, in the year 2009, participants to the online courses were asked to score six statements regarding collaboration on a five-point scale (1 – Do not agree, 2 – Partially agree, 3 – Neutral, 4 – Agree, 5 – Completely agree). The statements and the result of the scoring are presented in Table 4. When counting specifically the (completely) agreeing percentages, it can be seen that almost half of the participants tend to agree on having had good collaboration. For the DSS course we had 66 participants, of which 50 responded the questionnaire about collaboration, and in the case of the FMM course we had 94 participants and 80 responded the questionnaires. Looking at the overall numbers presented in Table 4 for the FMM course the statement of "I had lively and stimulating discussions with other participants in the pilot" has relatively the lowest score, but still one third agree (completely). In the DSS course 45.2 % participants tend to agree (completely) with the same statement. The overall numbers for this course also give an indication of high collaboration, except that in this case smaller portion of participants agree that they have themselves provided help to others and received feedback (last two rows in Table 4).

In case of the face to face courses the evaluation of collaboration is done at the end of the MSc course, by direct discussion with students and by regular evaluations on the learning load and learning objectives.

6 Conclusions

When comparing face to face and online versions of the same courses the main challenge seems to be maintaining adequate structure of presentation of the course materials. While this is straight forward during face to face course delivery, for online courses this can be achieved by gradual introduction of the course material supported with adequate exercise assignments. This is especially important for modelling courses in which generic concepts need to be introduced before actual hands on exercises with modelling tools. In online courses the discussion forum plays a critical role for enhanced learning through peer collaboration as well as active participation of the involved lecturers. When adequately designed and implemented this approach can act as a substitute to the group problem solving sessions normally introduced in face to face courses. The evaluation of the courses presented in this article indicates a high level of students' appreciation of the collaboration realised in this manner.

There is still much to learn from the experiences with online education in hydroinformatics, such as how to produce interactive materials much more cheaply and effectively. UNESCO-IHE's experience is proving valuable in helping it to fulfil its international remit in education in hydroinformatics. The main challenging questions for the academic staff involved in conducting the online courses are how to measure students' learning and how to set up the online course in such a way that students' learning is facilitated.

The further development of the hydroinformatician depends upon adequate preparation, education and training. The latest technological developments that will determine the success or failure of major water-related projects are explicitly taken into consideration during the development of the hydroinformatics programme, and their achievements are directly used in the educational process and in implementation of numerous research projects. Dozens of the trained specialists have experienced that this programme is a very challenging, while at the same time rewarding undertaking that opens new horizons in their professional career.

Edited by: J. Seibert

References

Abbott, M. B.: Hydroinformatics: Information Technology and the Aquatic Environment, Aldershot, UK/Brookfield, USA: Ashgate, 1991.

Abbott, M. B.: Introducing hydroinformatics, J. Hydroinform., 1, 3–19, 2001.

Abbott, M. B. and Jonoski, A.: Promoting collaborative decision-making through electronic networking, Proc. 3rd Int. Conference on Hydroinformatics, Copenhagen, Denmark, 1998.

Calizaya, A., Meixner, O., Bengtsson, L., and Berndtsson, R.: Multi-criteria Decision Analysis (MCDA) for Integrated Water Resources Management (IWRM) in the Lake Poopo Basin, Bolivia, Water Resour. Manag., 24, 2267–2289, 2010.

Cheetam, G. and Cheevers, G.: Competence and informal learning, Edwar Elgar Press, 2005.

Gonzalez, J. and Wagenaar, R. (Eds.): Tuning Educational Structures in Europe, Final Report Phase One, Universidad de Deusto, Bilbao, 2003.

Jonoski, A.: Hydroinformatics as Sociotechnology: Promoting Individual Stakeholder participation by Using Network Distributed Decision Support Systems, Sweets & Zeitlinger B.V., Lisse, The Netherlands, 2002.

Jonoski, A. and Popescu, I: Distance Learning in Support of Water Resources Management: an Online Course on Decision Support Systems in River Basin Management, Water Resour. Manag., 26, 1287–1305, 2012.

Kaspersma, J. M., Alaerts, G. J., and Slinger, J. H.: Competence formation and post-graduate education in the public water sector in Indonesia, Hydrol. Earth Syst. Sci., 16, 2379–2392, doi:10.5194/hess-16-2379-2012, 2012.

Molkenthin, F., Holz, P. K., Belleudy, P., Jozsa, J., Price, R. K., and van der Veer, P.: HydroWeb: "WWW based collaborative engineering in hydroscience" – a European education experiment in the Internet, J. Hydroinform., 3 239–243, IWA Publishing, 2001.

Odgaard, A. J.: Trends and Current Developments in Hydraulic Engineering, Keynote lecture, National Chiao Tung University, Taiwan of China, 15–16 October 2001.

Popescu, I., Jonoski, A., Bhattacharya, B., and Keuls, C.: On-Line Competence Based Learning in Hydroinformatics at UNESCO-IHE, Proceedings of the 8th International Conference on Hydroinformatics, Concepción, Chile, January 2009.

Price, R. K., Popescu, I., Jonoski, A., and Solomatine, D. P.: Fifteen years of experience in hydroinformatics at UNESCO-IHE Institute for Water Education, Proceedings of the 7th International Conference in Hydroinformatics, Nice, France, Research Publishing, 3101–3108, 2006.

Price, R. K., Bhattacharya, B., Popescu, I., and Jonoski, A.: Flood modelling for management: UNESCO-IHE's online course in hydrology, World Meteorological Organisation Bulletin WMO, 56, 102–106, 2007.

UNESCO-IHP: Water-Education-Training (W-E-T), A Sector Vision of Educators and Those to be Educated, Working Paper, UNESCO, Paris, 1999.

Wagener, T., Kelleher, C., Weiler, M., McGlynn, B., Gooseff, M., Marshall, L., Meixner, T., McGuire, K., Gregg, S., Sharma, P., and Zappe, S.: It takes a community to raise a hydrologist: the Modular Curriculum for Hydrologic Advancement (MOCHA), Hydrol. Earth Syst. Sci., 16, 3405–3418, doi:10.5194/hess-16-3405-2012, 2012.

Environmental Modelling & Software 31 (2012) 37–46

Contents lists available at SciVerse ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

River cross-section extraction from the ASTER global DEM for flood modeling

T.Z. Gichamo [a,1], I. Popescu [a,*], A. Jonoski [a,1], D. Solomatine [a,b,1]

[a] UNESCO-IHE Institute for Water Education, P.O. Box 3015, 2601 DA, Delft, The Netherlands
[b] Water Resources Section, Delft University of Technology, Delft, The Netherlands

ARTICLE INFO

Article history:
Received 18 January 2011
Received in revised form
5 December 2011
Accepted 9 December 2011
Available online 24 December 2011

Keywords:
ASTER GDEM
River cross-sections
HEC-RAS/GeoRAS
Optimization
Tisza River

ABSTRACT

An approach to generate river cross-sections from the Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model (ASTER GDEM) is discussed. The low resolution and the inadequate vertical accuracy of such global data present difficulties in differentiating features of hydraulic importance, which necessitate pre-processing of the DEMs before they are used. A vertical bias correction carried out by comparison of elevation points with a high accuracy terrain model produces a considerable improvement to the cross-sections obtained. In a situation where there are some flow/stage measurements at either end of the river reach, an optimization routine combined with a conceptual flow routing method can provide an additional tool to identify the parameters of an equivalent river section. The extracted cross-sections were used in a 1D river modeling tool HEC-RAS/GeoRAS to simulate flooding on a part of the Tisza River, Hungary. Model results are encouraging and show good potential for using the suggested method in the areas of topographic data scarcity.

1. Introduction

A successful river flood model requires a sufficient representation of the river channel and floodplain geometries, with an accurate description of the model parameters, to make it possible to predict the flow magnitude and water levels along the reach accurately. Software tools have been (and are being) developed and updated to extract spatial features that are useful for hydraulic models, from topographical data sources, both in GIS (Merwade et al., 2008; Tesfa et al., 2011) and non-GIS environments (Schwanghart and Kuhn, 2010). However, obtaining detailed topographical data for every river basin under study is a difficult task as the process involves an expensive and time consuming survey campaign (ground or airborne) and a painstaking post-processing of the survey data (Liu, 2008; Mandlburger et al., 2009; Merwade et al., 2008). A number of studies have been carried out that try to deal with topographical data scarcity in river flood modeling. Most of them rely on the integration of GIS with digital elevation models (DEM) obtained from remote sensing satellites or other globally available data sets (Asante et al., 2008; Herath et al., 2003; Merwade et al., 2005; Sanders, 2007); while others try to use data assimilation techniques to identify a (synthetic) cross-section that is hydraulically equivalent to the real river geometry (Honnorat et al., 2006, 2009; Roux and Dartus, 2008). In addition to topographical data for flood propagation modeling, satellite images provide the locations and types of structures in the area of interest. This is useful for spatially distributed flood damage quantification, particularly in urban areas (Qi and Altinakar, 2011).

This study presents two approaches for the extraction of river cross-sections from a freely available, satellite-based DEM. The first method involves reading the dimensions of triangular cross-sections from the DEM and applying vertical bias correction to improve the constructed cross-sections. In the second method, an optimization routine applied to conceptual flow routing equations is used to identify the equivalent channel geometry parameters from the observed flow and water levels for a given flood event. With the obtained river cross-section data, flood simulation and analysis was carried out on a part of the Tisza River Basin, Hungary, using the HEC-RAS/GeoRAS modeling software package. The following main tasks were carried out:

- Extracting river cross-sections from a satellite DEM, and comparing the outputs from two hydraulic models, one using surveyed data and the other with the satellite DEM;
- Correcting the vertical bias of these cross-sections;
- Identifying the equivalent (synthetic) cross-section using parameters determined from the DEM and the optimization routine;

* Corresponding author. Tel.: +31 15 2151895; fax: +31 15 3122921.
E-mail addresses: tseganeh@gmail.com (T.Z. Gichamo), i.popescu@unesco-ihe.org (I. Popescu), a.jonoski@unesco-ihe.org (A. Jonoski), d.solomatine@unesco-ihe.org (D. Solomatine).
[1] Fax: +31 15 3122921.

1364-8152/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2011.12.003

- Constructing the rating curve at the downstream end of the river;
- Simulating the April 2006 flood and assessing the performance of the methods.

The Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model (ASTER GDEM) was downloaded from the website of NASA Land Processes Distributed Archive Center (LP DAAC) as tiles of 1°-by-1°. The tiles were merged together (mosaic) for the whole river stretch and processed using GIS tools to convert them into Raster and Grid formats. The cross-sections extraction was attempted both automatically using the HEC–GeoRAS tools and manually. The automatic extraction brought problems as it often resulted in unrealistic cross-sections. Due to their lack of adequate horizontal resolution and poor vertical accuracy with respect to the accuracy required by river flood modeling (Vaze et al., 2010), the topographic data obtained from these global sources need correcting for vertical and horizontal biases (Hengl et al., 2010; Pramanik et al., 2010). Elevation value differences were calculated at a number of points along the river channel and floodplain between the ASTER GDEM and a high resolution, high vertical accuracy Digital Terrain Model (DTM) available for a limited portion of the river reach. A simple vertical bias correction by adjusting the cross-sections for the average error of these points produced a considerable improvement in the model outputs.

Before trying to construct a synthetic cross-section, an appropriate hypothesis should be formulated about the relationship between the hydraulic parameters and the geometric variables (Roux and Dartus, 2008). The approach adopted here follows (Price, 2009a) who used the cross-section formulation given in (Knight, 2006) and combined it with a conceptual model for one dimensional flow routing (Price, 2009b). This formulation defines the channel bed width and depth, channel side slope, flow and storage width for in-bank and out of bank areas which vary with the depth of flow. The values for these parameters are to be obtained through an optimization routine that minimizes the discrepancy between simulated and reference hydrographs (stage or discharge) at the downstream boundary.

Moreover, the methodology (for the construction of synthetic cross-sections) requires a prescribed (at least, in-bank) rating curve either at the upstream or downstream end of the river reach. When there is an accurate in-bank rating curve, parameters to extend it for flows exceeding bank-full discharge can be identified through the optimization process. The work reported here involves a situation when it was assumed that there is no rating curve at either end of the river. An approximate rating curve was constructed for the downstream boundary using an approach outlined in (Fenton and Keller, 2001). The work demonstrates the value of following sound hydraulic principles to generate a rating curve in the presence of only limited gauged data.

2. Methodology

The flow chart in Fig. 1 outlines the steps followed to construct the river cross-sections based on ASTER GDEM data and synthetic cross-sections by utilizing optimization. Each block points to the relevant section in the paper where detailed description is provided. ASTER GDEM products can be downloaded from the websites supported by the Ministry of Economy, Trade and Industry (METI) of Japan (http://www.gdem.aster.ersdac.or.jp/index.jsp) or the National Aeronautics and Space Administration (NASA) (http://asterweb.jpl.nasa.gov/gdem-wist.asp). The DEM is available in $1^\circ \times 1^\circ$ tiles, which can be combined in Arc-GIS environment. To ensure obtaining a smooth terrain, the DEM can be processed in Arc-GIS to fill or trim unrealistically too low or too high spot elevations.

The river scheme is formed in HEC–GeoRAS by manually 'digitizing' the river centerline, bank locations, flow directions, river cross-section locations, and other features. HEC–GeoRAS is a freely available add-on to ARC-GIS, which can be obtained from the US Army Corps of Engineers- Hydrologic Engineering Center (http://www.hec.usace.army.mil/software/hec-ras/hec-georas.html). The triangular cross-section is formed by reading the elevations of five points along a river cross-section, while for the synthetic cross-section the bottom elevations at the upstream and downstream ends of the reach are required, from which the slope of the channel is calculated.

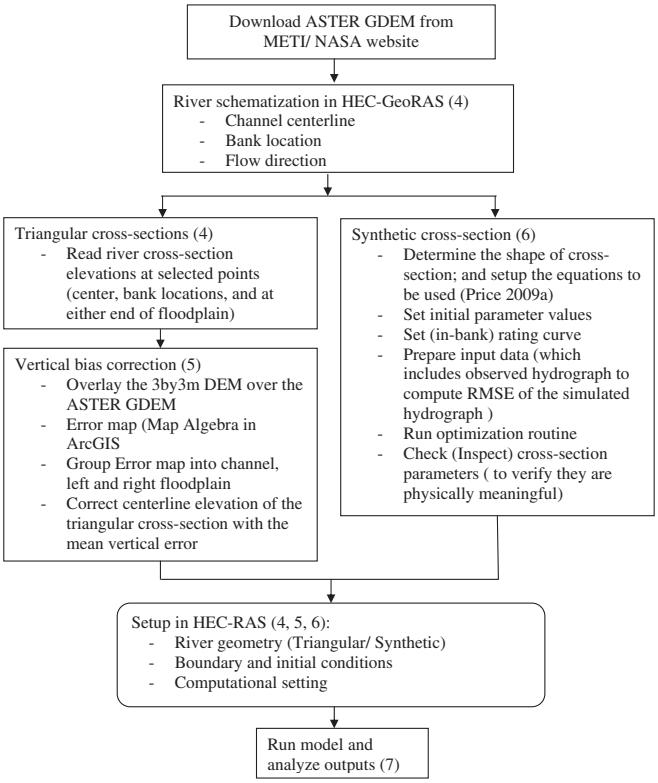

Fig. 1. Methodology (numbers in brackets indicate the section of this paper where the specific task is explained).

To undertake vertical bias correction for the triangular cross-section, a raster map is produced with values of elevation difference between a 3 m-by-3 m DEM and ASTER GDEM. This map is then divided into three raster maps, one for the channel area and one for each of the floodplains (left and right). The mean elevation difference for each area is computed and used to adjust the elevations of the triangular cross-sections. The configuration of the synthetic cross-section is determined by equations given in (Price, 2009a). The equations relate a number of channel parameters with the geometry of the channel as a function of the flow depth. The parameters are identified through optimization, which tries to minimize the root mean square error between observed and simulated discharge or water level at the downstream end of the reach. The parameter ranges must be constrained and the final values should be checked to verify that physically acceptable values of the parameters are obtained from the optimization.

3. Model setup and data used

A 1D hydrodynamic model of the middle Tisza River system consisting of the River Tisza spanning from Kisköre to Szeged, and the rivers Zagyva and Hármas- Körös joining it near the cities Szolnok and Csongrád respectively, was used as a starting model. This river model (hereafter referred to as the 'Bigger model') was obtained from theMiddle-Tisza District Environment and Water Management (KÖTIKÖVIZIG), Hungary, through the FLOODsite project (Bakonyi, 2009). The model was built using the HEC-RAS software, and it consisted of 8 rivers with stretches of flood defense embankments, flow obstruction structures, bridges and culverts, and proposed flood detention areas. Detailed river cross-sections at reasonable spacing are available in this model, together with spatially distributed Manning roughness coefficients.

The values of the Manning roughness coefficients for the bigger model were calibrated using the flood of spring 2000. The calibration procedure, which is described in detail in (Kovács et al., 2006), starts

by dividing the river cross-sections into strips of uniform roughness coefficients based on land cover classification obtained from aerial photos and field surveys. Each land cover class was assigned a range of roughness values which correspond to values suggested in literature, the Hungarian standard, and refined by the experience of the modellers in applying the HEC-RAS model (Tóth and Kovács, 2007). During the calibration runs, the roughness values for each strip were allowed to vary within the prescribed range, and comparisons were made between model calculated water levels and time series of hourly measured water levels at 50 gauge locations along the river. The calculated maximum water levels were within 10 cm of the observed ones (Kovács et al., 2006).

Model verification was carried out for the flood of spring 1999 (the same season as for the calibration, hence similar vegetation condition expected). The maximum discrepancy between the observed and simulated water levels during the verification was 24 cm which the authors reported to be within the expected error range due to measurements. The model was then used to simulate the flood of April 2006 to test the effectiveness of proposed measures aimed at cutting the flood peak, such as relocation of dykes and provision of flood detention basins. The full account of the bigger model setup, calibration, verification, and its performance in predicting the effect of the flood peak reduction measures is reported in (Tóth and Kovács, 2007; and Kovács et al., 2006). Some background information about flooding in Tisza river basin can also be found in (Bálint and Tóth, 2006). Fig. 2 shows the map of the study area with Szolnok – Csongrád reach of the Tisza River overlaid on it, together with a 3 m-by-3 m DEM covering a shorter span of the reach which was used for vertical bias correction (discussed in Section 5).

For this particular study we focus only on the part of the Tisza River that runs between the cities of Szolnok and Csongrád (stations 335.600 and 244.000 in the bigger model). First a model with all the information as in the bigger model was built (named here as the 'Reference model'). This is a single river – single reach system with one planned storage area towards the downstream end of the river stretch (Köröszug), two bridges at the upstream section near the city of Szolnok (stations 335.564 and 331.460) and one bridge downstream at station 267.624 Its length is 88.4 km, and it has all the characteristics of the bigger model – the same (spatially distributed) values of Manning coefficient, cross-sections from surveyed data,… etc., except it only spans between the cities of Szolnok and Csongrád.

Next, the models with river cross-sections extracted from limited data are built for the Szolnok – Csongrád reach (we simply call them 'the Szolnok – Csongrád test models' as they are the focus of this study). The purpose of the study is to combine techniques that can help identify the river cross-sections from freely available satellite maps and records of discharge or water level upstream and downstream of the river reach. Hence, apart from the locations and dimensions of the bridges and the storage areas all other information is assumed to be non-existent, and will be determined through the use of the freely available data sources. The outputs from the reference model are used as a benchmark to measure the performance of the test models. To define the river geometry (alignment and cross-sections) a satellite-based DEM (ASTER GDEM) is used with the HEC–GeoRAS tool in the ArcGIS 9.2 environment. In addition a 3 m-by-3 m resolution DTM for a shorter (30 km) span of the river reach was obtained from KÖTIKÖVIZIG. The latter was composed from orthophotos and low water cross-section surveys by ultrasonic depth measurements complemented by geodetic and GPS surveys. This will be used for bias correction of the satellite DEM. Two sets of tests were carried out. In the first test (Sections 4 and 5), a triangular cross-section is constructed from elevation readings from the ASTER GDEM. The second test (Section 6) involves identifying parameters of a conceptual (synthetic) cross-section through optimization.

As mentioned earlier, the reference model has distributed roughness coefficients, which vary both longitudinally (from one river section to the other) as well as across a given cross-section. However, owing to the assumption of limited data, single values of Manning coefficient for the channel and floodplain are used in the test models. Tóth and Kovács (2007; see in the cited document Table 4, p. 31) show that the range of the main channel roughness coefficient varies between 0.017 and 0.06. For the floodplain, seven land use/land cover types are given, with the friction coefficient values ranging from 0.02 for plough land to 0.2 for forest with undergrowth. Inspection of the calibrated reference model shows that the channel roughness coefficient predominantly lies in the range 0.025–0.036. For the overbank areas it varies between 0.027 and 0.19, most of the values being close to 0.1. For the tests reported in this paper Manning roughness coefficients of 0.034 and 0.097 are used for the main channel and the overbank areas respectively. These values were approximated as weighted averages of the spatially distributed values. To test the degree of sensitivity of the model outputs to the use of lumped roughness coefficients, the reference model was run with the spatially averaged Manning coefficients instead of the distributed values. The model outputs reported in section 7 (Results and Discussions) show that the errors due to the use of 'lumped' roughness coefficient are not significant compared to the overall model output errors.

The upstream boundary condition of the reference model was set with flow hydrograph computed by the bigger model at station 335.600. For the downstream boundary condition, time series of measured water levels available near Csongrád was used. The initial outputs from this (reference) model run give stable flow and stage hydrographs for most part of the reach except for stations at and near the downstream boundary where the flow hydrographs show oscillations about the line of the moving average (Fig. 3 shows one such hydrograph). An inaccurate choice of time/space steps and/or inadequate river cross-section spacing might be the possible reasons for the inaccuracies; therefore these two factors were checked first. Ideally the time step would be set using the Courant number criteria for stability with respect to the time/space steps. However, because this might be too demanding, the HEC-RAS users' manual (Brunner, 2008 pp 8–63) suggests choosing time steps that fulfill the following condition:

$$\Delta t \leq \text{Time for the rise of hydrograph}/20$$

This condition was satisfied by the time step adopted in the model. The river cross-section spacing was also within the recommended ranges (Brunner, 2008 pp. 8–60; Castellarin et al., 2009; Samuels, 1990). Where the spacing was too wide, interpolated cross-sections were used.

The use of flow hydrograph at station 335.600 and a rating curve at station 244.000, both computed by the bigger HEC-RAS model, as the upstream and downstream boundary conditions respectively give smooth results. The same rating curve was used as downstream boundary condition for the first test, i.e., the Szolnok – Csongrád model with triangular cross-sections from the ASTER GDEM. For the second test (involving synthetic cross-sections), an approximate rating curve was constructed for the downstream boundary as explained below.

By considering the flow theory for the various cross-section geometries of a river and based on observations, Fenton and Keller (2001) showed that in many cases the stage varies as a function of the square root of discharge, and plot of the square root of discharge versus stage would fall into an almost straight

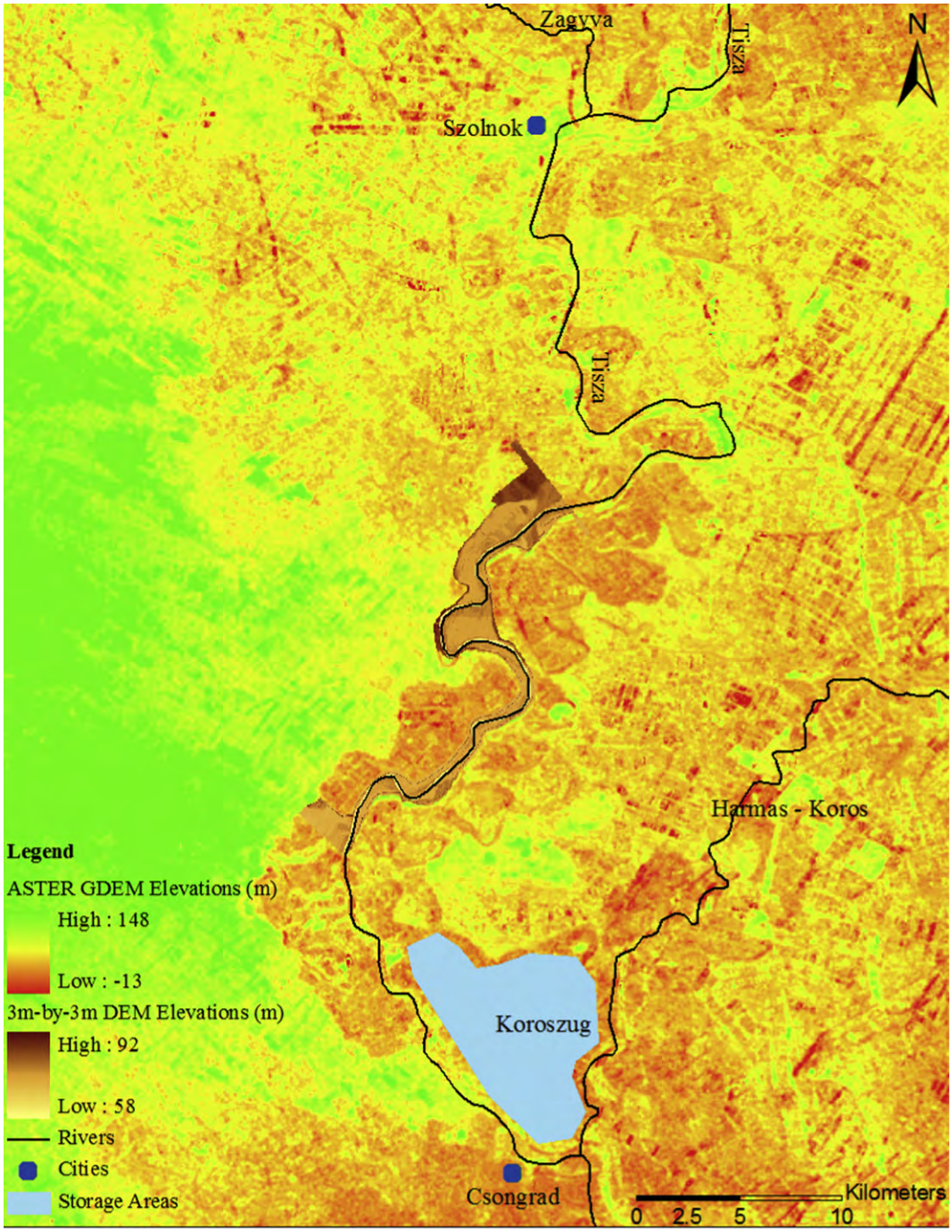

Fig. 2. ASTER GDEM of the study area (a 3 m-by-3 m DEM is overlaid at the center).

line. This can be used to generate the rating curve in the case of limited data. Thus, to construct the rating curve at the downstream boundary we start with the equation of the form:

$$Q = \alpha(H - \mathrm{Ho})^e$$

Where Q and H are discharge and water level (stage) respectively, α is constant and Ho is the zero-flow water level, i.e., the y-intercept of the stage-discharge rating curve. The coefficient e is assumed to be equal to 2 or values close to 2; therefore developing the rating curve primarily involves getting the values of α and Ho. The values $\alpha = 23.2$ and $\mathrm{Ho} = 76.1$ m were obtained from the best fit into the limited stage-discharge measurements available at the downstream section.

4. River geometrical data with HEC–GeoRAS

HEC–GeoRAS is a set of GIS tools used to prepare HEC-RAS geometric input data and to process its outputs, such as preparing flood maps based on the water surface profile calculations (Ackerman, 2009). With the preprocessor tool of HEC–GeoRAS, terrain data (TIN or Grid) are used to digitize the river centerline, banks, flow direction, and the start and end stations of a river reach.

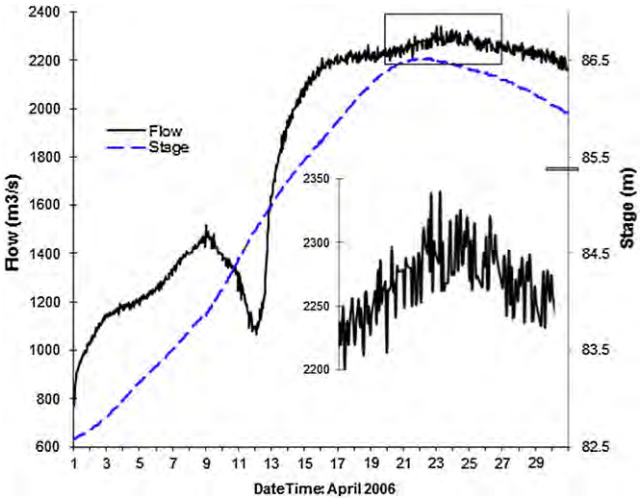

Fig. 3. Flow and stage hydrographs at station 244 for the reference model with observed water level as downstream boundary condition (Flow hydrograph has oscillations; part of it is magnified in the smaller graph).

The positions of flow structures like bridges, culverts and embankments, as well as other flow obstruction structures are also specified and, if a land use map is available, it can be processed to obtain the friction coefficient distribution across a section. The cross-section is digitized by using the stream lines and bank locations. These GIS layers are then used to produce an import file for HEC-RAS (a geometry file).

This is a quick and an easy approach for situations where the flow is predominantly in a single direction and the flood depth is the primary sought-after variable. The results however are likely to depend on the type, quality and resolution of the terrain data. Tiles from the ASTER GDEM are used here to prepare the input files for the HEC-RAS model. During the process of preparing geometric input files for HEC-RAS, river cross-sections are digitized by using the stream lines and bank locations. For user-selected locations along the stream, HEC–GeoRAS will read the elevation values at stations along a line defined as cross-section cut line. For the considered case study, such an automated extraction of river cross-sections proved to be not satisfactory (so of no practical use) as unrealistic elevation values were obtained in many locations. This might be because the satellite images read the water surface elevations while elevations of important structures like embankments are missed. Sometimes a cross-section where the centerline elevation is higher than the river banks was obtained.

To get around this problem a manual extraction of the elevations at each cross-section cut line was carried out after the DEM was first processed with GIS to fill or trim unrealistically too low or too high spot elevations. Elevation values at pixels were read, and the number of cells between two points of elevation reading was used to approximate the distance between them. To begin with a simple cross-section shape, a triangular cross-section was first assumed. This requires finding elevations at a minimum of 5 points: at the stream centerline, at the bank stations, and on the boundary of the floodplain either side of the river. When the pixel value appears to be practically invalid, the nearest cell value is read and used. When the whole cross-section that was obtained in this way became unrealistic another cross-section was taken nearby.

A total of 38 cross-sections were defined this way (in the reference model there are 84 sections) and the rest were interpolated to satisfy the requirements of cross-section spacing. The model simulation was then carried out for the April 2006 flood, and the downstream hydrographs and the water surface profiles inside the river cross-sections were compared with those of the reference model outputs. Generally higher water levels were obtained compared to the reference case, and the flow hydrograph exhibits some instabilities which are likely caused by the inaccurate elevation data.

5. Vertical bias correction

Generally the cross-sections obtained from the ASTER GDEM are shallower (in depth) and the resulting water surface elevations are higher than the one in the reference model. To characterize the possible vertical bias of the DEM, elevation value differences between the 3 m-by-3 m resolution DEM which is available for the middle portion of the river reach and the ASTER GDEM were computed. Then the mean errors were determined for the channel and the left and right overbank areas separately. The assumption here is that the average computed topographical error for this 30 km stretch can be representative for the whole reach. This is not an unrealistic assumption considering the size of the river stretch compared to the satellite image swath covered in a single flight.

Fig. 4 shows the raster map of the computed elevation differences. Some of the statistical descriptors of the vertical errors are summarized in Table 1, and the vertical error distribution for the channel is plotted in Fig. 5. The ASTER GDEM elevations at the channel are, on average, 8.2 m higher than the reference elevations. ASTER GDEM Validation Team (2009; see in the cited document Table 9, pp 13 – 15) provides vertical errors of the ASTER GDEM analyzed for selected areas around the globe. The study shows that for the geographic location 48°N 13°E the ASTER GDEM is 7.48 m lower than a reference DTM (DTM Germany) and for the geographic location 40°N 23°E the ASTER GDEM elevations are 8.51 lower than SRTM elevations used as reference. These are two of the closest geographical areas to the Tisza River considered in the validation study. One possible explanation for the difference between the vertical bias calculated above and the ones reported in the validation report can be the land cover. The validation report suggests that, while the ASTER GDEM elevations are lower than the true/reference elevations in general, they are likely to be higher in areas covered with forest because the canopy heights rather than bare earth elevations are captured.

The river cross-sections obtained from the GDEM were corrected by accounting for the vertical bias calculated above, and model simulations were carried out using the corrected river cross-sections. An attempt to incorporate a horizontal bias correction by using the GIS stream delineation features was not successful here as the streams obtained from the GDEM were not clear enough to compare with the streams obtained from the higher resolution DEM. However the location of the river alignment might have appreciable uncertainty (Hengl et al., 2010); hence, making horizontal bias correction in the future might help in further refining the model outputs.

6. Identification of synthetic cross-sections through optimization

The cross-sections that were extracted from satellite DEM in the previous section all have triangular shapes. In the second test optimization methods were used to find the river cross-section shapes. Due to the nature of the 1D flood routing models, synthetic/standard cross-sections might replace true/real cross-sections while preserving sufficient description of the channel hydraulic characteristics (Roux and Dartus, 2004). The synthetic cross-sections can be identified by starting with some random shape and progressively refining it, employing certain data assimilation techniques. The success of such an approach may depend on the type of flow routing method used, for example when the

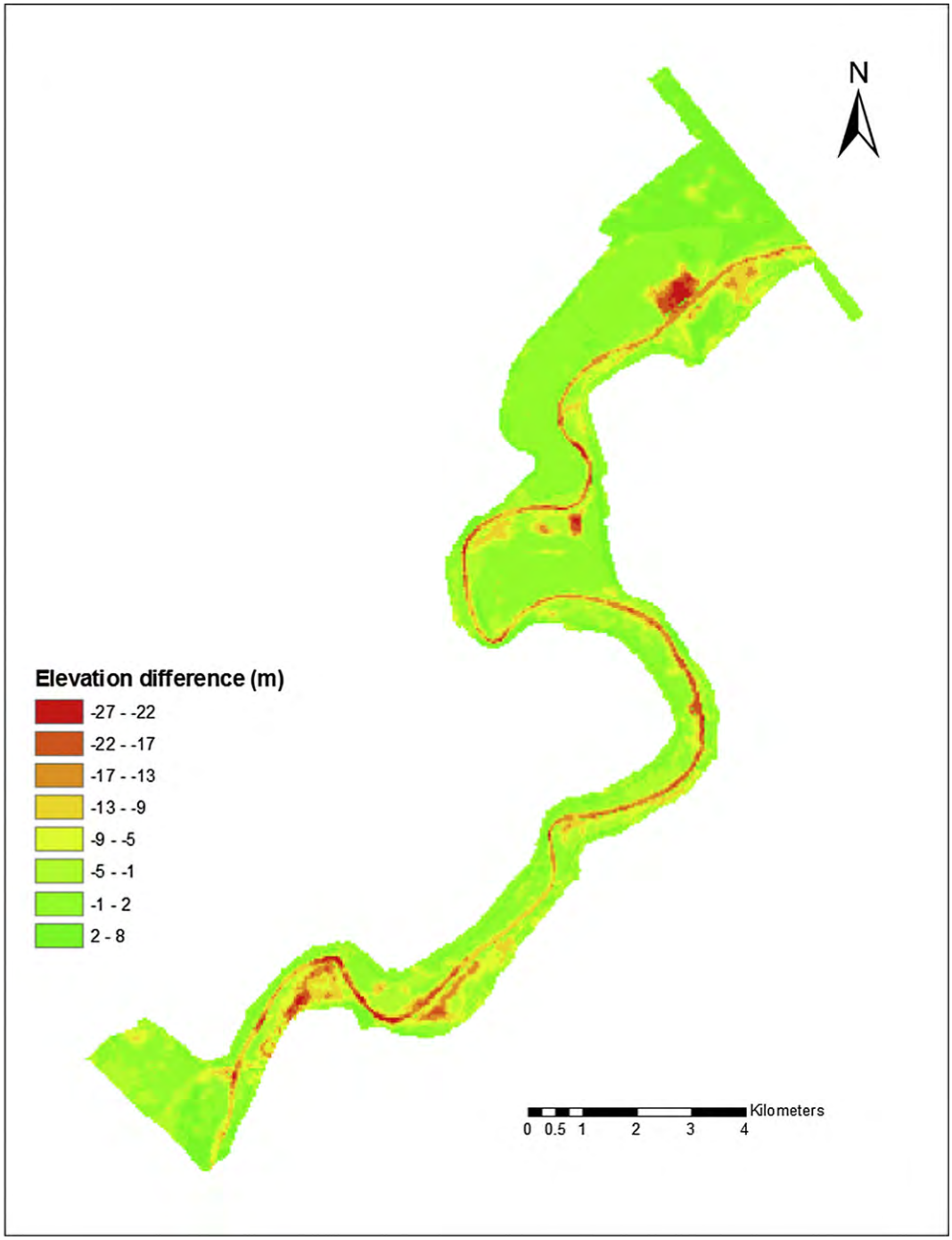

Fig. 4. Raster map showing the elevation difference between 3 m-by-3 m DEM and ASTER GDEM for part of the study area where the 3 m DEM is available.

routing method requires only the gross area of flow rather than the definite shape of the cross-section.

A method of identifying river cross-section for a 1D flood routing making use of recorded flows and/or stages as proposed by (Price, 2009a) was tested here for the model of the Szolnok-Csongrád reach of Tisza River (See also Knight, 2006; Price, 2009b). The method requires determination of the channel depth (*Ych*), channel semi bed width (*Bb*), channel side slope (*Ss*), semi storage width for in-bank flow (*Bch*), semi storage width for out of bank flow (*Bfl*), and dimensionless numbers (coefficients) *k1*, *k2*, *k3*, *k4*. Equations are formulated in terms of these geometric variables and coefficients to define curves that determine the active flow (semi flow width *Bf*) and the semi storage width (*Bs*) as a function of the depth of water. The discharge and water levels at the upstream and downstream ends of the reach are related through in-bank rating curves, which are extended by the optimization routine for flows

Table 1
Basic statistical descriptors of vertical error.

	Channel	Left overbank	Right overbank
Mean (m)	−8.2	−2.5	0.7
Minimum (m)	5.0	−24.0	−27.0
Maximum (m)	−27.0	8.0	7.0
Standard deviation (m)	8.0	5.0	3.7

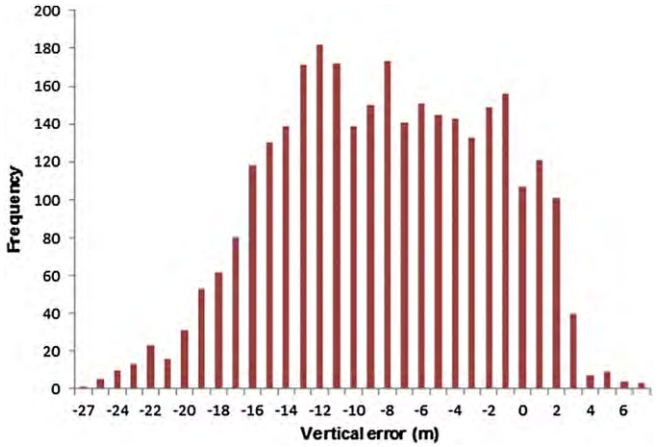

Fig. 5. Vertical error distribution plot for the channel.

beyond bank-full discharge. Fig. 6 shows the schematization of the synthetic cross-section. For detailed description of the cross-section formulation and the equations used see (Price, 2009a).

For a given reach and flood event these parameters can be determined by optimization routine using the available discharge/stage measurements during the event, together with a limited amount of information on the river reach. A total of thirteen parameters of optimization resulted, which include nine parameters for the channel and floodplain dimensions listed above plus four parameters for the extension of the in-bank rating curves for above-bank flows. The root mean square error of the predicted downstream discharge is selected as the objective function for the optimization. Two optimization algorithms, the Controlled Random Search CRS2 (Price, 1983) and the Adaptive Cluster Covering with local search ACCOL (Solomatine, 1999) were used in the Global Optimization Tool GLOBE (Solomatine, 1995; http://www.data-machine.nl/). There was no marked difference between the performances of the two algorithms.

After the cross-sections were plotted for the upstream and downstream ends of the Szolnok-Csongrád reach, the intermediate sections were obtained by interpolation. Two challenges are faced during this process: first, in the parameterization followed here, the storage width (which corresponds to the ineffective flow area width in HEC-RAS) varies with the depth of flow. As there is no way to enter varying storage area width into the HEC-RAS model, the value estimated for the maximum flood level was used. The second challenge is related to the range of values the parameters involved in the optimization can assume; i.e., there is no single/unique value a given parameter should take, and there is also a possibility that the optimization might result in values that do not represent the true cross-section. Therefore adequate physical constraints should be placed on the parameter ranges to avoid too long optimization run time and to prevent reaching at unrealistic values.

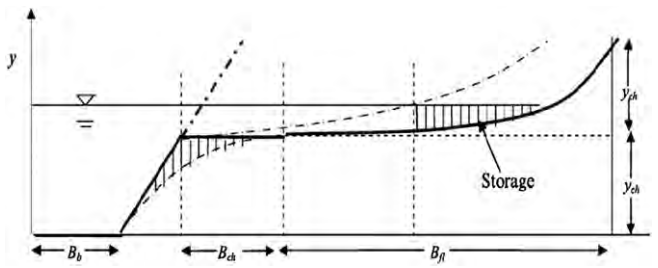

Fig. 6. Half channel for the synthetic cross-section (Adapted from Price, 2009a).

Table 2
Parameter values for synthetic cross-section obtained by optimization.

Optimization parameters	Unit	Value
Channel depth (*Ych*)	m	13.0
Channel semi bed width (*Bb*)	m	10.0
Channel side slope (*Ss*)		2.6
Semi storage width for in-bank flow (*Bch*)	m	203.2
Semi storage width for out of bank flow (*Bfl*)	m	3829.8
In-bank storage width index (*k1*)		0.98
Above-bank storage width index (*k2*)		0.70
Above-bank flow width index (*k3*)		0.52
Above-bank storage width factor (*k4*)		0.50
Above-bank rating curve index upstream ($\beta_{ab,up}$)		7.57
Above-bank rating curve control depth upstream ($y_{ab,up}$)	m	0.38
Above-bank rating curve index downstream ($\beta_{ab,dn}$)		10.0
Above-bank rating curve control depth downstream ($y_{ab,dn}$)	m	0.16
Objective function (RMSE)		0.1735

Since the primary concern here is to determine the channel dimensions that are below the water surface (i.e., the ones that cannot be picked up by the DEM), the channel and floodplain geometries which can be obtained from the freely available DEM with a relatively good accuracy were directly used. These include the total width of the floodplain and the location of channel banks (hence the channel top width). These two variables are not in the optimization parameter list; however they are used to define the limits of semi flow width (*Bf*) and semi storage width (*Bs*), hence constraining the channel extents. Constraining the channel extent this way would speed up optimization and it helps to obtain realistic values. The parameters identified by the optimization and the value of the objective function RMSE (root mean square error between observed and simulated discharge hydrographs at the downstream boundary) are reported in Table 2.

7. Results and discussions

The model outputs are presented in Fig. 7 through 10. Figs. 7 and 8 show that the downstream flow and stage hydrographs are reasonably well predicted by both test models. The simulations initially over-predict the hydrographs (both flow and stage), but they come closer to the reference model outputs when approaching the peak. The simulated peak discharges are very close to the corresponding value for the reference model. The peak water level from the model with synthetic cross-section is 13 cm less than the reference peak water level, while the model with triangular cross-section predicts a value 10 cm higher than the reference. These are

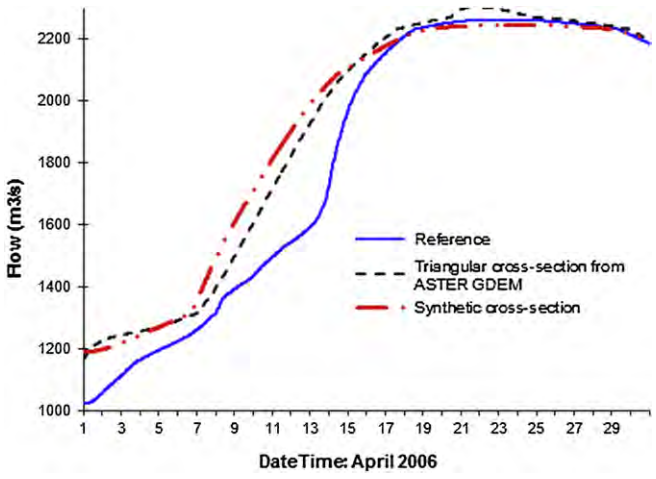

Fig. 7. Flow hydrographs at station 244.000 (downstream end).

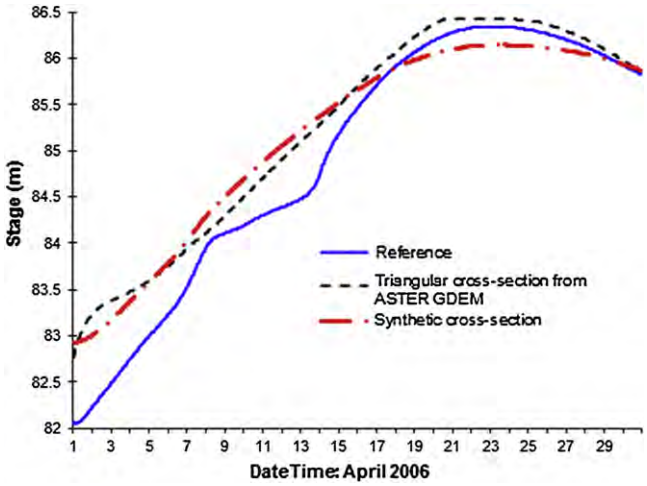

Fig. 8. Stage hydrographs at station 244.000 (downstream end).

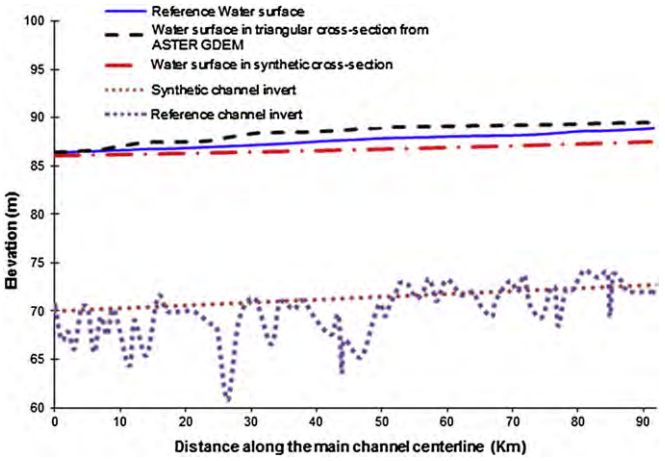

Fig. 10. Profile plot showing the maximum water levels along the river channel (The horizontal axis starts at the downstream end).

acceptable deviations considering the maximum stage errors during the calibration and verification of the bigger model (10 and 24 cm respectively).

Figs. 9 and 10 evaluate model performance at a station away from the downstream boundary. The simulated flow hydrographs at station 292 (midway between the upstream and downstream boundary) show good agreement with the reference model output and they perfectly capture the peak (Fig. 9). The stages, however, do not follow the same trend. As can be seen from the profile plots in Fig. 10, the simulated maximum water levels deviate from the reference maximum water level away from the downstream boundary. At the point of maximum error, the deviations reach about 130 cm, where the model with synthetic cross-section under-predicts the water surface profile while the model with triangular cross-section over-predicts it.

Fig. 10 also shows the bed elevation along the channel centerline for the reference and synthetic cross-sections. By the nature of its definition, a reach with synthetic cross-section has a continuous slope while the real channel bed is irregular with dips and bumps. Another important factor worth considering is the definition of storage (ineffective flow areas) in the synthetic cross-section. The optimization routine (Price, 2009a) defines storage (ineffective flow area) that varies with the depth of flow. However once the optimized cross-section is brought into HEC-RAS, a constant ineffective flow area was used (because it is not possible to define varying ineffective flow areas in HEC-RAS). These factors might contribute to the difference in outputs between the reference model and the one with the synthetic cross-sections. Generally it appears reasonable to advise that when using the synthetic cross-sections one should strictly follow all the assumptions and requirements set out in its formulation, rather than, for instance, trying to use it with a model (HEC-RAS) that was not part of the optimization, as was attempted here.

The over-estimation of water levels by the model with triangular cross-sections is likely caused by elevation errors that are still remaining after vertical bias correction. Pramanik et al. (2010) show that the vertical error of a satellite DEM at a point can be a function of its elevation value. The errors can also be related to the rate of elevation change or relief (ASTER GDEM Validation Team, 2009). Hence undertaking the bias correction by grouping the study area based on elevation ranges, relief, and other factors like land cover types may result in better simulation outputs. In the absence of accurate topographical data (like the 3 m-by-3 m DTM used in this study or other landmark elevation spots), the vertical error estimates provided by the validation study might be used for preliminary vertical bias correction, but care should be exercised to account for the other important factors mentioned above.

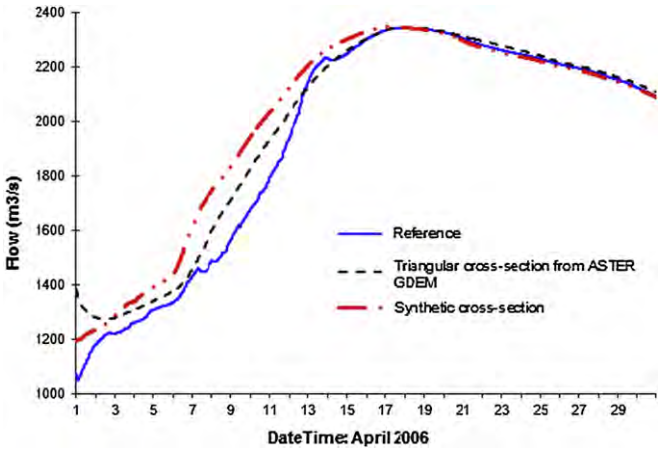

Fig. 9. Flow hydrographs at station 292 (mid way along the reach).

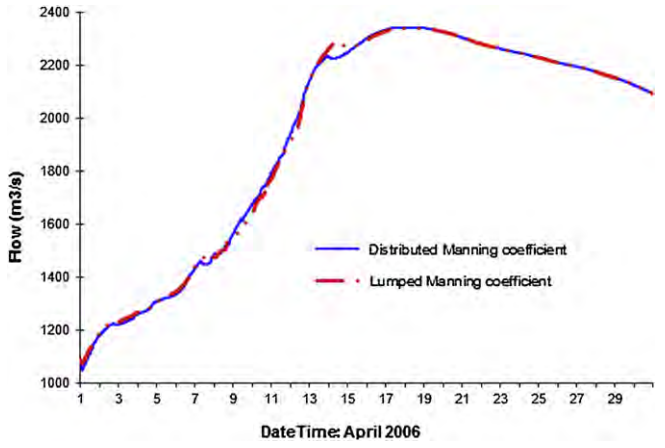

Fig. 11. Flow hydrographs of reference model at station 292 with distributed and lumped Manning coefficients.

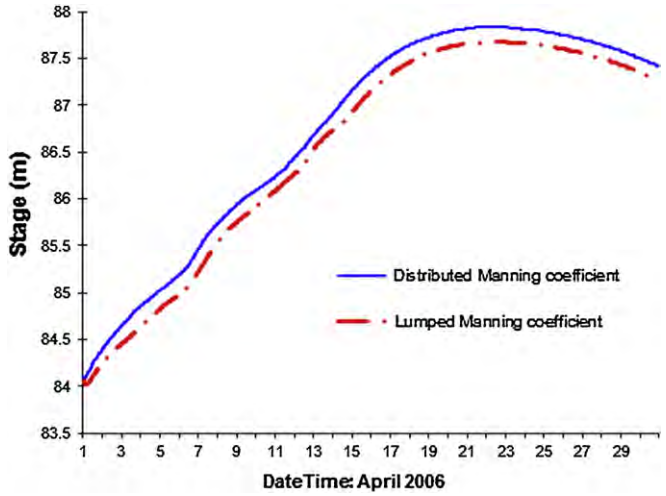

Fig. 12. Stage hydrographs of reference model at station 292 with distributed and lumped Manning coefficients.

Figs. 11 and 12 compare the reference model outputs at station 292 with distributed and lumped Manning coefficients. The flow hydrographs are essentially the same while the stage hydrograph is lower when the averaged Manning coefficients are used. The difference is less than 10 cm at all points on the stage hydrograph, which implies the influence of using lumped friction values in this study is insignificant when compared to the overall model output errors. It is also interesting to note that the water level is more sensitive to the use of spatially averaged Manning coefficient than the discharge, which is consistent with the results in Figs. 8–10 (i.e., stage errors are larger than discharge errors). It is difficult to argue that the results suggest the model outputs are insensitive to the use of average Manning coefficients. Rather they likely indicate that the influence can be masked by other factors, in this case the error in the geometry of the river section.

The methods presented herein have some limitations that need attention. The first is that simplified flow routing equations are used in the optimization routine. Such approximate methods based on the advection-diffusion equations may give incorrect results when the river reach is on a gentle slope (Cunge, 1969; Also see the argument in Fenton, 2010). This problem would be solved when the optimization is coupled with solutions of the full hydrodynamic equations (a possible extension of this work). The other challenge is related to the approximate rating curve estimated using the available records of water level and flow at the downstream end. Ideally, the rating curve should be constructed using long historical records of flows/stages — not just a single flooding event — before it can be used for future flood predictions. And finally, the tests carried out in this study were carried out manually, i.e., taking approximate cross-sections from the satellite DEMs, using visual inspection to drop those unrealistic values, determining the parameters of the river cross-section, and running the conceptual model for the various sets of geometric parameters,…etc. Translating the whole process into a semi-automatic algorithm might make the methodology more effective when used in similar situations.

8. Conclusions

Two approaches for the extraction of river cross-sections from freely available satellite DEM (the ASTER GDEM) for flood routing were discussed. The methods are useful where accurate river cross-sections and detailed floodplain information from surveying or high resolution aerial images are not available. In the first test, the dimensions of triangular cross-sections were determined from the DEM, which were subsequently corrected for vertical bias based on elevation error computation between the ASTER GDEM and a 3 m-by-3 m DEM along the river channel and in the floodplain. The vertical bias correction was a very important component of the methodology followed, even though elevation comparisons were done only for a limited part of the river reach where a high resolution/high accuracy DEM was available.

The second method uses an optimization routine that requires discharge and/or water level measurements at the downstream boundary together with an in-bank rating curve at either upstream or downstream (or both) of the river reach. The rating curve was also assumed to be non-existent here, and it was constructed using an approach based on flow theory together with the limited information at the downstream end. To speed up the identification of optimization parameters and to provide better physical constraints, the optimization is complemented by the direct use of the channel and floodplain dimensions which can be determined from the DEM relatively accurately. The approaches presented have limitations, and they are not a substitute for the use of accurate topographical data when available; however they can provide a valuable tool for river flood routing in areas with lack of sufficient topographical and flow data.

Acknowledgments

We would like to thank KÖTIKÖVIZIG, Dr. P. Bakonyi, and the FLOODsite Project for providing the data and the initial model of the Tisza River used in this paper. And we are very grateful to Prof. R. Price (UNESCO-IHE) for his valuable comments and the discussions we had during the course of this study. The comments and suggestions from the anonymous reviewers have helped improve the paper considerably.

References

ASTER GDEM Validation Team, 2009. ASTER Global DEM Validation Summary Report. METI/ERSDAC, NASA/LPDAAC, USGS/EROS.
Ackerman, C.T., 2009. HEC-GeoRAS GIS Tools for Support of HEC-RAS Using ArcGIS User's Manual Version 4.2. US Army Corps of Engineers Institute for Water Resources Hydrologic Engineering Center (HEC).
Asante, K., Arlan, G., Pervez, S., Rowland, J., 2008. A linear geospatial streamflow modeling system for data sparse environments. International Journal of River Basin Management 6 (3), 233–241.
Bálint, Z., Tóth, S., 2006. Flood protection in the Tisza river basin. In: Schanze, J., Zeman, E., Marsalek, J. (Eds.), Flood Risk Management: Hazards, Vulnerability and Mitigation Measures. Springer, pp. 185–197.
Bakonyi, P., 2009. Pilot study of the "Tisza River Basin" Executive Summary, FLOODsite Report Number T22-08-01. KÖTIKÖVIZIG Kht.
Brunner, G., 2008. HEC-RAS River Analysis System User's Manual Version 4.0, Report CPD-68. US Army Corps of Engineers, Hydrologic Engineering Center, Davis, CA.
Castellarin, A., Di Baldassarre, G., Bates, P., Brath, A., 2009. Optimal cross-sectional spacing in Preissmann scheme 1D hydrodynamic models. Journal of Hydraulic Engineering 135 (2), 96–105.
Cunge, J., 1969. On the subject of a flood propagation computation method (Musklngum method). Journal of Hydraulic Research 7 (2), 205–230.
Fenton, J., 2010. Accuracy of Muskingum-Cunge Flood Routing.
Fenton, J., Keller, R., 2001. The Calculation of Streamflow from Measurements of Stage. Cooperative Research Centre for Catchment Hydrology.
Hengl, T., Heuvelink, G., Van Loon, E., 2010. On the uncertainty of stream networks derived from elevation data: the error propagation approach. Hydrology and Earth System Sciences 14 (7), 1153–1165.
Herath, S., Dutta, D., Wijesekera, S., 2003. A coupled river and inundation modeling scheme for efficient flood forecasting. Geoinformatics 14 (1), 37–41.
Honnorat, M., Lai, X., Monnier, J., Le Dimet, F., 2006. Variational Data Assimilation for 2D Fluvial Hydraulics Simulations. CMWR XVI — Computational Methods in Water Resources XVI International Conference, Copenhagen, Denmark.
Honnorat, M., Monnier, J., Le Dimet, F., 2009. Lagrangian data assimilation for river hydraulics simulations. Computing and Visualization in Science 12 (5), 235–246.
Knight, D., 2006. River flood hydraulics: validation issues in one-dimensional flood routing models. In: Knight, D., Shamseldin, A. (Eds.), River Basin Modelling for Flood Risk Mitigation. Taylor and Francis, Leiden, Netherlands, pp. 335–385.

Kovács, S., Kiss, A., Szekeres, J., 2006. Experiences in application of HEC-RAS model under circumstances of flood waves. In: Schanze, J., Zeman, E., Marsalek, J. (Eds.), Flood Risk Management: Hazards, Vulnerability and Mitigation Measures, pp. 47–58.
Liu, X., 2008. Airborne LiDAR for DEM generation: some critical issues. Progress in Physical Geography 32 (1), 31.
Mandlburger, G., Hauer, C., Höfle, B., Habersack, H., Pfeifer, N., 2009. Optimisation of LiDAR derived terrain models for river flow modelling. Hydrology and Earth System Sciences 13 (8), 1453–1466.
Merwade, V., Maidment, D., Hodges, B., 2005. Geospatial representation of river channels. Journal of Hydrologic Engineering 10 243.
Merwade, V., Cook, A., Coonrod, J., 2008. GIS techniques for creating river terrain models for hydrodynamic modeling and flood inundation mapping. Environmental Modelling & Software 23 (10–11), 1300–1311.
Pramanik, N., Panda, R., Sen, D., 2010. One dimensional hydrodynamic modeling of river flow using DEM extracted river cross-sections. Water Resources Management 24 (5), 835–852.
Price, W., 1983. Global optimization by controlled random search. Journal of Optimization Theory and Applications 40 (3), 333–348.
Price, R., 2009a. An optimized routing model for flood forecasting. Water Resources Research 45 (2), W02426.
Price, R., 2009b. Volume-conservative nonlinear flood routing. Journal of Hydraulic Engineering 135 (10), 838–845.
Qi, H., Altinakar, M.S., 2011. A GIS-based decision support system for integrated flood management under uncertainty with two dimensional numerical simulations. Environmental Modelling & Software 26 (6), 817–821.
Roux, H., Dartus, D., 2004. Data Assimilation Applied to Hydraulic Parameter Identification, British Hydrological Society International Conference 2004, Hydrology: Science and Practice for the 21st Century. London, United Kingdom, pp. 354–361.
Roux, H., Dartus, D., 2008. Sensitivity analysis and predictive uncertainty using inundation observations for parameter estimation in open-channel inverse problem. Journal of Hydraulic Engineering 134 (5), 541–549.
Samuels, P., 1990. Cross section location in one-dimensional models. In: White, W.R. (Ed.), International Conference on River Flood Hydraulics. Wily, Chichester, UK, pp. 339–350.
Sanders, B., 2007. Evaluation of on-line DEMs for flood inundation modeling. Advances in Water Resources 30 (8), 1831–1843.
Schwanghart, W., Kuhn, N.J., 2010. TopoToolbox: a set of Matlab functions for topographic analysis. Environmental Modelling & Software 25 (6), 770–781.
Solomatine, D., 1995. The Use of Global Random Search Methods for Models Calibration. XXVIth Congress of the International Association for Hydraulic Research (IAHR), London, UK, pp. 224–229.
Solomatine, D., 1999. Two strategies of adaptive cluster covering with descent and their comparison to other algorithms. Journal of Global Optimization 14 (1), 55–78.
Tóth, S., Kovács, S., 2007. River capacity improvement and partial floodplain reactivation along the Middle-Tisza: scenario analysis of intervention options, FLOODsite Report Number T22-07-01. HEURAqua/KÖTIKÖVIZIG.
Tesfa, T.K., Tarboton, D.G., Watson, D.W., Schreuders, K.A.T., Baker, M.E., Wallace, R.M., 2011. Extraction of hydrological proximity measures from DEMs using parallel processing. Environmental Modelling & Software 26 (12), 1696–1709.
Vaze, J., Teng, J., Spencer, G., 2010. Impact of DEM accuracy and resolution on topographic indices. Environmental Modelling & Software 25 (10), 1086–1098.

Environmental Engineering and Management Journal December 2012, Vol.11, No. 12, 2305-2311
http://omicron.ch.tuiasi.ro/EEMJ/

"Gheorghe Asachi" Technical University of Iasi, Romania

DECISION SUPPORT SYSTEMS FOR FLOOD MANAGEMENT IN THE TIMIS BEGA CATCHMENT

Ioana Popescu[1*], Andreja Jonoski[1], Lucretia Bociort[2]

[1]*Department of Integrated Systems and Governance, UNESCO-IHE Institute for Water Education, P.O. Box 3015, 2601 DA Delft, The Netherlands*
[2]*Romanian Water National Administration Banat, Regional Water Branch, Timisoara, Mihai Viteazul Street 32, Romania*

Abstract

To better monitor the progression of a major flood and its likely impact on people and properties, Romanian Water Boards are in need to build sophisticated Flood Emergency Decision Support Systems (DSS). The aim of such a DSS is to integrate the hydrologic, hydraulic and GIS flood maps, assessed over the last 10 years and extract and present vital flood emergency decision making information during an actual event. This paper outlines the DSS strategy and shows how the DSS should extract and display all the various components of information to various groups, which are required to respond to a flood emergency event in the Timis Bega catchment. Example is given of a framework of a collaborative project between Romania and the Netherlands, where a demonstrator of a flood forecasting system was developed, which can support operational water management under extreme conditions when actions have to be taken quickly.

Key words: decision support systems, flood, flood mapping, hydrodynamic modeling, hydroinformatics

Received: September 2011, Revised final: December 2011, Accepted: December 2011

1. Introduction

1.1. Background

Floods remain one of the most frequent and devastating natural hazards worldwide. Though nowadays forecasting and warning systems significantly contribute to the reduction of losses, further prevention and avoidance of losses remains considerable, if looking at the potential given by the technological advances (Balica et al, 2013; Barredo, 2007; Jonoski et al, 2012b; Popescu et al, 2012a).

Flood modeling helps to understand flood generation and identification of the potential areas to be inundated, thus allowing for damage reduction and allowing to early warning communities downstream, which will be affected, especially in floodplains (Stewart and al., 1999; Teodosiu et al., 2009; Van et al, 2012). Moreover, modeling can be used to evaluate various flood mitigating measures in order to determine which alternative will be economically and environmentally feasible given the prevailing conditions (Jonoski and Popescu, 2011; Knight and Assad, 2007; Palomino Cuya et al., 2013; Popescu et al, 2012c).

Romanian Waters (The National Agency responsible for overall water resources management) is responsible for managing the surface and ground water resources on the Romanian territory, in accordance with both national and EU regulations, regarding water resources management and the preservation of aquatic ecosystems and water areas (Popescu et al, 2012b; Souter et al., 2009).

Flood is one of the main issues to be addressed by Romanian authorities. Their current flood forecasting and warning system is based on empirical models, which show relationships of the flow at a downstream point to that at the upstream

* Author to whom all correspondence should be addressed: e-mail: i.popescu@unesco-ihe.org; Phone: +31152151895

station (Gichamo et al, 2012; Hartanto et al, 2011; Marsalek et al., 2006; Zoccatelli et al., 2010). For example in case of managing a flood event from Banat region (in Romania) in the year 2005, there were three significant time periods with high rainfalls, with warnings provided in advance, still flooding and consequently damages occurred, which proved that Romanian Waters need better models in order to better manage flooding crisis situations (Popescu et al., 2010; Quinn et al, 2010).

In order to show the potential of using new technologies, for flood prevention, a demonstrator of a flood forecasting system was developed for the Timis and Bega river catchments, in Banat region in Romania, within a framework of a collaborative project between Romania and the Netherlands. The developed demonstrator can support operational water management under extreme conditions when actions have to be taken quickly. Timis and Bega rivers were considered jointly, because of the way the existing hydraulic structures are operated to transfer water between them, which makes their hydrodynamic unseparable. The Timis - Bega basin discharges water in Serbia beyond Romanian borders. Bilateral agreements between the two countries specify the discharge and water quality conditions downstream of the border (Stanescu and Drobot, 2005).

The present paper describes the flood decision support demonstrator build within the Dutch-Romanian project, and shows how such a system can be used for operational flood management in Romanian context. Initially the catchment selected for demonstration is described along with its past flood problems, followed by a description of the architecture of the implemented decision support system. The functionality of the demonstrator is presented in the end of the paper, along with suggestions on how be extended to prediction of dike bursts which will enable taking of proper and timely preventive and mitigation measures (Popescu et al., 2010).

1.2. Basin selection and description

One of the most flood vulnerable regions in Romania is in the west part of the country. Many rivers in this region are of trans-boundary nature, having their basins in either Romania and Serbia, or, Romania and Hungary. Any event occurring on these rivers is advected and affects the downstream country. A typical example is the Timis River, which in recent past caused severe flooding in Romania and Serbia. The flood propagation was quicker than in the past and caused severe flood damages, due to the dikes, which have been constructed along the river for flood protection.

The Timis Bega basin is located in South-West of Romania in Banat province, between 44°30' and 46° latitude, and 20°20' and 22°40' longitude. The climate of the region is temperate continental with influences from the Mediterranean basin and under the Carpathian Mountains protection, east and north, which diminish the climate influence of Eastern Europe. In Fig. 1 the basin of the two rivers is depicted.

River Bega starts in the northern part of Banat area, the at the junction headwaters of Bega Luncanilor and Bega Poieni. After starting of to the north, the river bends to the west at Coşava, finally entering the low Banat plains. Since this is the area that was frequently flooded in the past, the Bega canal was constructed with a length of 114 km, parallel to and existing canal (Bega Veche – 97 km). These two canals are connected again, downstream, on the Serbian territory. The Bega canal runs through the city of Timişoara and continues to the South-West. It has a draining area of 2,878 km².

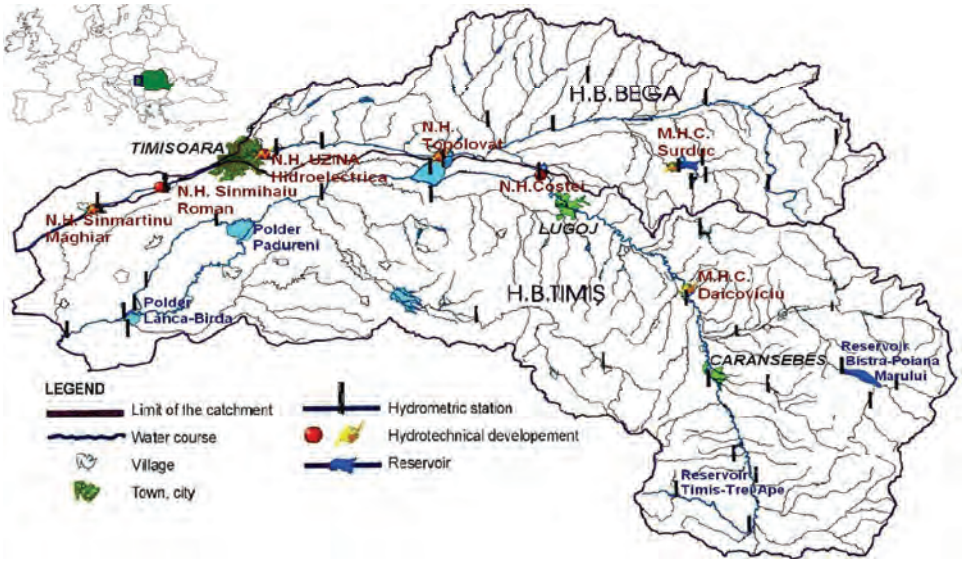

Fig. 1. Timis- Bega catchment

The Timiș River is 359 km long, rising in the Semenic Mountains (southern Carpathian Mountains) Caraş-Severin County, Romania. It flows through the Banat region and flows into the Danube in Northern Serbia. The drainage area covers 13,085 km² (in Romania 8,085 km², in Serbia 5,000 km²). After entering Banat, the river becomes slow and meandering, causing floods in rainy years (Popescu et al., 2010).

Due to recurring floods in the area, technical measures for relieving the lowlands have already been initiated 250 years ago, when flood regulation has been achieved by correcting the path of some water courses; dykes construction; construction of several permanent and non-permanent reservoirs on the watercourses or laterally; drainage of swamps; development of a drainage canals network, together with pumping stations. The most important measure was the establishment of double canal connection between Timis and Bega Rivers that allow gravitational flow diversion from one river into the other, depending on the flow conditions. Under normal flow conditions water flowing form Timis to Bega ensures minimum water supply for the city of Timişoara, while under highwater conditions on Bega, water is diverted from Bega to Timis ensuring flood protection of Timişoara.

Inspite of all these efforts, in the last century, various major flooding events have been recorded in the Timis-Bega catchment. These events were caused by dike overtopping and breaching and have been of a catastrophic nature. The recorded floods were in 1912 and 1966, at Lugoj city (the second largest city in the area) and in 2000 and 2005 at Grăniceri station, close to the border between Romania and Serbia. In 2005 there were three dyke failures; one along Timis, and two breaches along Bega, upstream of Timişoara (the largest city in the area), that resulted in the flood of about 25,000 ha spilling a water volume of about 300 - 350 million m^3 which led to severe damages (Moya Quiroga et al, 2013).

As such the flood mitigation measures proved not to be able to control and manage floods sufficiently.

2. Design of a flood decision support systems (DSS)

To better monitor the progression of a major flood and its likely impact on people and properties, the Banat Water Boards needs to concentrate in building a Flood Management Decision Support System (DSS) for floods emergencies in the area. It is anticipated that the DSS will help Romanian Water Boards to better understand the behaviour of floods and their impacts on community, and most importantly, to provide the authorities with the proper tools to take action and/or plan in case of flood events (Dinh et al, 2012).

The main objective of a flood management DSS for the Romanian Water Boards would be to extract and present vital flood emergency decision making information during an actual flood event. This can be done by generating a flood surface for an actual event by interpolating from a library of pre-computed surfaces, or by running a flood model in real time using the predicted hydrograph for the event and converting the results (Blackburn, 2002).

In essence the main aim of a DSS is to integrate the hydrologic and hydraulic data and models, of past year events, and map them using GIS tools. This does not only require extended knowledge on ICT and software development, but also attention is required for the quality of modeling and the users' demands concerning presentation and communication of model results. This is a very important aspect to be addressed while developing a DSS, because it ensures the use of the system by users (Jonoski and Popescu, 2004, Popescu et al, 2012b).

Most of DSSs consist of a three-layer system, as presented in Fig. 2, a presentation layer along with a calculation and data layer. The presentation offers the information in form of reports (tables, graphs, maps, texts), while the calculation layer uses data for preprocessing, calculatins, after-processing and saving model results in the data layer. The data layer represents a collection of relevant data and information, useful for the good functionality of the other two layers.

The three layers function independently of each other, which makes it possible to add functionality to each layer without significantly influencing the other layers. They can function on separate computer systems, and therefore management and use can be distributed (multi-computer / multi-user systems), while the layers continue communicating to each other (Popescu et al., 2010).

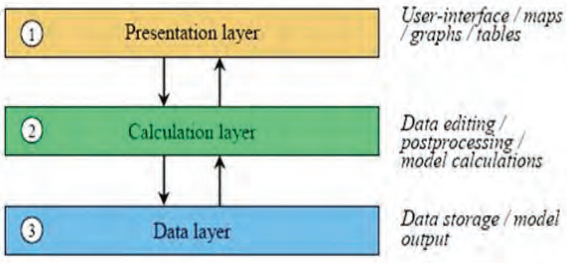

Fig. 2. Typical Decision Support Structure

Presentation layer visualizes data from a database of the data layer, comming from both model calculations as well as from on site measurements. In general the elements of a water system have spatial characteristics; therefore a geo-graphically oriented user interface should normally used for the presentation layer.

The main elements of a water system, such as rivers and canals, as well as other topographical features like roads and/or cities can be represented in a map.

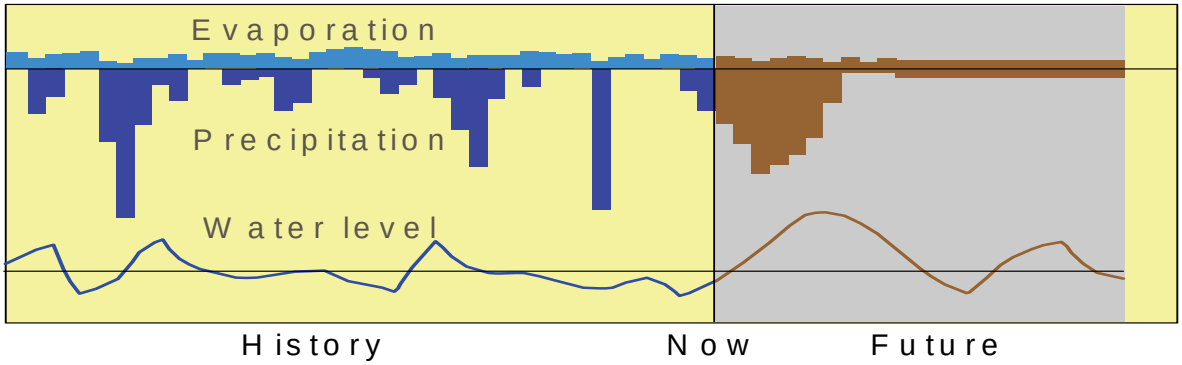

Fig. 3. Example of a generated graph of measured and forecasted water level

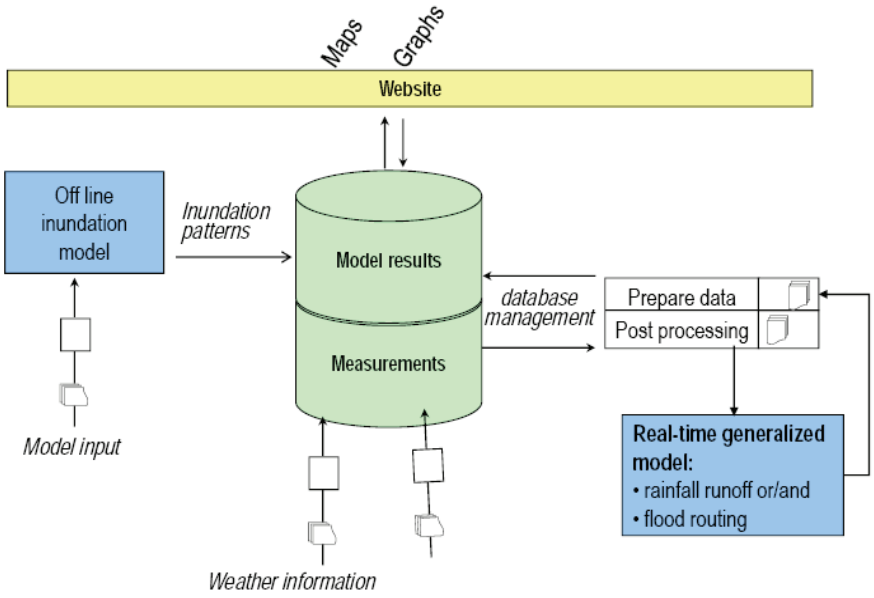

Fig. 4. The architecture of the designed Romanian flood DSS

All calculations of models (hydrological and hydraulic) are done in the calculation layer and than sent to the data layer.

The same map where the elements of the water system were represented can be further on used to present modelling results, coming from the calculation layer. These results are in form of current and forecasted water levels, discharges in river reaches, and sub-catchments hydrographs, using colored symbols in the map.

The presentation layer can use desktop GIS software like ArcGIS or a web-interface. Desktop GIS software has the advantage of being flexible and has more features, but is in general used by a limited group of users, while a web-interface can be used by a large group of users, however with fewer features and less flexibility (Rabuffetti, 2005; van der Berg and Popescu, 2005).

In the case of a flood management DSS the calculation layer is the place where the hydrological and hydrodynamic models are running. These models will calculate the behaviour of a water system, based on measured and forecasted precipitation, evaporation measured, water levels or discharges in the rivers and canals. Fig. 3 shows and example of a graph in which measured and forecasted water levels are presented in a presentation layer, based on the feedback coming from the calculation layer.

3. Flood management DSS for Timis-Bega

The Dutch-Romanian project had as aim the development of a demonstrator Decision Support System for the Timis-Bega basin. Because of the high stakes involved concerning water management during a crisis, it is important that a forecast of the emerging situation occurs in a structured and reproducible manner. Preparedness for an actual emergency involves all of the aspects of the hazard reduction, it concerns all the preparations made in anticipation of an actual operational event (Jonoski et al., 2012a).

The architecture of the Romanian flood management DSS (shown in Fig. 4) follows the basic principles of the DSS design presented above. A web interface was used, which offered the possibility to open graphs with model results, reports or tables.

The developed DSS is a flood forecasting system demonstrator, which can support operational

water management under extreme conditions when rapid action has to be taken. The demonstrator has two components:

- an on-line component consisting of a simple rainfall-runoff model, developed with the USACE HEC-HMS modelling system (a hydrological model), focusing on the on-line integration of this component with meteorological and hydrological data, which can be used for the short term operational DSS (real-time flood forecasting); and
- an off-line component, focusing on integrated flood modeling for long term analysis and design of flood mitigation alternatives, which is used for planning. The off-line component consisted of combining the HEC-HMS model for the on-line component with a 1D hydrodynamic model developed with HEC-RAS and a SOBEK 1D-2D model for flood inundation modelling.

In the modelling approach the Timis and Bega rivers were considered jointly, since their joint hydrodynamic response are conditioned by the operation of existing hydraulic structures used for water transfer between the two rivers.

As part of the on-line DSS component a hydrological model has been build for the Timis – Bega system. The model is used to forecast water levels and discharges in identified critical points of the basin. The hydrological model of the Timis- Bega catchement was built using the HEC-HMS modelling system, by discretising it in 20 sub-basins based on their common physical characteristics. The water is routed from upstream to downstream via the other components of the system such as junctions and reaches. HEC Geo-HMS which uses ArcView GIS extension was used to delineate the sub catchments from an available 30x30 m^2 digital elevation.

In the data layer of the DSS, measured discharges and precipitation, originating from the hydraulic / meteorological measurement stations are stored. These data form the dynamic input data that feed the on-line HEC-HMS model. The workflow of processing and publication of the on-line model is shown in Fig. 5.

In order to disseminate the results of the on-line model the DSS demonstrator has a web interface, which offers the possibility to open graphs with model results, reports or tables. The DSS demonstrator presents forecasted discharges for Timis and Bega rivers obtained from the developed HEC-HMS model. The webbased part of the DSS-demonstrator makes it possible to view the results on any computer with an internet connection. The webpage is secured with a password.

The pages that are part of the main Timis – Bega DSS are the following:

- main page, which shows the forecasted situation, in the river basin in a map using colored dots (where green means a normal situation is forecasted, yellow is critical and red is an alarm situation) (Fig. 6);
- page about Romanian Waters ("Apele Romane");
- page with background information on the DSS;
- page with precipitation information;
- page with computed discharges and water levels (Fig. 7);
- page with links.

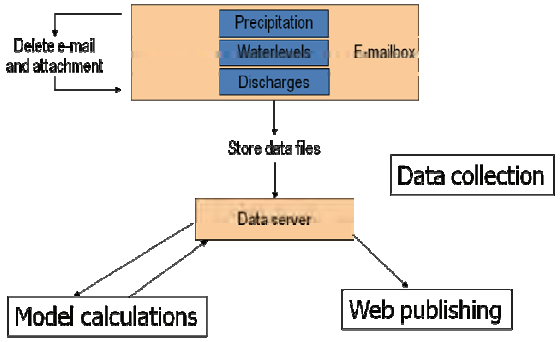

Fig. 5. Processing and publishing data of the on-line component of the DSS

The integrated model used in the off-line analysis of the system is based on a sequential combination of hydrological and hydrodynamic models. The integrated model starts with the already developed HEC- HMS hydrological model, integrated with a one-dimensional hydraulic model of the major rivers in the catchement. The hydraulic model uses the generated hydrographs from the HEC-HMS hydrologic model. The third used model is a refinement of the one dimensional model by integrating the 1D-2D SOBEK model for modelling river-floodplain interactions. The whole suite of models used comprised of HEC-HMS (0D), HEC-RAS (1D) and a 1D-2D SOBEK. All the models were calibrated and checked for sensitivity.

The rainfall runoff model HEC–HMS was used in the whole catchment and, the 1D hydraulic model HEC-RAS was used in the main channel starting at the gauging stations at Balint and Lugoj for the Bega and Timis rivers respectively and up to their outlets. The SOBEK 1D-2D was used for the floodplain starting at the gauging stations for Remetea and Brod.

The hydrograph obtained from the rainfall runoff model was fed into the hydraulic model HEC-RAS that computed the water surface elevations. This computed discharge hydrographs, of the HEC-RAS model were used as input into the SOBEK1D-2D model that simulated the floodplain inundation.

The main purpose of the off-line model was to test different mittigation alternatives, an exercise which can be done during long term planning of the flood management in the area.

4. Conclusions

In on-line situations flood modelling is central in forecasting and warning because it can help to better understand flood generation and identify the potential areas to be inundated. This allows early and on-time warning of the downstream communities located in the floodplains, which will be affected by the flood.

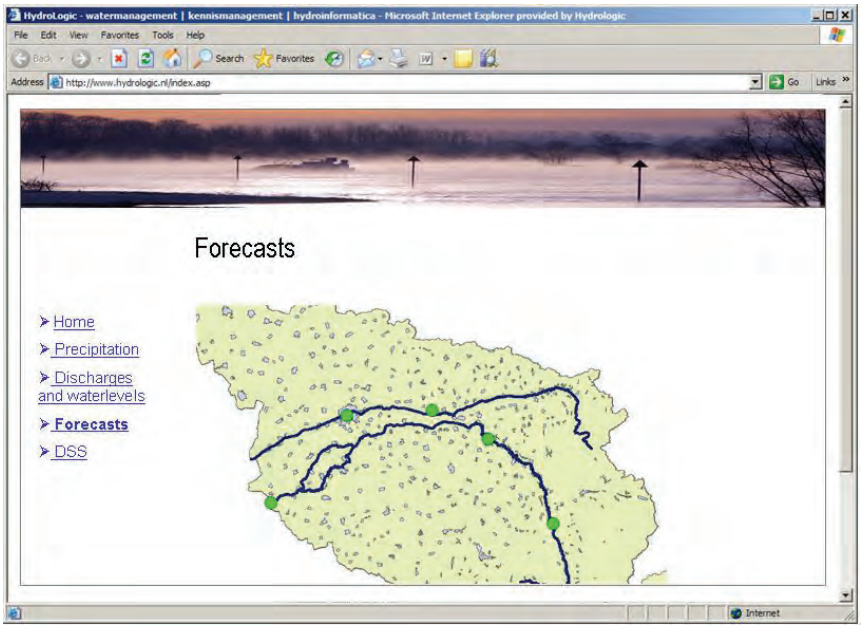

Fig. 6. Forecasted water level and their alarm levels in critical points in the basin

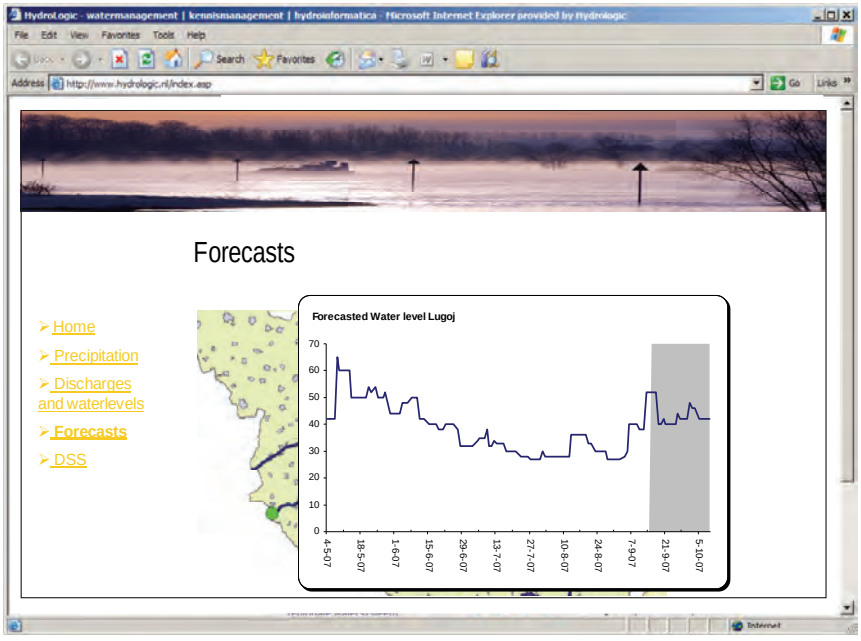

Fig. 7. Computed and foracasted water levels

In off-line situations flood modelling enables long-term planning for flood damage reduction, by determining which alternative is environmentally and economically feasible, given the prevailing conditions.

The requirements of on-line and off-line models are different. The on-line models, for flood forecasting, require fast and accurate simulation of discharge peaks for a known water system, while off-line models need to provide a physically-based reliable simulation of the water system behaviour for a wide range of conditions and for changes in the modelled water system itself, to allow design and analysis of structural flood management measures.

The DSS demonstrator developed in this study shows the potential for using such systems by Romanian authorities in their operational management of floods. The system can be further improved to present design and historic flood results integrated with property, infrastructure and community data and thus assist in the preparation of emergency disaster plans. This is an efficient way to demonstrate the possibilities of implementation of a model, with the aim of helping decision makers to understand flood propagation and take the appropriate mitigation measures in case of a flooding event. In Romania these models could be of great help to the Romanian authorities.

Aknowledgements

The work presented herein was financialy supported by the Dutch Government, through Partners for Water. Modelling data was supplied by Romanian Waters, Banat region.

References

Balica, S. F., Popescu, I., Beevers, L., Wright, N., G., (2013), Parametric and physically based modelling techniques for flood risk and vulnerability assessment: a comparison, *Journal of Environmental Modelling & Software*, **41**, 84-92.

Barredo J.I., (2007), Major flood disasters in Europe: 1950–2005, *Natural Hazards and Earth System Science*, **42**, 125-148.

Blackburn J.F.E.H., (2002), Combined flood routing and flood level forecasting, *Canadian Journal of Civil Engineering*, **29,** 64-75.

Dinh, N.Q., Balica, S., Popescu, I., Jonoski, A., (2012), Climate change impact on flood hazard, vulnerability and risk of the Long Xuyen Quadrangle in the Mekong Delta, *Journal of River Basin Management*, **10**, 103-120

Gichamo Z., G., Popescu, I., Jonoski, A., Solomatine D.P., (2012), River Cross Section Extraction from ASTER Global DEM for Flood Modeling, *Journal of Environmental Modelling & Software*, **31**, 37-46

Hartanto, I.M., Beevers, L., Popescu, I., Wright, N.G., (2011), Application of a coastal modelling code in fluvial environments, *Journal of Environmental Modelling and Software*, **26**, 1685-1695,

Knight D.W.S., Asaad Y, (2006), *River Basin Modelling for Flood Risk Mitigation*, Taylor & Francis Group, London, UK, 181-196.

Jonoski A., Popescu I., (2004), *The Hydroinformatics Approach to Integrated River Basin Management*, Proceedings of Water observation and Information System for Decision Support" Conference, Ohrid, Macedonia.

Jonoski A., Popescu I., (2011), Distance learning in support of water resources management: an online course on decision support systems in river basin management, *Water Resources Management*, **26**, 1287-1305

Jonoski A., Almoradie A., Khan K., Popescu I., Andel S.J., (2012a), Google Android mobile phone applications for water quality information management, *Journal of Hydroinformatics*, DOI 10.2166/hydro.2012.147.

Jonoski A., Alfonso L., Almoradie A., Popescu, I., van Andel, S.J., Vojinovic, Z., (2012b), Mobile phone applications in the water domain, *Environmental Engineering and Management Journal*, **11**, 919-930.

Marsalek J., Stancalie Gh., Balint G., (2006), *Transboundary Floods: Reducing Risks Through Flood Management*, Vol. 72, NATO Science Series: Earth and Environmental Sciences, Springer.

Moya Quiroga V., Popescu I, Solomatine D., Bociort L., (2013), Cloud and cluster computing in uncertainty analysis of integrated flood models, *Journal of Hydroinformatics*, **15**, 55-70, DOI 10.2166/hydro.2012.017.

Palomino Cuya D.G., Brandimarte L., Popescu I., Alterach J., Peviani M., (2013), A GIS-based assessment of maximum potential hydropower production in La Plata basin under global changes, *Journal of Renewable Energy*, **50**, 103-114.

Popescu I., Jonoski A., van Andel S.J., Onyari E., Moya Quiroga V.G., (2010), Integrated modelling for flood risk mitigation in Romania: case study of the Timis-Bega river basin, *International Journal of River Basin Management*, **8**, 269-280.

Popescu, I, Jonoski, A, Bhattacharya, B. (2012a), Experiences from online and classroom education in hydroinformatics, *Hydrology and Earth System Sciences*, **16**, 3935-3944.

Popescu, I., Archetti, F., van Andel, S.J., Giordani, I., (2012b), Lenvis: A user centric, web services based system to retrieve, analyze and deliver environmental and health information, *Environmental Engineering and Management Journal*, **11**, 889-897.

Popescu, I., Brandimarte, L., Perera, M.S.U., Peviani, M., (2012c), Assessing residual hydropower potential of the la Plata Basin accounting for future user demands, *Hydrology and Earth System Sciences*, **16**, 2813-2823.

Quinn, P., Hewet, C., Popescu, I., Muste, M., (2010), Towards new types of water-centric collaboration: instigating the Upper Mississippi River Basin observatory process, *Water Management*, **163**, 39–51.

Rabuffetti D.B.S., (2005), Operational hydro-meteorological warning and real-time flood forecasting: the Piemonte Region case, *Hydrology and Earth System Sciences*, **9**, 457-466.

Soutter M., Alexandrescu M., Schenk C., Drobot R., (2009), Adapting a geographical information system-based water resource management to the needs of the Romanian water authorities, *Environmental Science and Pollution Research*, **16**, 33-41.

Stewart M.D., Bates P.D., Andersona M.G., Price D.A., Burt T.P., (1999), Modelling floods in hydrologically complex Lowland River reaches, *Journal of Hydrology*, **223,** 85-106.

Stanescu V.A., Drobot R., (2005), Viitura din perioada 14 – 30 aprilie 2005 in bazinul hidrografic Timis-Bega, *Hidrotehnica*, **50,** 3-17.

Teodosiu C., Ardeleanu C., Lupu L., (2009), An overview of decision support systems for integrated water resources management, *Environmental Engineering and Management Journal*, **8**, 153-162.

Van P.D.T , Popescu, I., van Griensven, A., Solomatine, D., Trung, N. H., and Green, A., (2012), A study of the climate change impacts on fluvial flood propagation in the Vietnamese Mekong Delta, *Hydrology and Earth System Sciences.*, **16**, 4637–4649, DOI 10.5194/hess-16-4637-2012

Van der Berg C.C., Popescu I., (2005), An experience in knowledge mapping, *Journal of Knowledge Management*, **9**, 123 -128.

Zoccatelli D., Borga M., Zanon F., Antonescu B., Stancalie G., (2010), Which rainfall spatial information for flash flood response modelling? A numerical investigation based on data from the Carpathian range, Romania, *Journal of Hydrology*, **394,** 148–161.

Water Resour Manage (2012) 26:273–293
DOI 10.1007/s11269-011-9917-8

Model-Based Optimization of Downstream Impact during Filling of a New Reservoir: Case Study of Mandaya/Roseires Reservoirs on the Blue Nile River

K. Hassaballah · A. Jonoski · I. Popescu · D. P. Solomatine

Received: 27 August 2010 / Accepted: 20 September 2011 /
Published online: 1 October 2011

Abstract The aim of this paper is to develop a methodology based on coupled simulation-optimization approach for determining filling rules for the proposed Mandaya Reservoir in Ethiopia with minimum impact on hydropower generation downstream at Roseires Reservoir in Sudan, and ensuring power generation at Mandaya Reservoir in Ethiopia. The Multi-Objective Optimization (MOO) approach for reservoir optimization presented in this paper is a combination of simulation and optimization models, which can assist decision making in water resource planning and management (WRPM). The combined system of reservoirs is set in MIKE BASIN Simulation model, which is then used for simulation of a limited set of feasible filling rules of the Mandaya reservoir according to the current storage level, the inflow, and the time of the year. The same simulation model is then coupled with Multi-Objective optimization Non-dominated Sorting Genetic Algorithm (NSGA-II), which is adopted for determining optimial filling rules of the Mandaya Reservoir. The optimization puts focus on maximization of hydropower generation in both the Mandaya and the Roseires Reservoirs. The results demonstrate that optimal release- (and correspondingly filling-) rules for Mandaya Reservoir which maximize the hydropower generation in both Mandaya and Roseires reservoirs can be found. These rules are determined along the Pareto frontier obtained by the optimization algorithm, which can serve as a decision support tool for choosing the actual filling rule. The results also showed that the NSGA- II is an efficient and powerful tool that could assist decision makers for solving optimization problems in complex water resource systems.

Keywords Trans-boundary River · MIKE BASIN Simulation model · Reservoir operation · Multi-objective optimization (MOO) · Genetic Algorithm · Hydropower

K. Hassaballah
Ministry of Irrigation and Water Resources, Hydraulics Research Station, Wad Medani, SudanP.O. Box 318

A. Jonoski · I. Popescu (✉) · D. P. Solomatine
Department of Hydroinformatics and Knowledge Management, UNESCO-IHE Institute for Water Education, P.O. Box 3015, 2601 DA Delft, The Netherlands
e-mail: i.popescu@unesco-ihe.org

1 Introduction

Trans-boundary Rivers can bring cooperation or conflict. The choice between the two will in large part be determined by perceptions of their relative benefits or losses. Hydropower and irrigation are usually associated with these international rivers disputes. The attention and arguments between the upstream country Ethiopia and the downstream countries Sudan and Egypt, on the Nile River is one obvious example.

With construction of a new Mandaya hydropower dam and other proposed dams on the Blue Nile, in Ethiopia, attention must be given for optimizing the operation of those new dams and improving the operational efficiency of the existing reservoir systems for maximizing the beneficial uses of these reservoirs.

Quite many approaches exist to solve optimization problems in general. These approaches depend on the character of the objective function, the decision variables and the constraints. Each approach also has a different solution algorithm. If the objective function is known analytically, five main approaches are in common use: Linear Programming, Integer Programming, Dynamic Programming, Non Linear Programming and Stochastic Programming.

Reservoir operation systems analysis techniques have been reported in literature in a large number of publications during the past several decades. It has been viewed as an area of water resource planning and management having particularly high potential for beneficial application of optimization models.

Since most of the optimization problems in engineering and water resources are non linear, have multiple extreme values (local minima or maxima) and may not be known analytically (so that gradients cannot be calculated), new techniques and approaches of adaptive heuristic methods have been developed (Baretto et al. 2006). Recently these methods are used coupled with simulation models, when the objective function is not known analytically, and its values are provided by running given simulation model, which is included within the optimization loop. These recent optimization methods are mainly inspired by certain principles of nature, such as the evolutionary principle for example, which is the basic idea for the class of Evolutionary optimization methods.

1.1 Mathematical Models

The literature related to mathematical models, in general, and its application on reservoir operations, in particular, is extensive. Water related mathematical modelling, as well as the programming associated with the implementation of these models, is covered by numerous operations research and mathematics books, as well as by water resources system analysis textbooks such as Loucks, et al. (1981) and Mays and Tung (1992).

In case of reservoir operation models, a comprehensive in depth state-of-the-art review with strong emphasis on optimization techniques is presented by Yeh (1981). The article reviews optimization models applied to reservoir operation and management that involved traditional methods such as; linear programming (LP), non-linear programming (NLP), and dynamic programming (DP). In Needham et al. (2000), while solving a flood control problem on Iowa and DeMoines rivers, using linear programming, emphasizes that the curse of dimensionality is a significant obstacle for solving optimization problems by using dynamic programming (DP) when the dimension of decision variables is large.

Celeste et al. (2004) recommended that the nonlinear programming (NLP) is a more general formulation and effective for handling the nonlinearities. However the mathematical requirements of the NLP are more complicated and its standard algorithms are usually

based on calculus and not always efficient for complex problems containing non-convex, discontinuous and non-differentiable functions.

Cristiano et al. (2006) offered a numerical solution to a large-scale water reservoir network of a 30-dimenional optimization problem using stochastic dynamic programming (SDP). The results showed how dimensionality issues given by the large number of basins and realistic modelling of the stochastic inflows can be mitigated by employing neural approximations for the value functions, and efficient discretizations of the state space.

Joshi and Gupta (2009) used a simulation model for the operation of the multipurpose multi-reservoir systems of the Narmada River Reservoir system in India. The model has been formulated for monthly operation of reservoir to meet industrial, domestic, irrigation and hydropower generation water requirements. Multisite Lane's Applied Stochastic Techniques (LAST) (Lane and Frevert 1990) inflow generation models, have been used for the generation of the monthly inflow series.

Mohan and Raipure (1992) developed a linear multi-objective programming model to derive water releases from five reservoirs for irrigation and hydropower production for optimal operation of the reservoirs in the Chaliyar River Basin, Karnataka State, India.

1.2 Evolutionary Computations

Evolutionary Computation is a subfield of artificial intelligence, specifically computational intelligence. Evolutionary Computation uses iterative progress, such as growth or development in a population. This population is then selected in a guided random search to achieve the desired end. Such processes are often inspired by biological mechanisms of evolution.

The most popular of such evolutionary techniques is Evolutionary Algorithm (EA), which mimics nature's evolutionary principle to drive its search towards an optimal solution. One of the unique features of EAs is that, in each iteration it uses a population of multiple solutions instead of a single solution. Therefore, the outcome is also a population of solutions. In this way if the optimization problem has a single optimum, then all population members can be expected to converge to that optimum solution. However, if the optimization problem has multiple optimal solutions, an EA can give multiple optimal solutions in its final population. This ability of an EA to find multiple optimal solutions in a single simulation run makes EAs distinctive in solving multi-objective optimization problems (Deb 2002).

Ngo et al. (2007) proposed control strategies for the Hoa Binh Reservoir, Vietnam, by applying a combination of simulation and optimization models. The control strategies are setup in the MIKE 11 simulation model to guide the reservoir releases according to the current storage level, the hydro-meteorological condition and the time of the year. A heuristic global optimization tool, the Shuffled Complex Evolution (SCE) algorithm, is used for optimizing the reservoir operation.

Oliveira and Loucks (1997) presented an approach for optimizing operation rules for a multi-reservoir system using genetic algorithms (GAs). Chen (2003) successfully applied a real coded GAs in combination with a simulation model to optimize 10-day operation rule curves of a major reservoir system in Taiwan. The results showed that the method can be powerfully used to optimize the rule curves, and it is not limited by the type of objective function and simulation model applied.

Celeste et al. (2004) used genetic algorithms for the operation of a water resource system in real time. A genetic algorithm was developed and applied to solve an optimization model for the operation of the system responsible for the water supply of Matsuyama City, in

Japan. For comparison purposes, the same model was solved by a technique based on calculus and the Shuffled Complex Evolution (SCE) algorithm.

Azamathulla et al. (2008) developed and carried out a comparison of two models - a Genetic Algorithm (GA) and Linear Programming (LP), to be applied to real-time reservoir operation in an existing Chiller reservoir system in Madhya Pradesh, India. Their performance is analysed, and from the results, the GA model is found to be superior to the LP model.

East and Hall (1994) applied GA to a water resource system of four reservoirs. The objective is to maximize the benefit from power generation and irrigation water supply, subject to constraints on storage and releases from reservoirs. Sharif and Wardlaw (2000) applied a genetic algorithm for the optimization of a multi-reservoir system in Brantas Basin, Indonesia by considering the existing development situation in the basin and two future water resource development scenarios. They proved that GA is able to produce solutions very close to those produced by dynamic programming.

Kim and Heo (2004) applied multi-objective GAs to optimize the multi-objective system of the Han River Basin in South Korea. Ahmed and Sarma (2005) developed a GA model to derive the optimal operation policy and compare its performance with that of stochastic dynamic programming (SDP) for a multi-purpose reservoir.

Chang (2008) proposed a reservoir flood control optimization model with a linguistic description of requirements and existing regulations for rational operating decisions. The approach involves formulating reservoir flood operation as an optimization problem and using the genetic algorithm (GA) as a search engine.

Solomatine and Avila Torres (1996) applied an approach of models approximation to optimize a system of three reservoirs located in the Apure river basin in Venezuela having a multi-criteria decision making (MCDM) problem. The hydrodynamics and hydrology of the basin was modelled using MIKE-11 modelling software. The model of the basin was approximated using an artificial neural network generator (ANN) in order to run it in the optimization loop.

Different evolutionary algorithms have been tested to solve the multi-objective optimization problems. Deb et al. (2002) developed NSGA-II as an improvement of NSGA, and compared its performance with two other elitism multi-objective evolutionary algorisms: Pareto-Archived Evolution Strategy (PAES), and Strength- Pareto Evolution Algorithm (SPEA). The result shows that, NSGA-II (Non-dominated Sorting Genetic Algorism) was able to maintain a better spread of solutions and converge better in the obtained non-dominated front compared to the others two elitisms.

Hakimi-Asiabar et al. (2010) proposed a new model of Self-Learning Genetic Algorithm (SLGA), which is an improved version of the SOM-Based Multi-Objective GA (SBMOGA). The proposed model is used to derive optimal operating policies for a three-objective multi-reservoir system. To evaluate the applicability and efficiency of the proposed methodology, it is used for developing optimal operating policies for the Karoon-Dez multi-reservoir system, which includes one-fifth of Iran's surface water resources.

The differences between the above presented applications there are two distinctive approaches on how decision variables are selected, one approach uses the reservoir releases as decision variables, while others are using the operating rule parameters.

In view of the above presented recent literature review the main conclusion is that, while traditional optimization methods such as LP, NLP or DP are still being used, very promising results can be obtained by search-based algorithms (e.g. GA), in which the optimal solution is sought through intelligent search through the solution decision space, leading to continuous improvement of the objective function. Based on this conclusion, the

present research looked into the analysis of the reservoir operation using evolutionary algorithms. Frequently the evaluation of the objective function is provided by a simulation model. Given that the simulation model is not too demanding in terms of computing time, this coupled simulation-optimization approach can lead to successful solutions for a variety of problems.

2 Case Study Description

The water resources of the trans-boundary river basin represent a single entity, and must be treated as a whole to avoid conflicts over the utilization of this limited resource. Water resources management of the trans-boundary river basin requires optimal utilization of the available resources given the competing demands specially when a new reservoir was constructed or planned to be constructed. The water demands, distributed in space and time, must be matched with the resources, with their own temporal and spatial distribution.

In recent years a number of hydropower dams have been planned to be constructed on the main Blue Nile River within the Ethiopian plateau. In Ethiopia four large hydroelectric dams, and reservoirs, were proposed, on the Blue Nile. Those dams, from upstream to downstream, are: Karadobi, Mabil, Mandaya, and Border.

Figure 1 shows the location of Roseires and Mandaya reservoirs on the Blue Nile.

The existing Roseires reservoir, in Sudan (Fig. 1.), started to operate in 1966, in the town of Roseires on the Blue Nile, about 500 km southeast of the Sudanese capital Khartoum. The lake resulted, from the dam construction, is 75 km in length and has a total area of 280 km^2, with maximum averaging depth of 50 m. After the completion of the dam heightening, the area is expected to increase. The total storage volume of the reservoir at the time of start of operation was approximately 3,024 million m^3. Due to the high sediment load of the Blue Nile during the flood season, from June to September, the reservoir is operated for about 4 months at Normal Minimum Operating Level (467 masl), to route the majority of the sediment through the reservoir with minimum deposition. At the end of the flood season, the reservoir is impounded to the Maximum Operating Level (481 masl), to store sufficient water for the dry season, from October to May. In spite of this particular operation policy, a storage volume of approximately 1,000 million m^3 are already lost, due to sedimentation.

The proposed Mandaya dam site, in Ethiopia (Fig. 1), is located at the downstream boundary of the Blue Nile gorge (about 700 km from Lake Tana) and at a point where the steep topography of the gorge opens out into gentle slopes and more rising and falling terrain. The lake which will result from the proposed dam has a total area of 736 km^2, with maximum depth of 200 m. The total storage volume of its reservoir is expected to be about 49 billion m^3 with Minimum Operating Level (MOL) of 760 masl, and Maximum Operating Level of 800 masl.

In the pre-feasibility study of the Mandaya dam, conducted by EDF-Generation and Engineering Division, and Scott Wilson Consultancy, simulations have been performed to assess the impact on the hydropower generation downstream within Sudan and Egypt during the filling period of the proposed Mandaya Reservoir. The results demonstrate that the hydropower generation within the existing Sudanese dams, is significantly reduced by about 25%, in the first 2 years. Developing an optimal filling policy for Mandaya Reservoir is required in order to to maximize the hydropower generation in the existing downstream Roseires reservoir, in Sudan. Moreover during the period of filling of Mandaya reservoir, power generation needs to be ensured at Mandaya.

Fig. 1 The study area

In this paper, model-based optimization for multi-objective reservoir optimization is used for determining filling policy for the new reservoir, Mandaya, on a trans-boundary river between Ethiopia and Sudan.

3 Modelling Approach and Methodology

Simulation and optimization modelling processes provide quantitative information for use in a typically complex decision making process. Problem formulation for model-based

analysis requires a good understanding of the real problem and the overall decision making process.

A method is required to analyze the distribution of the water resources and demands, firstly for the existing condition of development with reservoirs, irrigation schemes, water supplies and diversions, and secondly for expected future development states in the basin.

In order to analyze the present and predict the future water resources in the Blue Nile basin, in the case of the connected system of Roseires and Mandaya reservoirs, first a simulation model, developed in Mike Basin was used, and secondly the simulation model was coupled with the NSGA-II optimization model in order to determine possible optimal filling rules of Mandaya reservoir.

3.1 The Simulation Model

Since the sources and demands are spatially distributed throughout the Nile Basin basin, a Geographical Information System (GIS) represents an appropriate environment in which to prepare the input data, and present the results of a simulation model. MIKE BASIN tool was chosen to built the simulation model of Roseires and Mandaya reservoirs.

MIKE BASIN is a simulation model for water allocation representing the hydrology of the basin in space and time, as defined in the DHI 2005 user manual. This tool is an integrated water resource management and planning computer model that integrates GIS (specifically ArcGIS) with water resource modelling. This gives managers and stakeholders a framework within which they can address water allocation and environmental flow in a river basin. In this particular study GIS was not used for analysis, it was just used to prepare the input data for the simulation model.

To represent the current and future situations of the reservoirs system and operation policies as closely as possible, substantial efforts have been made to collect data from relevant institutions in Ethiopia and Sudan. The collected data include hydrological data and the physical characteristics of the existing and proposed reservoirs. The available data for the existing Roseires site covered a period of 39 years (1965–2003) in daily intervals, while the available data for the proposed Mandaya site covered a period of 50 years (1954–2003) in monthly intervals. Because the data on the operating rules of the existing Roseires reservoir, is available on a 10- day interval, all the collacted data was converted into 10-day average intervals (three periods per month). Microsoft spreadsheet and Delphi Code Gear have been used for data processing and MIKE Zero has been used for data storing.

The Mike BASIN model consists of two reservoirs (Fig. 1) connected via a reach (Blue Nile). The Blue Nile reach is divided in several computational points, in which the continuity of mass is applied in each computational time step. The input data for the model are hydrological data, physical characteristics of the reservoirs along with their operation rules.

3.1.1 Hydrological Data

The collected hydrological data consists of inflow, irrigation demands, evaporation and water levels at both the Mandaya and the Roseires reservoirs. Downstream river water level and flow discharge from Roseires reservoir have been used to derive rating curves relations. The rating curves are used to find the downstream water levels for a given discharge passing through the dam, then the downstream water level is subtracted from the upstream water level to get the head difference which is required for hydropower computation. Since Mandaya Reservoir does not exist yet (it is just proposed), no tailwater table has been

specified to be used for hydropower computations. Therefore, the simulation model computed the head difference automatically as a difference between the water level and the bed level of the reservoir.

3.1.2 Flood Control Operation Rules

One special characteristic of the Blue Nile River is the high flow during the flood season and the low flow during the dry season. The operation of the reservoirs is designed to pass the large annual flood volume through the dams during the flood period from Jun 15 to September 15. It is estimated that 80% of the annual flow of the river occurs in those 3 months.

During the flood season, the water level in the Roseires reservoir is lowered down to 467 masl using the deep sluices. The spillway will expand the deep sluice flow, when it becomes necessary to pass the flood peak. As can be seen in Fig. 2 the filling of the Roseires reservoir usually starts in September, after the peak flood has passed, and lasts for about 45 days, until top retentions level is reached at 480 masl. During the filling of the reservoir, the spillway gates are closed and the deep sluices control the flow, until the reservoir filling is completed. Since 1981, the top retention level has been increased by 1 m, to 481 masl, and later the minimum operation level has been increased to 467.6 masl.

The same policy of operation has being used for the Sennar reservoir, downstream of Roseires reservoir (see Fig. 1). For consistent operation of the system, the simulation model used the same operation policy of the existing Roseires and Sennar reservoirs, for the proposed Mandaya reservoir. with minimum operation level of 760 and maximum of 800 masl respectively.

3.1.3 Physical Characteristics of the Reservoirs

The Level-Area-Volume relation tables for Roseires and Mandaya reservoirs have been specified for the model in order to compute the water levels in the reservoirs, the evaporation losses from the reservoirs and the precipitation over the reservoirs areas.

3.1.4 Hydropower Plants Parameters

Mandaya Reservoir Mandaya hydropower plant is planned to operate with 8 Francis vertical shaft turbines, with rated output of 250 MW per machine, which gives a total

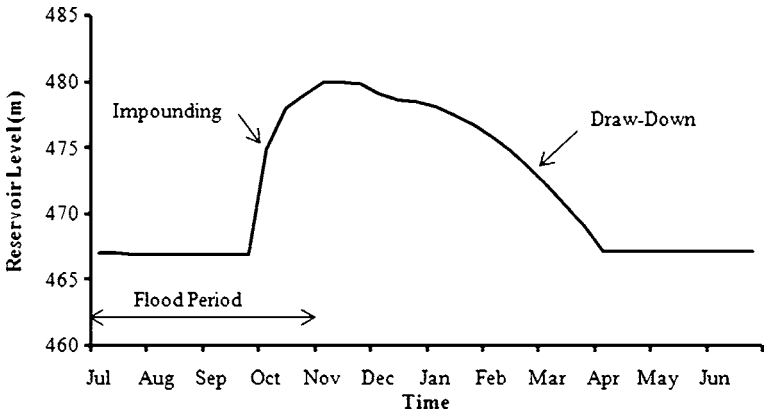

Fig. 2 Typical operation program of Roseirse reservoir

installed capacity of 2,000 MW. The design net head is assumed to be 171.9 m and the minimum head for operation is assumed to be 149 m. The main objective of this study was to maximize the hydropower production, therefore the reference value for the power, with respect to which the deficit was calculated, was set to be 2,000 MW (the same as the installed capacity) from the very beginning of the plant operation. This may be somewhat uncertain, because usually the power generation engines are installed gradually (in phases), but in absence of information about the installation plan, the assumption was made that the plant has full installed power from the very beginning of its operation. For all simulations it is assumed that every turbine is taken out of the service for a period of 2 weeks, each year, for maintenance purpose. Those weeks were selected to be during the dry period.

Turbine capacity (TC) is computed using the following formula:

$$TC = IC/g.\rho.\varepsilon.ANH = TC = 2000*10^6/9.81*1000*0.86*180 = 1317m^3/s \quad (1)$$

Where, TC is the turbine capacity (m^3/s), IC is the installed capacity in (MW), g is the gravitational constant (m/s^2), ρ is the density of water (kg/m^3), ε is the machine efficiency (−), and ANH is the average net head (m).

Roseires Reservoir At Roseires dam, there are seven vertical Kaplan turbines installed, with rated output of 40 MW per machine, which gives a total installed capacity of 280 MW. The average net head from the available data is found to be 31 m and the minimum head for operation is 18 m. The average power generated output for the last 15 years available data (1989–2003) is found to be 1026.12 GWh/year. The simulation model used the actual generated power as target power. As in the case of Mandaya powerplant, it is assumed, in all simulations, that every turbine is taken out of the service for a period of 2 weeks, each year, for maintenance purpose. Those 2 weeks are selected to be during the dry period.

Acording to Eq. (1) turbine capacity for Roseires is:

$$TC = 280*10^6/9.81*1000*0.95*31 = 969m^3/s$$

3.2 The Optimization Model

This study is focused on Multi-Objective Optimization (MOO) model framework development and its application in reservoir optimization. The computational tools used for the optimization process are the coupled models of MIKE BASIN and the NSGA-II genetic algorism optimizer. To link the MIKE BASIN simulation model engine and the optimization model (NSGA-II), a code was written in MATLAB to run the MIKE BASIN in a loop, and to guide the optimization process to calculate and pass the objective functions to NSGA-II for evaluation.

3.2.1 Objective Functions

Two conflicting objective functions are considered to be used by the MOO, namely:

- Maximization of generated power in Mandaya reservoir for a period of 5 years; and
- Maximization of generated power in Roseires Reservoir for a period of 2 years.

The objective functions are computed as $Max[P_{Mandaya}, P_{Rosaires}]$, where, P is the cumulative generated power. The decision on number of years for the two objective

functions is coming from the fact that the results of all four simulation scenarios, of the simulation model, show that in the worst case scenario, the deficit in hydropower generation in Roseires reservoir occurs only during the first 2 years of the filling of Mandaya reservoir and then the following there more years, will generate additional gain. Therefore the objective function is computed for a time horizon of 2 years for power generation in the existing Roseires reservoir. On the other hand, the simulation results show that in the worst case Mandaya reservoir can reach its normal operation level (between 780 and 800 masl) only after 5 years of starting operating. Therefore, it has been decided for Mandaya reservoir to maximize the power generation for a time horizon of 5 years.

3.2.2 Decision Variables

Decision variables are selected to be 36 monthly values of control flows from Mandaya reservoir. These 36 values are randomly generated, within the range of 100 m^3/s and 1,000 m^3/s, and they are used as the Minimum Remote Flow Control Rule. The 100 m^3/s is the minimum environmental release, computed as 10% from the annual average flow at Mandaya site, while the 1,000 m^3/s is the capacity of the bottom outlet of the same reservoir. These 36 values are monthly control flow values distributed over the first 3 years of the period of filling to ensure that at least a minimum environmental release downstream of Mandaya would be satisfied. Several simulation scenarios were performed with the simulation model, before using it in the optimisation model. The results of these simulations demonstrate that after the first 2 years of operation of Mandaya reservoir, the minimum environmental release would be satisfied through the turbine outflow, therefore it has been decided to select decision variables, by specifying the minimum environmental flow, only in the first 3 years, hence 36 decision variables. Figure 3 presents the upper and lower boundaries of the decision variables. The upper and lower boundaries of the decision variables are actually used by the simulation-optimisation model, as constraints beside other physical characteristics of the system, such as operating rules of the reservoirs, volume of reservoirs, etc.

3.3 The Interlink between the NSGA-II and MIKE BASIN Model

A code was written in MATLAB which plays a key interlinking role between the optimization model (NSGA-II) and MIKE BASIN simulation model, through its computational core "engine". The code incorporates analyses the calculated values of the objective functions and uses the results for the purpose of optimization. The general step by step procedure and the functions of the MATLAB interlink code, are:

- In the code, the decision variables which are randomly generated in NSGA II, are real values of 36 monthly control flows from Mandaya reservoir, with values ranging between 100 and 1,000 (m^3/s). Because the operating rules of both Roseires (existing reservoir) and Mandaya (proposed reservoir) are on the basis of a 10-day time step, the MIKE BASIN simulation model is set-up with a 10-day time step of computation. In order to run the MIKE BASIN simulation model, as per the set-up time step, the 36 decision values are assigned, each one three times, so that the total number of variables becomes 108 values of 10-day control flows over a period of 3 years.
- The code enables to introduce the 108 generated variables, as input values for the MIKE BASIN simulation model. At each run new values for decision variables are generated by the NSGA-II algorithm, and subsequently a MIKE BASIN simulation is carried out for each generated set of values (each population member).

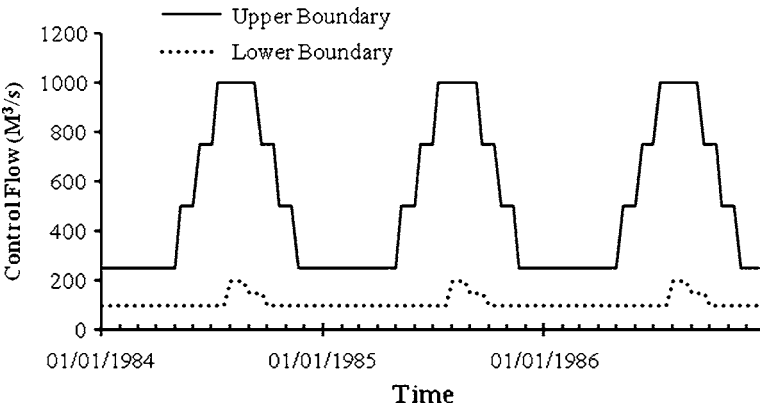

Fig. 3 Upper and lower boundaries of the decision variables

- Hydropower calculations are performed automatically by the MIKE BASIN run using head difference and release at each time step.
- The cumulative hydropower production values are extracted after each MIKE BASIN run.
- After each MIKE BASIN model run and, the results coming from the computed objective functions (cumulative generated power) are passed into NSGA-II for evaluation and generation of new set of decision variables.

Based on these steps, the general modelling workflow of the Multi-Objective Optimization model is shown in Fig. 4 and detailed below:

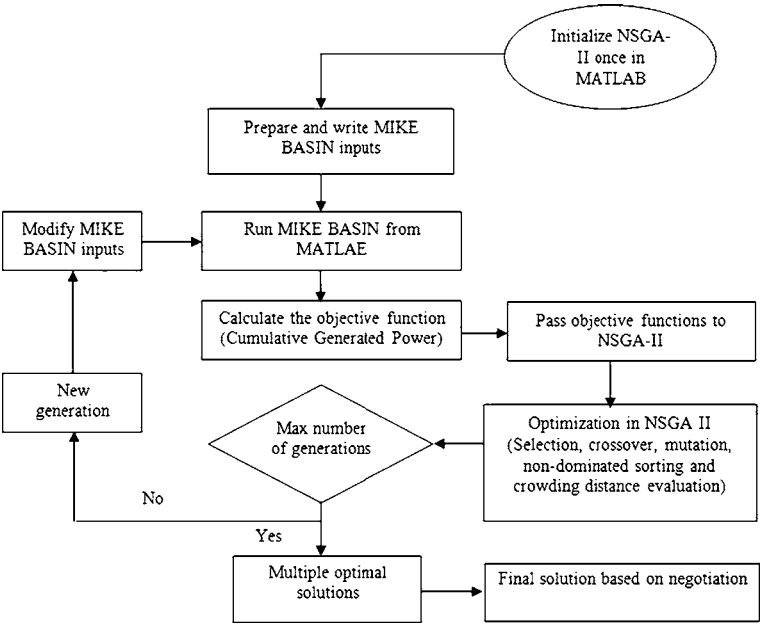

Fig. 4 General algorithmic framework of the multi-objective optimization model

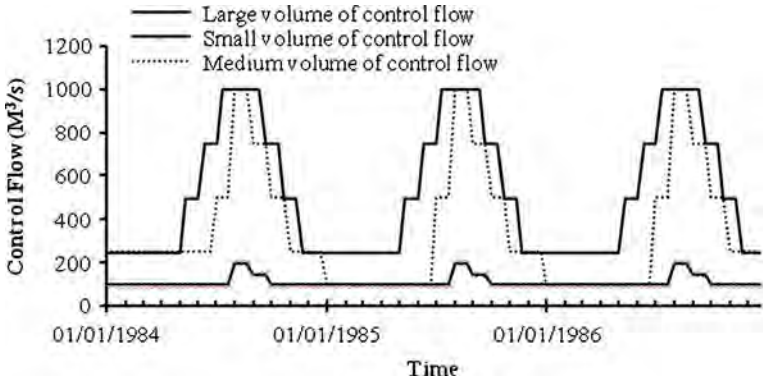

Fig. 5 Control flow rules

I. Initialize NSGA II for first generation
II. Modify MIKE BASIN input according to the decision variables.
III. Run MIKE BASIN model and calculate the objective functions (generated power) for each population
IV. Pass value of objective functions to NSGA II.
V. Perform non-dominated sorting and crowding distance sorting in NSGA II
VI. Generate new offspring population based on genetic operators (selection, crossover and mutation)
VII. Repeat loop steps from II to VI until all generations are run and provide outputs.

4 Results and Discussion

4.1 Simulation Model

The simulation model was tested for four scenarios of filling of the proposed Mandaya reservoir, in order to assess the sensitivity of the power generation in the existing Roseires reservoir, under different power plant operation policies and severe hydrological conditions, in terms of wet and dry years. Looking into the results of these four scenarios it has been found that if Mandaya filling would start in the driest year recorded as data (such as 1984), the losses in hydropower generation in Roseires reservoir are much more than in those

Table 1 Cumulative generated power during the period of filling with small volume of control flow

Reservoir	Roseires			Mandaya					
	1st year	2nd year	1st + 2nd	1st year	2nd year	3rd year	4th year	First 4 years	First 5 years
Target Power (GWh)	873	883	1759						
Generated Power (GWh)	681	944	1628	0	6150	9826	8535	24355	32900
Power deficit (%)	−22	+7	−7.5						

Positive deficit = gain

Table 2 Cumulative generated power during the period of filling with large volume of control flow

Reservoir	Roseires			Mandaya					
	1st Year	2nd Year	1st + 2nd	1st Year	2nd Year	3rd Year	4th Year	First 4 years	First 5 years
Target Power (GWh)	873	883	1759						
Generated Power (GWh)	767	1187	1958	0	0	4269	7671	4269	11897
Power deficit (%)	−12	+34	+11						

Positive deficit = gain

scenarios of filling starting in the wettest year (such as 1988), as it was expected. Therefore, it has been decided that the optimization should be performed in the worst case, with filling starting in a dry period.

In order to evaluate the optimization results, three other separate scenarios of filling of Mandaya reservoir, are performed, with different release rules. These releases rules, as specified in the previous section, are not modelled as actual releases from the Mandaya reservoir, but by using Remote Flow Control rule, which is special rule in MIKE BASIN. The Remote Flow Control rule is used to impose flow values in a node far downstream from the Mandaya reservoir. These rules act like constraints in the model, because it controls the amount of flow in that particular computational node. The simulation model can achieve the value of the required flow, in the control node, in two ways, either by making direct releases through dam outlets or by making realeases as water coming out of the turbines, after producing power. The three scenarios of filling have been used here are namely: 1) small volume of control flow, 2) medium volume of control flow, and 3) large volume of control flow. These control flow rules are manually created following the same trend of the inflow hydrograph at Mandaya site but, with different volumes as presented in Fig. 5, and their values are summarized in the Table 1, 2 and 3

By comparing the three above defined scenarios of filling of Mandaya reservoir, it is clear that the more control flow is released from Mandaya reservoir—the more hydropower generation is realised at Roseires powerplant. In contrast, for Mandaya reservoir, the more control flow is generated from Mandaya reservoir, the less power generation at Mandaya site will be.

As previously mentioned the objective function (maximization of power generation) is done for 2 years for Roseires reservoir and 5 years for Mandaya reservoir. A population of 60 and 15

Table 3 Cumulative generated power during the period of filling with medium volume of control flow

Reservoir	Roseires			Mandaya					
	1st year	2nd year	1st + 2nd	1st year	2nd year	3rd year	4th year	First 4 years	First 5 years
Target Power (GWh)	873	883	1759						
Generated Power (GWh)	702	1048	1752	0	2904	8250	8535	19537	28617
Power deficit (%)	−20	+19	−0.3						

Positive deficit = gain

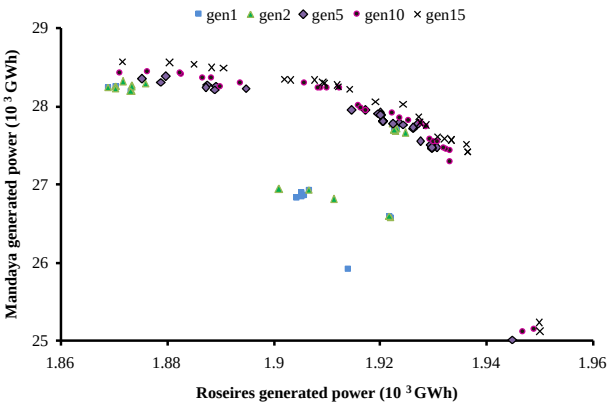

Fig. 6 Convergence of NSGA II for different generations

generations run have been carried out and it was observed that NSGA- II obtains a good shape of Pareto-optimal front and it converged reasonably quickly. The optimization run took about 3 h for one generation when it was run on a normal computer, having 2.8 GHz and 1 GB of RAM. Figure 6 shows the convergence behaviour of the algorithm in different generations, and Fig. 7 shows the result obtained from the last generation (generation 15).

The NSGA-II converges at the 10th generation and as the generations increase the Pareto optimal solutions are gradually improved- they are distributed in space and are evolving towards an optimal point. It is also observed that there is a big improvement in the solutions in early generations, while the improvement in the quality of solutions diminishes as the optimization converges. There is a big improvement between the 1st and 5th generation while the improvements between the solutions obtained at 10th and 15th generation are not significant. Therefore, it is reasonable not to use large number of generations when the

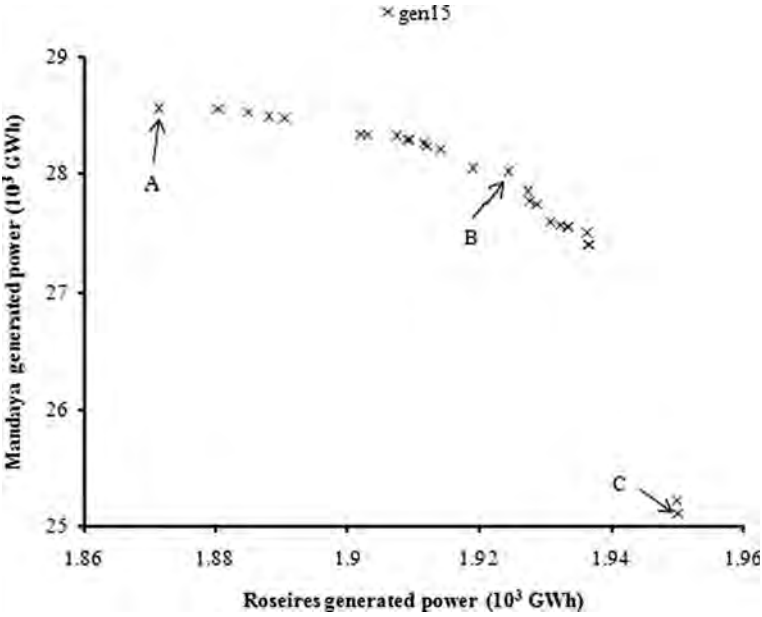

Fig. 7 Output of optimization model for 60 population and the15th generation

Table 4 Solution (A) better for Mandaya reservoir (upper left point of the Pareto)

Reservoir	Roseires			Mandaya					
	1st year	2nd year	1st + 2nd	1st year	2nd year	3rd year	4th year	First 4 years	First 5 years
Target Power (GWh)	873	883	1759						
Generated Power (GWh)	695	1172	1872	0	5343	7643	7611	20451	28561
Power deficit (%)	−20	+32	+6						

Positive deficit = gain

improvement becomes insignificant, which would help in saving extra time taken by making more iterations. On the other hand the number of populations should be increased for more improvement of the solutions.

Three feasible solutions from the Pareto front are indicated in Fig. 7: the solution from the upper left point of the Pareto front that gives highest weight to Mandaya objective function; the solution from the lower right point of the Pareto front that gives highest weight to Roseires objective function; and the solution from the middle point of the Pareto front that gives equal weight to both objective functions. The actual generated powers in both Roseires and Mandaya from these three solutions are presented correspondingly in Tables 4, 5 and 6.

4.2 Comparison between Simulation and Optimization Results

In order to evaluate the results from the optimization, three comparisons are carried out for comparing the three simulation scenarios of filling of Mandaya reservoir, as they have been defined in section 4.1 (small, medium and maximum flow control values at the control node downstream of Mandaya reservoir). The three optimization solutions, taken from the Pareto front (A,B and C) are presented in Fig. 7. The comparison is made for optimization time horizon of 2 years for Roseires reservoir, and 5 years for Mandaya reservoir.

Figures 8, 9 and 10 present the comparisons of the three cases of the control flows from the Mandaya Reservoir which have been used in the simulations (small, large, and medium control flows) and those have been obtained from the optimization for the three cases (A, B, and C). Looking at these figures three main conclusions can be drawn, which are elaborated below.

Table 5 Solution (C) better for Roseires reservoir (lower right point of the Pareto)

Reservoir	Roseires			Mandaya					
	1st year	2nd year	1st + 2nd	1st year	2nd year	3rd year	4th year	First 4 years	First 5 years
Target Power (GWh)	873	883	1759						
Generated Power (GWh)	687	1258	1950	0	4530	6749	6321	17466	25110
Power deficit (%)	−21	+42.5	+11						

Positive deficit = gain

Table 6 Solution (B) good for both Roseires and Mandaya reservoirs (Pareto midpoint)

Reservoir	Roseires			Mandaya					
	1st year	2nd year	1st + 2nd	1st year	2nd year	3rd year	4th year	First 4 years	First 5 years
Target Power (GWh)	873	883	1759						
Generated Power (GWh)	693	1121	1919	0	4677	7634	7720	19886	28051
Power deficit (%)	−20	+38	+9						

Positive deficit = gain

Firstly, by comparing the simulation scenario of filling (with small volume of control flow, which is better for Mandaya power generation) with that one from optimization (A), it is noticeable that the simulation result shows 22% deficit in power generation in Roseires reservoir during the first year of filling, while the optimization result shows 20% deficit with reduction of 2%. In the second year the simulation result shows gain of 7%, whilst the optimization result shows 32% additional gain in power generation as presented in Table 7.

Over the first 2 years the simulation model gives cumulative power of 1628 (GWh) with total deficit of 7.5%, while the optimization model gives an accumulative power of 1872 (GWh) with 13% difference in power production. In contrast for Mandaya reservoir the simulation model gives better result than the optimization as can be seen in Table 8, that is simply because the simulation model does not care about the losses of hydropower in Roseires reservoir, whilst the optimization model try to find the best solution for both Mandaya and Roseires reservoirs by taking into account maximization of hydropower in both reservoirs.

Secondly, by comparing the simulation scenario of filling (with large volume of control flow, which is better for Roseires power generation) with that one from optimization (C), it is observable that the simulation result shows 12% deficit in power generation in Roseires reservoir during the first year of filling, while the optimization result shows 21% deficit. In the second year the simulation result shows that power generation in Roseires grew rapidly to 1187 (GWh) having a gain of 34%, whilst in the optimization result power generation increased dramatically to 1258 (GWh) having 42% gain in power generation as presented in Table 9.

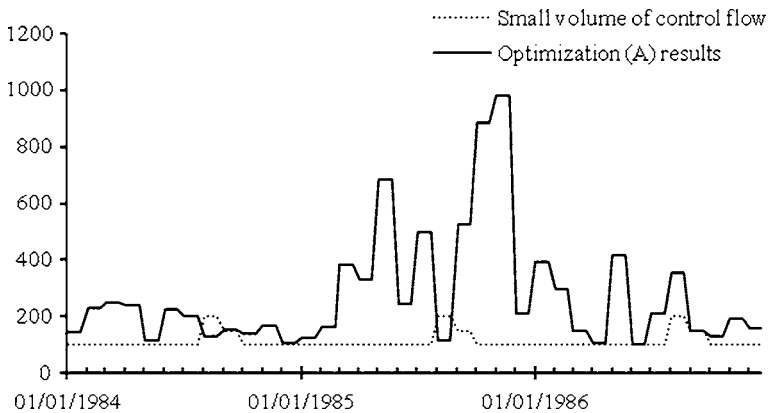

Fig. 8 Comparison of small volume of control flow with optimization (A) results

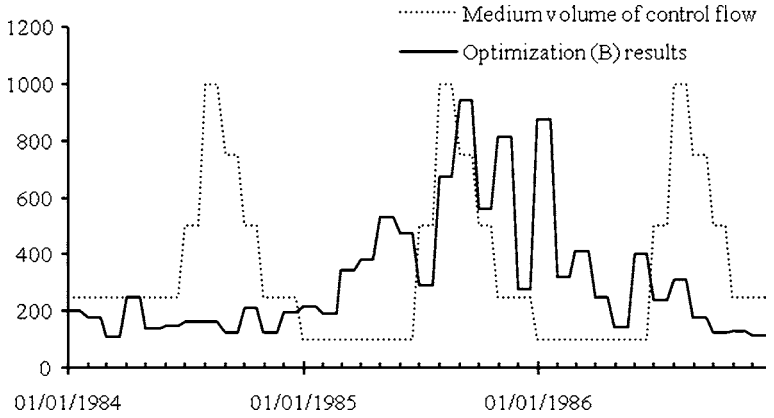

Fig. 9 Comparison of medium volume of control flow with optimization (B) results

Over the first 2 years both simulation and optimization models show total gain of 11%. On the other hand the simulation model gives an accumulative power of 11897 (GWh) in Mandaya reservoir during the first 5 years, while the optimization model gives cumulative power of 25110 (GWh) which is double the generated power produced by the simulation model as can be seen in Table 10. That is simply because the simulation model does not care about the hydropower production in Mandaya reservoir, whilst the optimization model try to find the best solution for both Mandaya and Roseires reservoirs by taking into account maximization of hydropower in both reservoirs.

Thirdly, by comparing the simulation scenario of filling (with medium volume of control flow, which is the middle solution for both Roseires and Mandaya power generation) with that one from optimization (B) (the point in the middle of the Pareto), it is noticeable that both simulation and optimization result show 20% deficit in power generation in Roseires reservoir during the first year of filling. In the second year the simulation result shows that power generation in Roseires grew rapidly to 1048 (GWh) having a gain of 19%, whilst in the optimization result power generation increased dramatically to 1121 (GWh) having 38% gain in power generation as ca be seen in Table 11. Over the first 2 years the simulation model gives cumulative power of 1752 (GWh) with insignificant deficit of 0.3%, while the

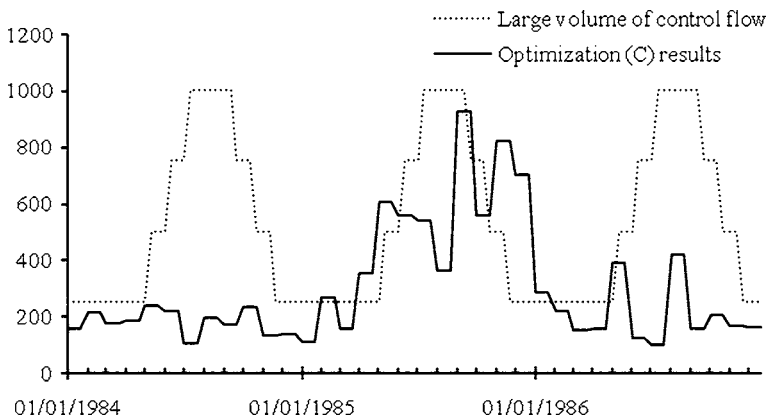

Fig. 10 Comparison of large volume of control flow with optimization (C) results

Table 7 Comparison of the simulation (small control flow) and optimization (A) results in first and second *years*

Roseires reservoir	Simulation results		Optimization results	
	1st year	2nd year	1st year	2nd year
Target power (Gwh)	873	883	873	883
Generated Power (GWh)	681	944	695	1172
Power deficit (%)	−22	+7	−20	+32

Positive deficit = gain

optimization model gives cumulative power of 1919 (GWh) with an additional gain of 9% as presented in Table 12. On the other hand the simulation model gives cumulative power of 28617 (GWh) in Mandaya reservoir during the first 5 years, while the optimization model gives cumulative power of 28051 (GWh) with no big difference.

Figures 8, 9 and 10 present the comparisons of the three cases of the control flows from the Mandaya Reservoir which have been used in the simulations (small, large, and medium control flows) and those have been obtained from the optimization for the three cases (A, B, and C).

5 Conclusion

The main objective of this study was to develop a methodology based on coupled simulation-optimization approach for determining filling policy for the proposed Mandaya Reservoir, in Ethiopia, with minimum impact on hydropower generation on the downstream Roseires Reservoir, in Sudan, while still ensuring power generation at Mandaya Reservoir. The developed approach of Multi-Objective Model-Based Optimization (MOO) can assist decision making in water resource planning and management. The normal output of MOO is a set of Pareto front optimal solutions which offer decision makers compromised/trade-off solutions to negotiate, and to agree on one solution.

A MIKE BASIN simulation model of the combined Mandaya-Roseires-Sennar reservoir system was developed prior to the optimization experiments, which served for providing initial insights into the problem by providing results from several simulation scenarios. The same model was then used for the MOO.

In this study MOO was applied for reservoir optimization. Two objective functions were considered to maximize the hydropower generation in both Roseires and Mandaya reservoir. The multi-objective optimization model developed here combined the MIKE

Table 8 Comparison of the simulation (small control flow) and optimization (A) results

Reservoir	Roseires (first 2 years)	Mandaya (first 5 years)
Generated power from simulation model(GWh)	1628	32900
Power deficit from simulation model (%)	−7.5	
Generated power from optimization model (GWh)	1872	28561
Power deficit from optimization model (%)	+6	

Positive deficit = gain

Table 9 Comparison of the simulation (large control flow) and optimization (C) results in first and second years

Roseires reservoir	Simulation results		Optimization results	
	1st year	2nd year	1st year	2nd year
Target power (Gwh)	873	883	873	883
Generated Power (GWh)	767	1187	687	1258
Power deficit (%)	−12	+34	−21	+42.5

Positive deficit = gain

BASIN simulation model and genetic algorithm NSGA-II optimizer with MATLAB environment.

Two objective functions were considered in the MOO model to maximize hydropower generation in both Roseires and Mandaya reservoirs. In future studies, it is recommended to maximize the overall benefit of hydropower generation in both Roseires and Mandaya by applying cost benefit analysis. For more improvement of the optimization solutions, it is recommended to increase the population size, but then the use of parallel computing systems might be required for reducing computational times.

This methodology can be extended to include some other existing reservoirs on the main Nile, e.g. Merawi in Sudan and High Aswan dam in Egypt. Also other proposed reservoirs can be included, e.g. Border, Karadobi or Beko Abbo in Ethiopia. In this study, due to the lack of data regarding Roseires dam heightening, the current volume of Roseires reservoir has been used. In future studies it is recommended to consider the volume of the reservoir after heightening.

From the above presented results, NSGA-II shows a good performance and capability of finding Pareto-set of optimal solutions, which helps for solving complex problems in water resources systems that could assist in decision making processes.

All the points along the Pareto front are optimal solutions. Point (A) is the best solution for Mandaya power generation, while point (B) is the best solution for Roseires power generation. Point (C) presents a solution in the middle, which can easily be agreed among the decision makers. However, these optimization results are supporting decisions rather than making them, and the final decision will depend on the negotiation. In a negociation situation, in general each party should be aware of the Best Alternative To a Negotiation Agreement (BATNA)(Soncini-Sessa et al. 2007). In such cases it is important to have extremes when negotiations, around Pareto solutions, are taking place. In the present case the two extremes cases that are: putting the releases downstream to zero at Mandaya reservoir, and the other one the long term filling of Mandaya just with the excess of water,

Table 10 Comparison of the simulation (large control flow) and optimization (C) results

Reservoir	Roseires (first 2 years)	Mandaya (first 5 years)
Generated power from simulation model (GWh)	1958	11897
Power deficit from simulation model (%)	+11	
Generated power from optimization model(GWh)	1950	25110
Power deficit from optimization model (%)	+11	

Positive deficit = gain

Table 11 Comparison of the simulation (medium control flow) and optimization (B) results in first and second years

Roseires reservoir	Simulation results		Optimization results	
	1st year	2nd year	1st year	2nd year
Target power (Gwh)	873	883	873	883
Generated Power (GWh)	702	1048	693	1121
Power deficit (%)	−20	+19	−20	+38

Positive deficit = gain

after the power generation in Roseires reservoir are not shown in the present analysis. These extremes, were not found by NSGA-II, because they are not really part of the solution space due to the way the constraints were defined. The simulation model explicitly considered the constraints that would reflect the existing agreements between Sudan and Ethiopia on releases, thus not allowing for zero releases when optimsing for Mandaya, and not allow for maximum releases when optimising for Roseires. This is not a reduction of the negotiation alternatives, because during negotiations only the possible real-life alternatives are presented not to create animosities that may even lead to water conflicts, if not something worse.

There are several issues which were not within the scope of this study, and which should be mention here, for future considerations.

Solutions found in this approach are based on the basis of the 4 years datasets, and therefore can be considered to be optimal only with assumption that the data set is representative as a worst-case scenario (along with two other scenarios). However, as presented, the choice of these 4 years was not random—the driest years within the available dataset were used. The analysis aims just to show the possible filling rules (optimal in a certain sense) for the three selected scenarios; which can form a good basis for negotiations. In reality the filling rules will be progressively updated with every forecast coming in, or possible change even due to changes in water management policies. In that case a new analysis has to be performed.

For future work more data is needed to improve the model results, especially the data that related to hydropower generation in Roseires reservoir.

Optimization carried here on was focusing on the operation of proposed Mandaya reservoir only, while it was assumed that the Roseires reservoir will be operated in the future according to the historical rules. A separate analysis might show that by changing the Roseires reservoir operation rules, as well, taking into account the new conditions, it might lead to good agreements between countries. This can be done by using the existing

Table 12 Comparison of the simulation (medium control flow) and optimization(B) results

Reservoir	Roseires (first 2 years)	Mandaya (first 5 years)
Generated power from simulation model (GWh)	1752	28617
Power deficit from simulation model (%)	−0.3	
Generated power from optimization model (GWh)	1919	28051
Power deficit from optimization model (%)	+9	

Positive deficit = gain

simulation model and extending the decision variable vector with releases from Roseires reservoir.

References

Ahmed J, Sarma A (2005) Genetic algorithm for optimal operating policy of a multipurpose reservoir. Water Resour Manag 19:145–161

Azamathulla H, Wu F-C, Ghani AA, Narulkar SM, Zakaria NA, Chang CK (2008) Comparison between genetic algorithm and linear programming approach for real time operation. J Hydro-environ Res 2:172–181

Barreto WJ, Price RK, Solomatine DP, Vojinovic Z (2006) Approaches to Multi-Objective Multi-Tier Optimization in Urban Drainage Planning. Proceedings to 7-th Conference in Hydroinformatics, Nice, France

Celeste A, Suzuki K, Kadota A (2004) Genetic algorithms for real-time operation of multipurpose water resource systems. J Hydroinform 6:19–38

Chang L-C (2008) Guiding rational reservoir flood operation using penalty-type genetic algorithm. J Hydrol 354:65–74

Chen L (2003) Real coded genetic algorithm optimization of long term reservoir operation. J Am Water Resource Assoc 39(5):1157–1165

Cristiano C, Victoria CPC, Aihong W (2006) Optimization of a large-scale water reservoir network by stochastic dynamic programming with efficient state space discretization. Eur J Oper Res 171:1139–1151

Deb K (2002) Multi-objective optimization using evolutionary algorithms. Wiley, Chichester, pp 81–112

Deb K, Pratap A, Agrawal S, Meyarivan T (2002) A fast elitist muti-objective genetic algorithm NSGA-II. IEEE Trans Evol Comput 6:182–197

East V, Hall MJ (1994) Water resource system optimizatiom using genetic algorithms. In: Proccedings of the first International Conference on hydroinformatics, Baikema, Rotterdam, The Netherlands. pp 225–31

Hakimi-Asiabar M, Ghodsypour SH, Kerachian R (2010) Deriving operating policies for multi-objective reservoir systems: application of self-learning genetic algorithm. Appl Soft Comput 10:1151–1163

Joshi GS, Gupta K (2009) A simulation model for the operation of multipurpose multireservoir system for River Narmada, India. J Hydro-environ Res 3:96–108

Kim T, Heo JH (2004) Multi-reservoir system optimization using multi-objective genetic algorithms. In: Proceeding of the World Water and Environmental Resources Congress, 27th Jun to 1st Jul. WERI, ASCE, Salt Lake City UT, pp. 1–10

Lane WL, Frevert DK (1990) Applied stochastic techniques. Personal computer version 5.2, User's manual. Earth science Division, U.S.Bureau of Reclamation, Denever

Loucks D, Stedinger J, Haith D (1981) Water resource systems planning and analysis. Prentice-Hall, Englewood Cliffs, p 559

Mays L, Tung Y (1992) Hydrosystems engineering and management. McGraw-Hill, New York

Mohan S, Raipure DM (1992) Multiobjective analysis of multireservoir system. J Water Resour Plan Manag 118(4):356–370

Needham JT, Watkins DW, Lund JR (2000) Linear programming for flood control in the Iowa and Des Moines rivers. J Water Resour Plan Manag 126(3):118–127

Ngo LL, Madsen H, Rosbjerg D (2007) Simulation and optimisation modelling approach for operation of the Hoa Binh reservoir, Vietnam. Journal of Hydrology, 336(3–4):269–281

Oliveira R, Loucks DP (1997) Operating rules for multi-reservoir systems. Water Rescour Res 33(4):839–852

Sharif M, Wardlaw R, (2000) Multi-reservoir system optimization using genetic algorithms case study. J. Comp. Civ. Eng. ASCE. 14(4):255–263

Solomatine D, Avila Torres LA (1996) Neural network approximation of a hydrodynamic model in optimizing reservoir operation. In: Proc. of the 2nd Int. Conf. on Hydroinformatics, Zurich pp. 201–206

Soncini-Sessa R, Cellina F, Pianosi F, Weber, (2007) Integrated and participatory water resources management—practice. Elsevier, p 405

Yeh W (1981) Reservoir management and operation models: Astae-of-the-Art review. Water Resour Res 21:1797–1818

Environmental Modelling & Software 26 (2011) 1685–1695

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Application of a coastal modelling code in fluvial environments

I.M. Hartanto [a], L. Beevers [a], I. Popescu [a,*], N.G. Wright [b]

[a] UNESCO-IHE Institute for Water Education, Hydroinformatics and Knowledge Management, P.O. Box 3015, 2601 DA, Delft, The Netherlands
[b] University of Leeds, Leeds LS2 9JT, United Kingdom

ARTICLE INFO

Article history:
Received 27 August 2010
Received in revised form
17 January 2011
Accepted 3 May 2011
Available online 31 July 2011

Keywords:
Numerical model
River system
XBeach
Shallow water equations
Open software

ABSTRACT

XBeach is an open source, freely available two dimensional code, developed to solve hydrodynamic and morphological processes in the coastal environment. In this paper the code is applied to ten different test cases specific to hydraulic problems encountered in the fluvial environment, with the purpose of proving the capability of XBeach in rivers. Results show that the performance of XBeach is acceptable, comparing well to other commercially available codes specifically developed for fluvial modelling. Some advantages and deficiencies of the codes are identified and recommendations for adaptation into the fluvial environment are made.

Software Availability

Name of software: XBeach
Developers: It is a public-domain model that has been developed with funding and support by the US Army Corps of Engineers, by a consortium of UNESCO-IHE, Deltares (Delft Hydraulics), Delft University of Technology and the University of Miami
Contact address: UNESCO-IHE Institute for Water Education, Westvest 7, 2611 AX, Delft The Netherlands
Availability and Online Documentation: Free download with manual and supporting material at: http://public.deltares.nl/display/XBEACH/Home
Year first available: 2004
Hardware required: IBM compatible PC
Software required: MS Windows (tested on Windows XP)
Programming language: FORTRAN 99
Program size: 4.9 MB

1. Introduction

Historically coastal modelling software has developed out of different constraints from fluvial software, due to the necessity of representing different characteristics of hydraulic behaviour. Parameters like wind and tidal forces, which have high influence in the coastal environment (de Vriend, 1991), have minor effects in fluvial environments. Conversely, a longitudinal slope and varying initial water level, which are very important in river modelling, are not considered important in coastal modelling. However, the hydraulic calculations are similar, hence coastal software can be applied in fluvial areas. The application of a code outside its original domain needs to be verified and tested comprehensively before wider application is attempted.

XBeach is open source coastal software developed to model coastal flooding, sediment transport and morphological changes in two dimensions. The software contains a number of sub-routines which solve the non-stationary two dimensional shallow water equations that are able to calculate a fluvial flood wave. Open source codes provide payment-free software (usually under the GNU Public License – http://www.gnu.org/licenses/gpl.html) to users, which is a key advantage in developing countries (Bitzer, 2004; Lanzi, 2009). This approach allows the user to modify the code to meet their specific requirements (Henley and Kemp, 2008) and can lead to a significant development and improvement of the code, whilst affording flexibility.

This research tests the validity of applying this freely available software in a cross-over domain (fluvial environments), which opens up its use to a greater number of professionals, and also permits use in coastal/river transition zones such as estuarine areas. Due to the morphological tools available within the software, specialists would also have the possibility of accessing free 2D

* Corresponding author. Tel.: +31 15 2151895; fax: +31 15 3122921.
E-mail addresses: isnaeni@gmail.com (I.M. Hartanto), l.beevers@unesco-ihe.org (L. Beevers), i.popescu@unesco-ihe.org (I. Popescu), n.g.wright@leeds.ac.uk (N.G. Wright).

1364-8152/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2011.05.014

sediment transport capabilities. XBeach has generally been used as a stand-alone model for small scale coastal applications. It has many capabilities such as: depth-averaged shallow water equations including subcritical and supercritical flow, time-varying wave action balance, wave amplitude effect and the depth-averaged advection-diffusion equations (Roelvink et al., 2009). This paper focuses solely on the depth-averaged shallow water equations solver.

The main objective of the development of the XBeach was to provide modellers with a robust and flexible environment where the concepts of dune erosion, over washing and breaching can be tested (Roelvink et al., 2009). During the code development, the stability of the numerical method was considered as a top priority. Consequently, first order accuracy was accepted since the software concentrated on representing near shore and swash zone processes which have strong gradients in time and space (Roelvink et al., 2008). Such accuracy is the norm in river modelling software.

The objective of this paper is to test the applicability of this coastal software (XBeach) in the fluvial environment. This is completed through a number of tests which are designed to recreate particular hydraulic problems encountered in fluvial flooding scenarios. The tests include comparison to semi-analytical calculations, other modelling codes and laboratory experimental results. The aim is to demonstrate that an open source approach is also applicable in the fluvial environment.

2. Theoretical background

2.1. Numerical methods

The increased demand for improved safety against flooding, prompted the development of mathematical models which describe flow propagation in rivers. These mathematical models, in most cases, do not have an analytical solutions and are solved using numerical methods (Ferziger and Peric, 1999). Flow description in rivers, lakes and coasts are long waves, which can be described by means of the so-called Shallow Water Equations. These are a hyperbolic set of partial differential equations depending on the nature of the problem to be solved. These equations describe the mass conservation and momentum conservation.

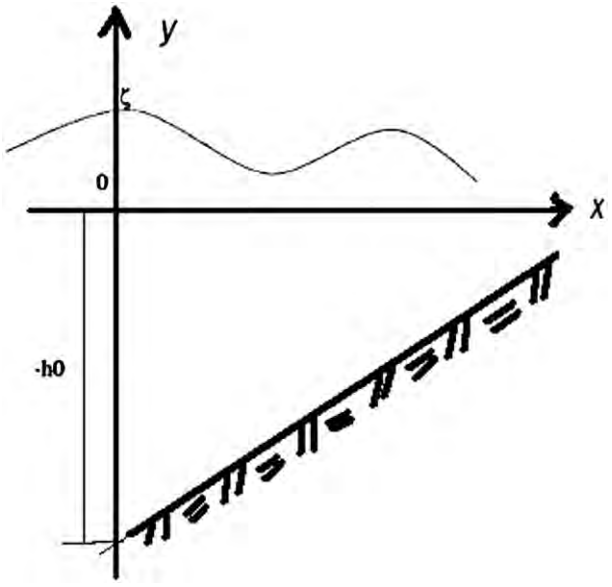

Fig. 1. Elevation and Depth for Shallow Water equations.

Significant effort during the 1980's and 1990's was devoted to defining efficient and accurate numerical methods for hyperbolic systems. Mathematically the hyperbolic equations permit discontinuous solutions and their numerical integration should lead to the computation of such discontinuities sharply and without oscillations.

The differential form of the Shallow Water equations, in the reference framework of Fig. 1, are:

$$\frac{\partial u}{\partial t}+\frac{\partial f}{\partial x}+\frac{\partial g}{\partial y}=s,\ \text{or}\ \frac{\partial u}{\partial t}+\mathrm{B}\frac{\partial u}{\partial x}+\mathrm{C}\frac{\partial u}{\partial y}=s\quad\text{with B}=\frac{\partial f}{\partial u}\ \text{and C}=\frac{\partial g}{\partial u} \tag{1}$$

Where:

$\Omega \subseteq \mathbb{R}^2$ is the domain of computation; σ is any open subset of $\mathbb{R}^2$ with boundary Γ n is the outward unit normal.

The vectors included in the equation are:

$$\mathrm{u}=\left(h,q_x,q_y\right)^T$$
$$\mathrm{f}=\left(q_x,\frac{q_x}{h}+\frac{g}{2}h^2,\frac{q_xq_y}{h}\right)^T,\quad \mathrm{g}=\left(q_x,\frac{q_xq_y}{h}+\frac{q_y^2}{h}\frac{g}{2}h^2\right)^T \tag{2}$$
$$\mathrm{s}=\left(0,gh\frac{\partial h_o}{\partial x},gh\frac{\partial h_o}{\partial y}\right)$$

With q(x,t) -the unit-width discharge,

h_o(x, y)-the depth under the reference plane in Fig. 1,

ζ(x, y, t)-the elevation over the same reference plane,

h(x, y, t) = ho + ζ

g − the gravitational acceleration

s − the source term which accounts for the bottom slope

Γ is the boundary of σ.

B and C the Jacobian matrices of the fluxes f and g respectively.

Equations (2) are the conservative form of the Shallow Water equations (all the spatial derivatives of the unknowns are in the form of a divergence operator). In the case of a flat bottom ($h_o = 0$) the right-hand side of the equation is 0 and the equation is the strong conservation form of the Shallow Water equations.

The Shallow Water equations have an infinite hierarchy of conservative forms (Ambrosi, 1995) expressing the conservation of mass, energy, discharge rate, velocity, etc. Any two of these equations

Table 1
Summary of tests completed.

Test no.	Name	Description	Figure no.
1a, 1b	Semi-analytical	Comparison of the model runs with semi-analytical solutions	Fig. 3
1c, 1d		M1, M2 curves (mild slope); S2, S3 curves (steep slope)	
2a	Idealised	Flow in a straight idealised channel	Fig. 4
2b		Flow in an embanked straight idealised channel	
2c		Flow in a meandering idealised channel	Fig. 5
3	EA case 1	Wetting and drying of a disconnected body	Fig. 6a & b
4	EA case 2	Low momentum flow	Fig. 7
5	EA case 3	Momentum conservation	Fig. 8
6	EA case 4	Flood propagation over a plain	Figs. 9 & 10
7	EA case 5	Dam break over a valley	Figs. 11 & 12
8a	EA case 6	IMPACT: Hydraulic jump and wake zone (laboratory scale)	Fig. 13
8b		IMPACT: Hydraulic jump and wake zone (realistic scale)	Fig. 14
9	Experimental	Dam break through an urban area	Fig. 15

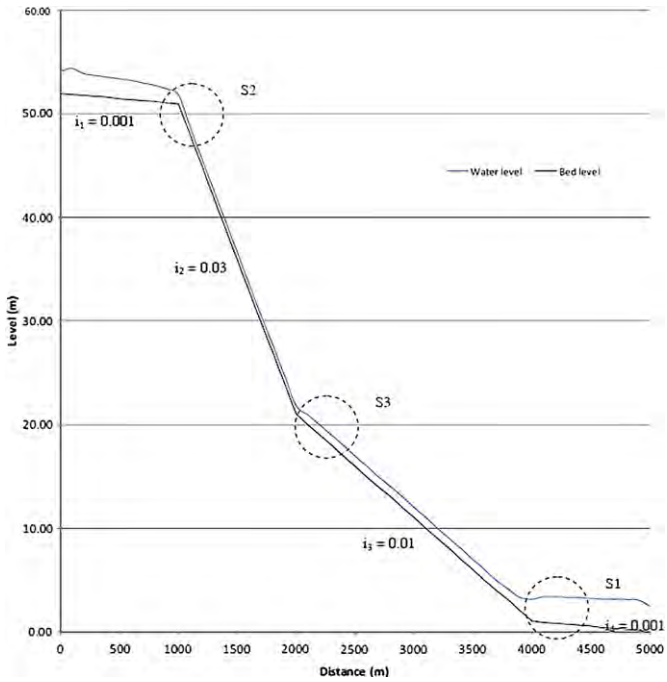

Fig. 2. Connections of the slopes on a river profile for the S1, S2 tests.

are equivalent to one another if the solution belongs to C. This equivalency is no longer valid when shocks (i.e. bores) are involved.

Current codes which solve the Shallow Water equations do so using different numerical methods. From the current literature, several numerical techniques for solving the Saint Venant Equations are available. These include the method of characteristics, explicit difference methods, semi-implicit methods (Casulli, 1990), fully implicit methods, and Godunov methods (Van Leer, 1979). The characteristic method transforms the Shallow Water partial differential equations into a set of ordinary differential equations, which are solved using finite difference methods. The explicit methods transforms the Shallow Water equations into a set of algebraic equations, which can be solved, in sequence, at each point of discretisation, at each time step, while implicit methods solve the equations simultaneously at all computational points at a given time. If, due to boundary conditions and assumptions, the set of Shallow Water equations are non-linear, iteration is needed in order to find the solution (Richtmeyer, 1957).

Numerical stability and convergence issues need to be addressed, while solving the Shallow Water equations numerically. In order to prevent error propagation in explicit methods, the Courant Frederichs Levy (C.F.L.) condition is imposed. This relates the time step to the spatial discretisation and the wave speed, i.e. the time step must be less than or equal to the ratio of the reach length to the minimum dynamic wave celerity $\Delta t \leq \Delta x/c$.

Godunov-type methods can be explicit or implicit. Generally, however, they are explicit in time and, accordingly, the allowed time step is restricted by the C.F.L. stability condition. These methods are in general based on non-staggered grids and can achieve first order accuracy. Godunov-type methods were originally developed for gas dynamics and were then later extended to hydrodynamics on the basis of the analogy between the equations for isentropic flow of a perfect gas with constant specific heat and the Shallow Water Equations (Toro, Leveque).

Semi-implicit methods can be unconditionally stable and still computationally efficient. A semi-implicit method that conserves the fluid volume, applied to channels with arbitrary cross-sections was introduced by Casulli and Zanolli, (1998). However, these methods have to be carefully considered, especially in the case when the physical conservation property of momentum is not satisfied, since incorrect results arise if the methods are applied to discontinuous problems. However, when a semi-implicit scheme using the efficiency of staggered grids is combined with the conservation of both fluid volume and momentum then problems addressing rapidly varying flow can be solved (Stelling and Duinmeijer, 2003).

There are a number of different numerical schemes embedded in different codes, for example: the weighted four point-Preissmann scheme (Preissmann, 1960), Godunov-based methods (LeVeque, 1992), the the weighted six-point Abbott-Ionescu scheme (Abbot and Ionescu, 1967), and TVD (Total Variation Diminishing) schemes (Toro, 1997). Each numerical scheme has its own advantages and disadvantages. Below schemes relevant to this paper are discussed.

The first schemes developed for hydrodynamic computational codes were the fully implicit schemes of Preissmann and Abott–Ionescu during the 1960s. These schemes have developed over time (most significantly in terms of graphical user interfaces GUI), but remain the most popular and widely used in commercially available software. Two examples of codes that use the Preissmann scheme are DAMBRK, which was developed by the US National Weather Service, and ISIS which was developed by Halcrow and HR Wallingford in the UK. An example of a code using Abbott-Ionescu scheme is Mike11, developed at Danish Hydraulic institute.

Table 2
Summary of model meshes, sizes and boundary condition types.

Test Case	Grid size	Grids counts	Roughness C(Chezy) N (manning)	Boundary conditions		Time step	Time span
				Upstream	Downstream		
1a	100 × 100 m	100 × 2 cells	$C = 50$	Constant Discharge	Constant Water Level	5 s	150 min
1b	100 × 100 m	100 × 2 cells	$C = 50$	Constant Discharge	Constant Water Level	5 s	150 min
1c	10 × 10 m	400 × 11 cells	$C = 50$	Constant Discharge	Constant Water Level	5 s	150 min
1d	10 × 10 m	400 × 11 cells	$C = 50$	Constant Discharge	Constant Water Level	5 s	150 min
2a	10 × 5 m	505 × 22 cells	$C = 50$	Varying Discharge	Constant Water Level	30 s	130 min
2b	10 × 5 m	505 × 22 cells	$C = 50$	Varying Discharge	Constant Water Level	30 s	130 min
2c	100 × 100 m	598 × 133 cells	$C = 50$	Varying Discharge	Constant Water Level	5 s	150 min
3	2 × 2 m (10 × 10 for other soft.)	350 × 50 cells	$n = 0.03$	Varying Water Level	N/A (wall)	60 s	20 h
4	20 × 20 m	100 × 100 cells	$n = 0.03$	Varying Discharge	N/A (wall)	60 s	48 h
5	5 × 5 m	60 × 20 cells	$n = 0.05$	Varying Water Level	N/A (wall)	2 s	15 min
6	5 × 5 m	200 × 400 cells	$n = 0.05$	Varying Discharge	N/A (wall)	60 s	5 h
7	50 × 50 m	160 × 340 cells	$n = 0.04$	Varying Discharge	N/A (wall)	60 s	30 h
8a	0.1 × 0.1 m	36 × 990 cells	$n = 0.01$	N/A (initial water level)		1 s	2 min
8b	2 × 2 m	36 × 990 cells	$n = 0.05$	N/A (initial water level)		15 s	30 min
9	0.05 × 0.05 m	72 × 716 cells	$n = 0.01$	N/A (initial water level)		1 s	2 min

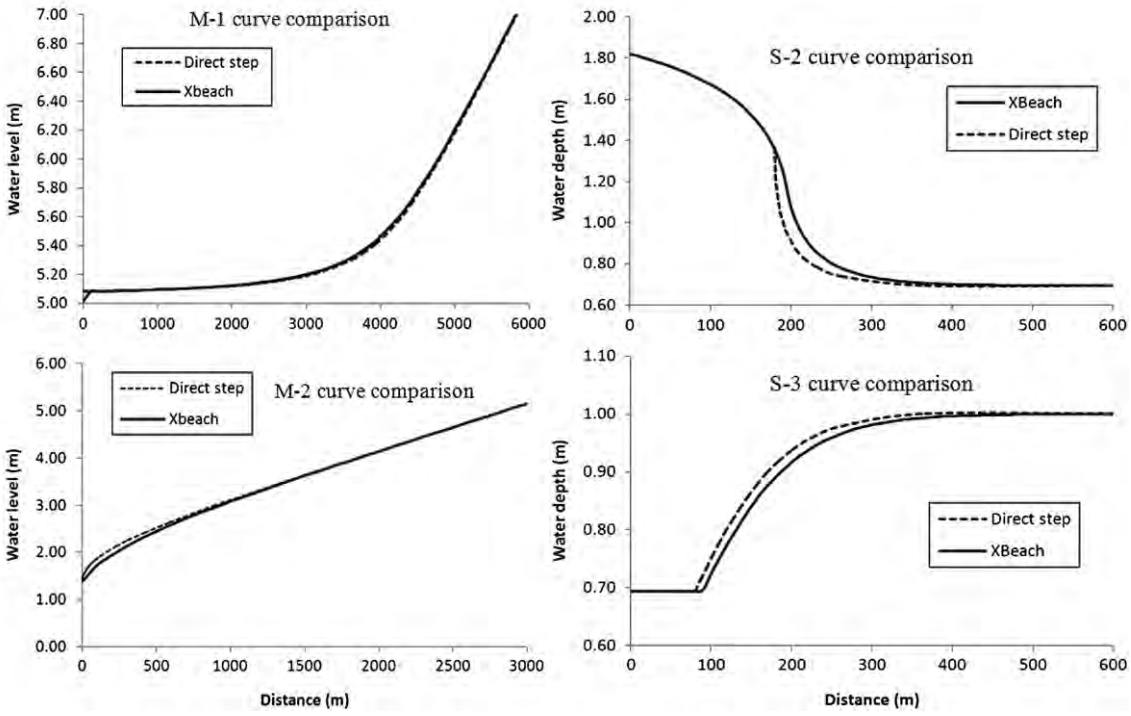

Fig. 3. Modelled and analytical solution of M1, M2, S2, and S3 types of flow (Test 1a, 1b, 1c, 1d).

Godunov developed a method to solve the non-linear systems of the hyperbolic conservation laws describing fluid flow. As a result, the scheme is able to solve the Riemann problem by including various approximate Riemann solvers (Toro, 1997). The Riemann problem is a discontinuity of the conservation law and of piecewise constant data (LeVeque, 1992). Godunov-based schemes with various Riemann solvers are used in river modelling software such as Infoworks RS 2D (Roca and Davison, 2009), TRENT (Villanueva and Wright, 2006) and BreZo (Begnudelli et al., 2008).

The TVD method is able to solve the competing requirement of high order of accuracy and the absence of unphysical oscillations in the vicinity of large gradients (Toro, 1997). A Riemann solver can also be integrated with this method to handle shock capturing. Furthermore, TVD upwind schemes, where the solution in space develops from left to right, are the extension of the Godunov first order upwind method. The TVD scheme has been implemented in the ISIS 2D software (Lin et al., 2006).

Finally, XBeach uses the Stelling and Duinmeijer scheme (Stelling and Duinmeijer, 2003), combining the efficiency of staggered grids with momentum conservation properties needed to ensure accurate results for rapidly varied flows and expansion and/or contractions. This method is very efficient in simulating large scale inundation (Stelling and Duinmeijer, 2003).

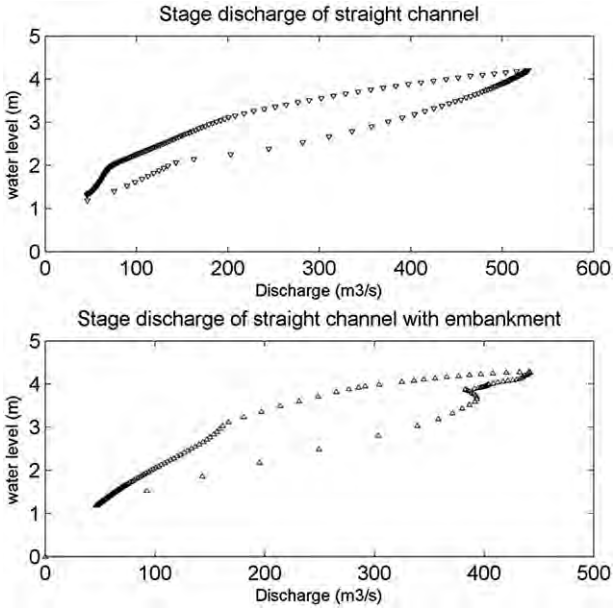

Fig. 4. Test 2a and 2b, Stage discharge in straight channel.

2.1.1. XBeach formulation

XBeach uses a rectilinear, non-equidistant, staggered grid. This discretisation calculates bed level, water level, water depth and concentration of sediment at cell centres while velocities and sediment transport are calculated at the cell border. Velocities at

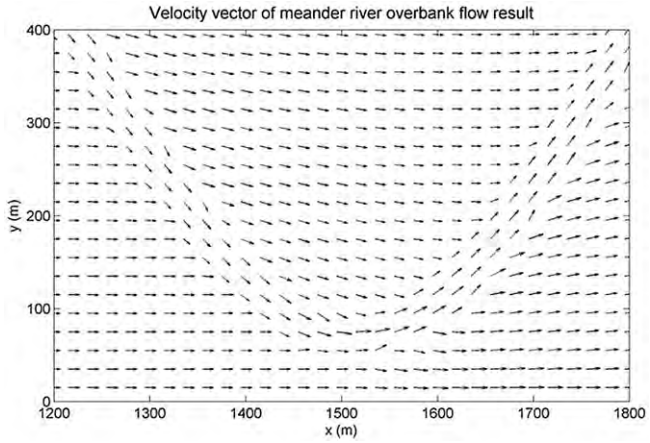

Fig. 5. Test 2c: Flow pattern in meandering channel.

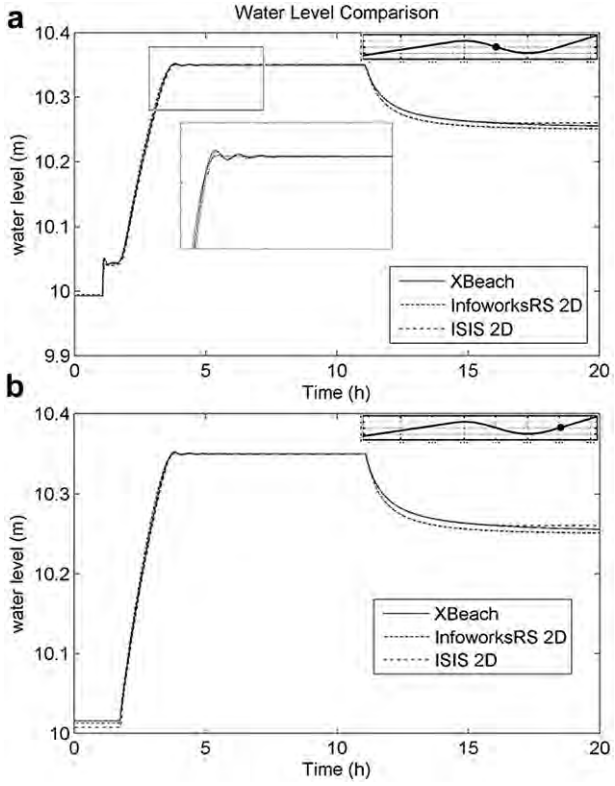

Fig. 6. a. Test 3 – wetting and drying of a disconnected body (results recorded on the downslope of the initial bump).

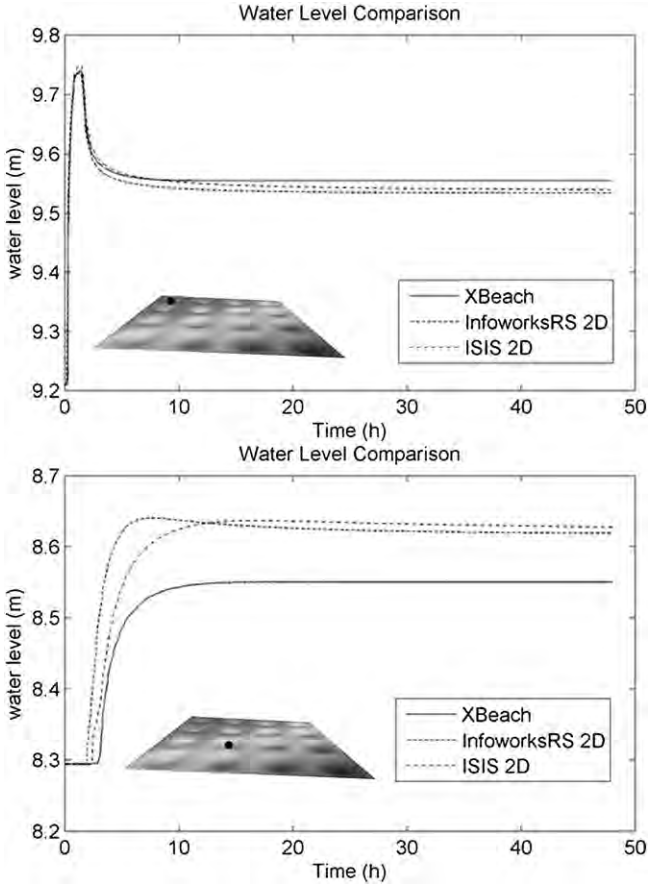

Fig. 7. Test 4 – Low momentum flow.

the cell centres are obtained by interpolating the results from the four surrounding points (Roelvink et al., 2009).

The shallow water equations that are used in XBeach are two dimensional, non-conservative, and are as follows:

Continuity:

$$\frac{\partial \eta}{\partial t}+\frac{\partial hu}{\partial x}+\frac{\partial hv}{\partial y}=0 \tag{3}$$

X momentum:

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}-f\ v-\nu_h\left[\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right]=\frac{\tau_{sx}}{\rho h}-\frac{\tau_{bx}}{\rho h}-g\frac{\partial \eta}{\partial x}+\frac{F_x}{\rho h} \tag{4}$$

Y momentum:

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}-f\ u-\nu_h\left[\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}\right]=+\frac{\tau_{sy}}{\rho h}-\frac{\tau_{by}}{\rho h}-g\frac{\partial \eta}{\partial y}+\frac{F_y}{\rho h} \tag{5}$$

Here τ_{bx}, τ_{by} are the bed shear stresses, η is the water level, F_x, F_y are the wave-induced stresses, ν_t is the horizontal viscosity and f is the Coriolis coefficient (Roelvink et al., 2009).

The other notations in the equations are:

η = water level
t = time
x, y = distance
u,v = water velocity
f = Coriolis coefficient
ρ = water density
g = gravity force per unit mass
νt = horizontal viscosity
h = water depth
τbx, τby = bed shear stresses
F_x, F_y = wave-induced stresses

A first order upwind explicit schematisation with an automatic time step is the preferred numerical method used in XBeach (Roelvink et al., 2008), due to the many shock-like characteristics which occur in hydrodynamic and morphodynamic behaviour (Stelling and Duinmeijer, 2003). The discretisation is similar to the one developed by Stelling and Duinmeijer in its momentum-conserving form, hence it is able to capture shocks and is very suitable for 'drying and flooding', allowing for combinations of sub- and supercritical flows.

The developers of XBeach selected upwind scheme in order to avoid numerical oscillations of many shock-like phenomena, which occur in coastal and flooding situations.

The scheme is able to avoid shock oscillations introduced by the additional dissipative term (Hibberd and Peregrine, 1979). As a result, the upwind scheme, together with a staggered grid, makes the model robust (Roelvink et al., 2009).

In this paper, XBeach is tested against a number of cases; firstly it is compared to the calculation results from semi-analytical solutions. Second, it is tested against the results from different fluvial codes based on various cases (Heriot Watt, 2009). The last comparison is against an experimental case in a laboratory environment (Soarez-Frazao and Zech, 2008). Table 1 provides an overview of the tests undertaken.

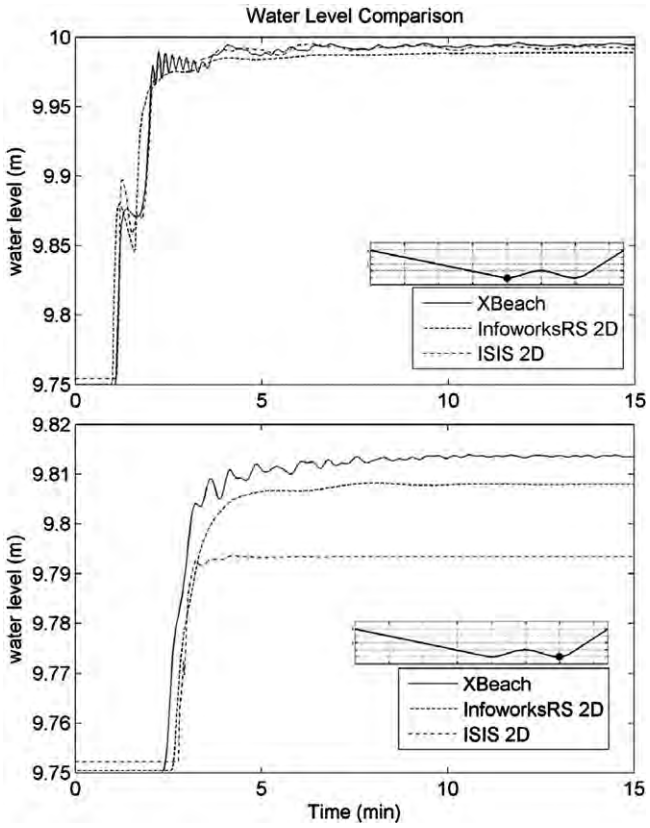

Fig. 8. Test 5 – momentum conservation.

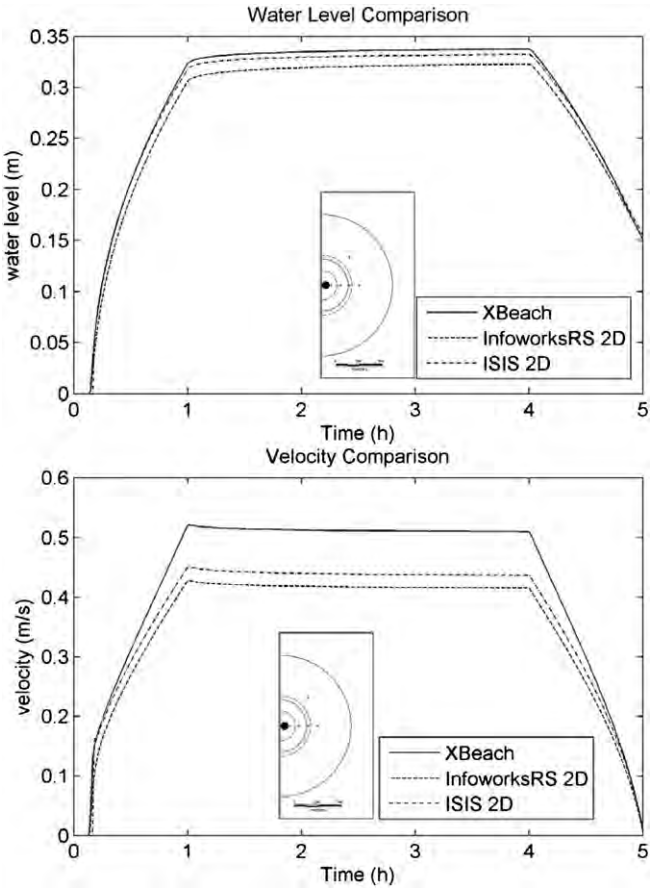

Fig. 9. Test 6 – flood propagation over a plain.

3. Test cases

3.1. Semi-analytical solution comparison

There are certain fluvial hydraulic scenarios which can be solved using semi-analytical methods. These cases provide the starting point for testing the capability of XBeach in the fluvial environment. In addition to the cases that have semi-analytical solutions, there are also cases where known fluvial behaviour can be tested.

The first test examines XBeach's capability to model simple cases such as gradually varied flow and backwater effects, Test 1a-d. The code was tested to model mild slope types (M1 and M2) and steep slope types (S2 and S3) (Chow, 1959; Cunge et al., 1980).

In the mild slope case (1a and 1b), a fragment of very wide river is modelled of a 100 m width with frictionless walls on both sides. A gentle slope of 0.001 is created in the model bathymetry over a 10 km distance. Uniform roughness with a Chézy coefficient of $C = 50$ is applied. For the M1 case (Test 1a), a constant discharge of 2 $m^3/s/m$ is introduced at the upstream boundary and a constant 5 m water level boundary condition in downstream. These conditions at the boundaries generate an M1 flow curve which varies from a water level of 1.17 m–5 m.

The M2 flow curve (Test 1b) was set up using a 5 $m^3/s/m$ discharge at the upstream boundary and a critical depth of 1.37 m at the downstream boundary of the model. These boundary conditions generate a normal water depth of 2.15 m at the upstream boundary.

The cases 1c and 1d are the steep slope cases, here the S2 and S3 flow curves are generated using a 5 km long channel with various bed slopes (i.e. 0.03, 0.01 and 0.001), as shown in Fig. 2. The upstream boundary condition is set as a constant discharge of 5 $m^3/s/m$, and the downstream boundary condition of 2.15 m water level. The S2 flow curve occurs after the transition from the 0.001 to the 0.03 slope and the S3 flow curve at the transition from 0.03 to 0.01 slope.

The second theoretical case (Test 2a) is a straight trapezoidal channel with a flat floodplain on both sides. The model uses a uniform value of Chézy roughness coefficient for both the main channel and floodplain area. The test investigates the two dimensional flow calculations of the software. Furthermore, Test 2b investigates the case of an embanked floodplain to examine the hydraulic representation of a disconnected waterbody.

Case 2a was modelled using a 5 km long, straight channel, with a 0.001 slope and a river cross-section of 30 m width (bottom), 2 m deep and floodplain of 30 m each side. The upstream boundary condition was a varying discharge from 50 to 700 m^3/s with the peak discharge occurring after 43.3 min and a minimum value after 60 min. The dimensions of the channel were set so as to allow an overflow at the peak flow. In Case 2b dikes are constructed on both banks of the channel.

More complex, and realistic, hydraulic behaviours throughout the domain are found in meandering channels. A perfect sinusoidal meander with no slope was modelled (Test 2c). This test investigates the water flows from the main channel to floodplain and vice versa, secondary flow in the curved channel and velocity distributions as well as flood wave behaviour.

For Case 2c the modelled reach uses a rectangular channel 50 m wide and 5 m depth, with a length of 4.50 km. The actual model domain was only 3 km as the meanders were introduced to create the extra length in the main channel. The total width of the floodplain was 600 m on both river banks. A zero bed slope was applied in this case, in order to model flow in the channel only as

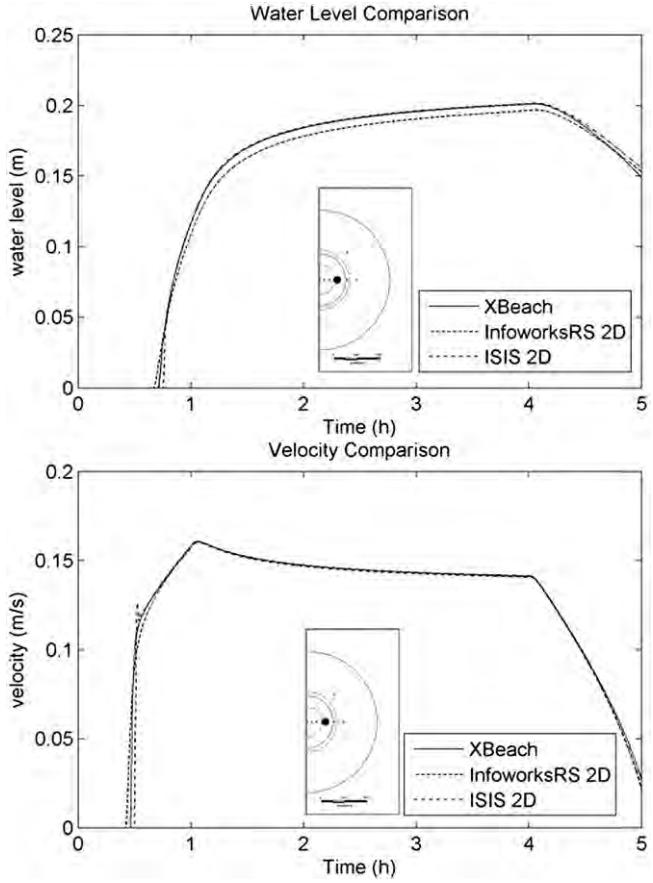

Fig. 10. Test 6 – flood propagation over a plain (cont).

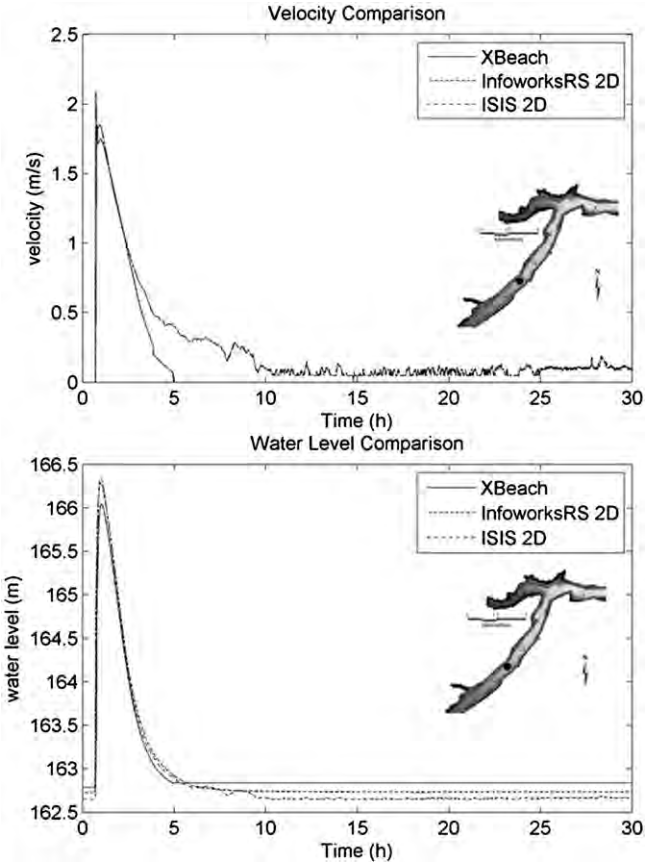

Fig. 11. Test 7 – dam break over a valley (head of the valley).

a result of the upstream boundary condition. This was set as a varying discharge (from 200 m^3/s to 1000 m^3/s),

3.2. The Environment Agency 2D benchmarking study (EA2D)

In 2009, the Environment Agency of England and Wales carried out a benchmarking study for 2D software. The benchmarking was undertaken to ensure that codes used for fluvial studies commissioned by the Agency were appropriate for use in assessing flood risk. The project was led by Heriot Watt University (Heriot Watt, 2009) and the simulations were set up and carried out by various model developers throughout the world. The final report of this study is available to the public (Neelz and Pender, 2010). Six out of eight EA2D tests were chosen to test the ability of XBeach in fluvial modelling situations. For this study comparisons were made to InfoWorks RS 2D and ISIS 2D.

Test 3 (the first test of the EA2D project) investigates the code's capability to handle wetting and drying of a disconnected waterbody. A water level fluctuation is introduced in the model with two low points disconnected by a bump. The modelled domain is 100 m × 700 m with a large bump in the middle of the bathymetry of 10.25 m elevation, in order to disconnect the water bodies once the water level is lower than the elevation of the bump. The boundary conditions imposed were varying water levels at the left side of the model (between 9.70 m and 10.35 m at time $t = 1$ h–11 h and than decreasing water levels back to the value of 9.7 m). The water level and velocity at two locations are compared.

Test 4 determines inundation extent with low momentum flows in a complex topography. Furthermore, it also examines the disconnected waterbody, wetting and drying of a floodplain, inundation extent and looks at final depth rather than maximum depth. The model domain is a sloped plain in two directions with 16 depressions included in the terrain to retain a portion of the water that flows in from upper corner of the domain. The model covers an area of 2000 m × 2000 m with 16 depression each of 0.5 depth. There is an overall slope of 1:1500 in the north direction and 1:3000 towards the east, resulting in a ~2 m drop of elevation between top left corners to bottom right corner.

The inflow boundary condition was located at the top left side of the domain over a length of 100 m, with a discharge value of 20 m^3/s for a period of 75 min starting at time $t = 10$ min. All other boundaries of the domain are closed boundaries.

The fifth test simulates momentum flow over a barrier. This capability is important in sewer or pluvial flood modelling in urban floodplain areas. The domain consists of a steep slope to accelerate the inflow and a bump to disconnect it from another depression. The boundary condition is a discharge of 65.5 m^3/s for 10 s starting at time $t = 5$ s with a peak at time $t = 15$. The model is 300 m long with a bump of 25 cm height. The domain consists of a steep slope to accelerate the inflow and a bump to disconnect it from another depression. The volume of the inflow is just enough to fill the depression. Water is expected to overtop the bump due to the force of momentum and settle in the depression behind the bump. This test differentiates codes which incorporate the full momentum terms and those that do not.

Case 6 tests the simulation of flood propagation over a wide floodplain following a dike failure. A high burst inflow is applied at the breach point, and a wide flat floodplain is modelled to test the propagation of a flood wave and velocities at the leading edge of the flood wave. The modelled area is a 1000 m × 2000 m of flat

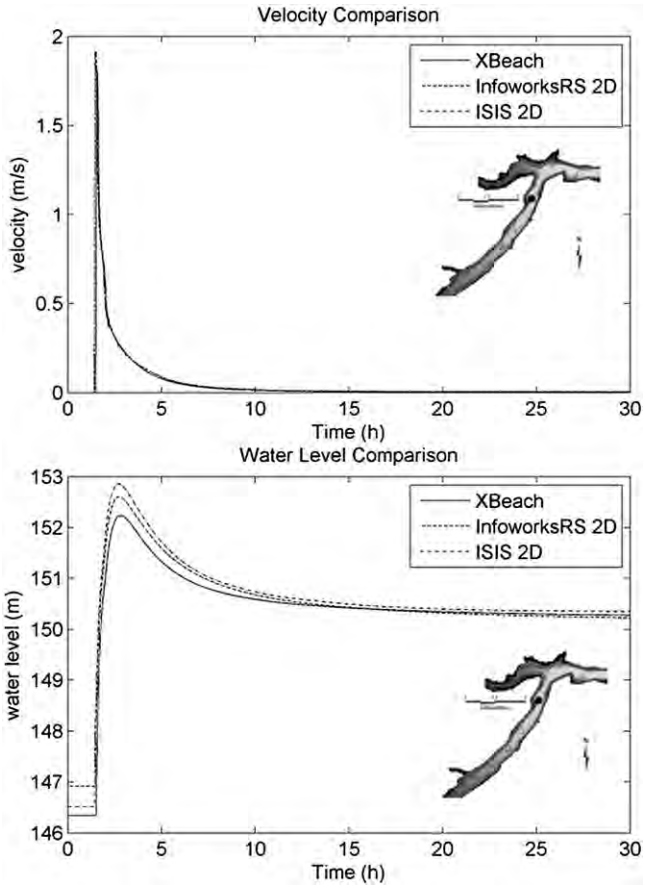

Fig. 12. Test 7 – dam break over a valley.

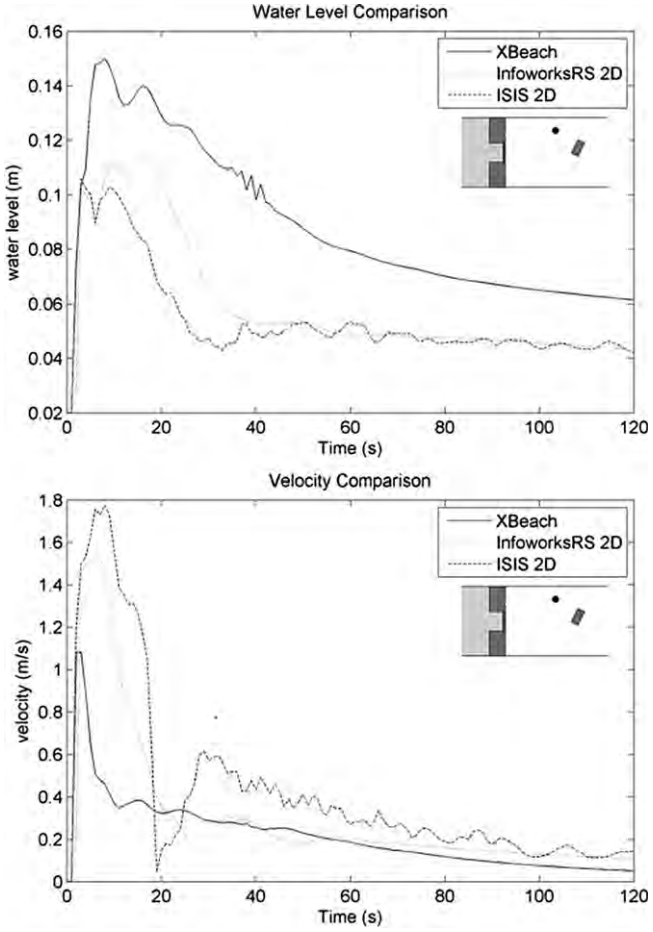

Fig. 13. Test 8a – dam break over a building, laboratory scale.

topography. The inflow boundary condition reaches a peak discharge of 20 m^3/s at time $t = 60$ min and continues constantly with this value for a further 180 min. The inflow is located in the centre of the left boundary of the model. The objective of the test is to examine the capability of XBeach to simulate the speed of flood wave propagation and predict transient velocities and depths. The test is applicable for fluvial and coastal flooding caused by a dike breach.

The penultimate EA2D test (Test 7) models the simulation of a flood wave propagation following a dam failure that flows through a river valley. The case tests the software's capability in simulating major flood inundation and flood hazard prediction that arises from a dam break scenario. The software was expected to be able to model a high burst discharge over steep and mild bed slopes involving both subcritical and supercritical flow. The test has a skewed discharge boundary with peak flow of 3000 m^3/s at time $t = 10$ min for 10 min and slowly decreases thereafter for a further 80 min.

Test 8 is adapted from a benchmark test case from the IMPACT project (IMPACT, 2005; Soares-Frazao and Zech, 2002). This test examines the capability of the software to simulate hydraulic jumps and the wake zone behind a building. The test consists of two cases, the laboratory scale (1:20) (Test 8a) and the realistic scale (Test 8b). The scale of the scaled model is 1:20. The dam size is 3.6 m × 99 m with a breach of 1 m wide in the middle of the dam (6.75 m from the left side of the dam). The initial water level in the reservoir behind the dam is 0.20 m, while the water level in the floodplain area is set to a value of 0.02 m (a wet bed domain). A model of a building is set in the floodplain in line with the dam breach location. Test 8b is a dam break case at real scale. The size of the computer model is obtained by multiplication with 20. Therefore the initial water level at the dam is 8.00 m and in the floodplain area is 0.4 m.

3.3. Experimental case comparison

The experimental test is based on the paper by Sandra Soares-Frazao and Zech (2008). The physical model represents an urban area with 25 buildings blocks flooded by a dam break simulation of a reservoir (Soarez-Frazao and Zech, 2008). The main difference between this test (Test 9) and the last test of EA2D project is the complexity of the obstacle. In this case there are many buildings and simulated streets. The capability to model hydraulic jump and complex flow through urban areas is therefore investigated.

Details on the models meshes, sizes and boundary condition types are given in Table 2. The Courant number used is not included in the table, but remains the same for all tests. The number used is 0.9.

4. Results and discussion

For the backwater cases comparing the modelled results with the semi-analytical results shows deviations (Fig. 3). In the M1 and M2 cases, a difference is observed at the boundary while for the S2 and S3 cases, the deviation occurs along the profiles, from the point of disturbance until the solution reaches normal depth. These anomalies occur at the transition from mild to a steep slope (M2, S2), while smaller differences are observed at the transition from

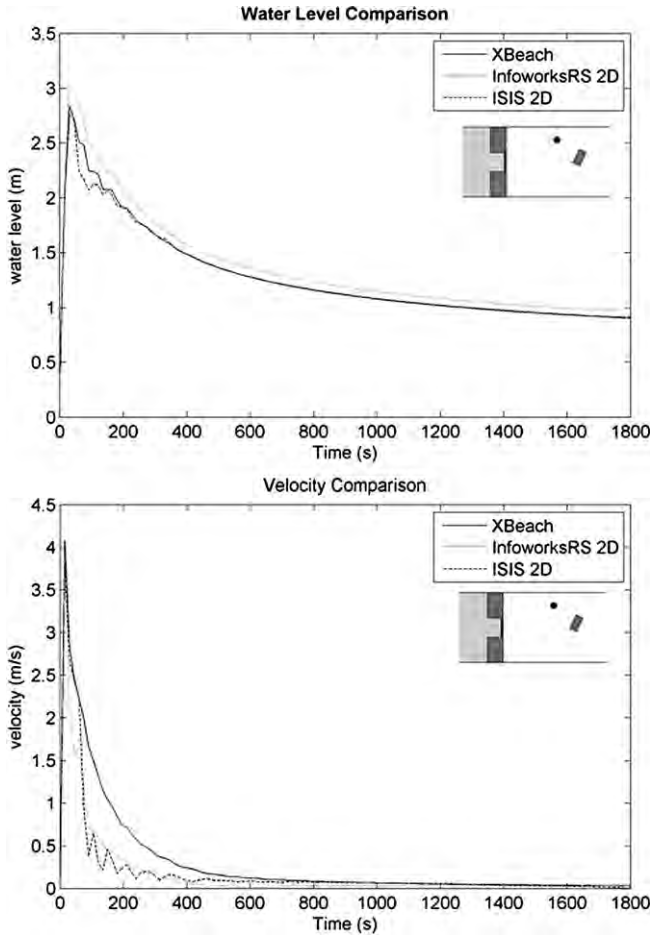

Fig. 14. Test 8b – dam break over a building, realistic scale.

steep to mild slope (M1, S3). This is due to the transition from subcritical to supercritical flow and shows the ability of the code to capture shocks. Boundary conditions also show inconsistencies. The inconsistency in results observed at the boundary is due to the implementation of a flow boundary condition, which is currently represented using velocity vectors. These are implemented at the centre of each cell situated in the boundary.

For Test 2a and 2b, the straight trapezoidal channel, good results are obtained in the case of overland flow prediction and Fig. 4 indicates that the hydraulic behaviour of flow on the floodplain is reasonably modelled. The hysteresis effects can clearly be seen in both figures (a and b), indicating the behaviour of the fluid as it flows out of bank. In the case of the embanked river, a distinctive transition can be observed in the hysteresis as the water overtops the levees and accommodates the available volume below. For Test 2c, Fig. 5 shows the patterns of flow in a meandering channel. Known behaviours such as water flows in to the main channel, and flow distribution characteristics in a channel with floodplain are observed (Muto et al., 1999). The flood wave progressing down the main channel ahead of the flow distribution on the flood plains can be observed. Tests 2a, 2b and 2c perform well and demonstrates that the celerity of propagation of a fluvial flood wave is represented well in the numerical scheme in a variety of scenarios.

For Test 3, the incoming water is expected to fill the depressions in the domain. Fig. 6a and b shows that XBeach compares well with ISIS 2D and Infoworks RS 2D for this test. A small instability can be observed following maximum depth, on the downslope of the first bump (Fig. 6a). Each computational code compared gives marginally different results. Additionally, differences are observed at initial and final water depths, which are of the order of a few millimetres. In general however, the results of Test 3 show a good comparison to other fluvial computational codes.

On the other hand, Fig. 7 shows the results of Test 4, where significant differences between codes can be observed. Due to the wet/dry threshold value that was set in XBeach, a greater number of dry depressions are observed than anticipated. The closer to the inflow location that the result is sampled the more favourable the XBeach comparison with other codes. However XBeach performs poorly in this test. The reason for this behaviour is that the threshold value of the wetting and drying algorithm is high. This means the domain retains 2 cm of water when it should actually be dry. This is due to the mathematical formulation of the problem whereby for Manning's equation the depth (d) calculation is completed with the Manning coefficient located in the denominator and hence the expression generates errors. The problem could be avoided if the Chézy equation is used instead of Manning. The use of the Manning equation was due to the test requirements where the Manning model is the preferred roughness representation in the fluvial environment for the Environment Agency. This issue can be neglected when modelling real rivers, since the 2 cm threshold does not impact significantly at the larger scale.

Fig. 8 shows the results of Test 5, momentum conservation, and indicates that there are minor differences between the codes. Firstly an instability occurred in the XBeach solution, close to maximum depth. Secondly, different final water levels for each code were observed, although all codes calculate water levels close to the expected 10 m mark. These differences are considered to be marginal for this case, and all codes demonstrate that momentum conservation is captured in the numerical formulation. The results of Test 6 are shown in Figs. 9 and 10. The results of the three codes give different values of velocity and water levels at the tested nodes. All show the general behaviour of half circle flood extents. From the results in Fig. 10 different values recorded at test nodes are observed. However, the arrival times for all nodes are approximately the same for all codes. After the arrival of the peak at the nodes differences in the predictions are observed. In general flood propagation over a plain and the momentum conservation tests (Tests 5 and 6) report good results. XBeach is able to model the

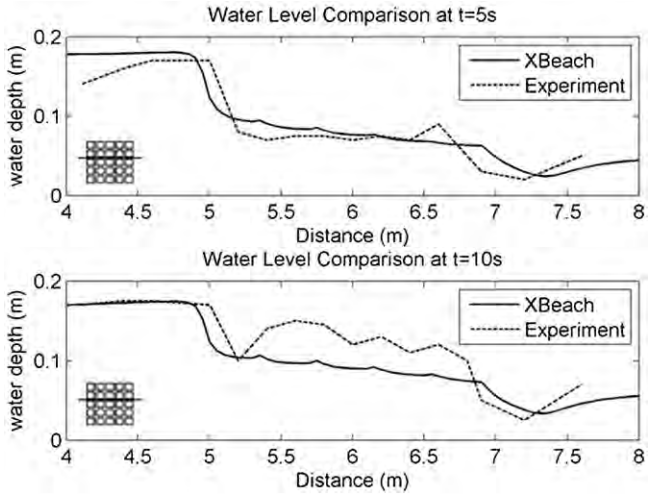

Fig. 15. Test 9 – Laboratory experiments.

required scenarios, and reproduce expected flow behaviour in a comparable manner to the other two river codes.

For the realistic scale dam break test (Test 7), XBeach gives a higher final water level result than the other tested codes and a lower value for flow velocities for nodes located at the head of the valley (Fig. 11). This trend is less evident further down the valley (Fig. 12). This is explained by the fact that Xeach is using a first order scheme, which is dissipative. However, for Test 8a (the EA2D laboratory small scale model), where the code's ability to reproduce hydraulic jumps at a laboratory scale is investigated, noticeable differences between the three codes for computed water depth and velocities can be seen (Fig. 13). When this is translated into the realistic scale (Test 8b) however, a similar result is observed for all three codes (Fig. 14).

The result from the dam break case through a valley (Test 7), and dam break over a building (Test 8b) gives comparable results to InfoWorks RS 2D and ISIS 2D. However, the small scale results show poor agreement (Test 8a), which is to be expected. Nevertheless, if Froude scaling is applied between the small scale measurement and simulation real scale computations for XBeach, a good match is found, which cannot be said for the other codes.

Finally, for Test 9, the comparison between XBeach and measured laboratory results of a dam break experiment over an urban area shows large differences between the two, especially at the street level. Improved results are likely if the computational mesh were to be refined significantly, however computational time would increase. This modification on the scale required however would be difficult. The structured rectangular grid that is used in XBeach can lead to a large number of cells, if it is applied to a detailed complex system at real scale such as a meandering channel, bifurcation or an urban area, however it gives a quick solution when modelling low complexity fluvial system. Highly complex river systems, which are solved only by using large domains, detailed meshes and long computational times, can be addressed by running such cases in a parallel manner, on multiple processor computers or by making use of grid computing.

5. Conclusions

The Shallow Water equation solver in XBeach has been seen to work well for river modelling scenarios, when compared to other codes developed for the fluvial environments. Furthermore, with an open source licence, any user may improve the software or add flexibility. However, some deficiencies are acknowledged. The existing representation of the boundary conditions, while adequate for coastal environments, was not always applicable for fluvial modelling purposes, especially in the case of upstream flow conditions. A small sub-routine was implemented, however further refinement is warranted.

There are several cases for which the software can be further tested, such as spillways, chute blocks, etc, which were not tested in this study. The main objective of this benchmarking study was to see the capability of the software to represent floods in rivers using this coastal software. The assumption under which the Shallow Water equations are solved using Xbeach restricts application of these equations for flow problems with a steep bed slope, particularly for the supercritical cases. Consequently testing spillways with this code would imply changes were implemented so that supercritical cases can be tested as well. Pipe flow was not tested in this study.

Codes treating coastal problems address the wetting and drying of computational cells differently from 2D river modelling which result in different inundation patterns. This can be overcome by imposing a different threshold value than that used in coastal applications of XBeach. Similarly in the coastal environment Chézy is the roughness representation of choice. Transferring to the fluvial environment for this study requires the implementation of a sub-routine to change the roughness coefficient, in this case to Manning's. Further modification should be implemented.

XBeach uses a structured staggered grid which can be inflexible for representing complex fluvial geometries. As a recommendation for further development of XBeach, it is suggested that an unstructured grid option is investigated so as to avoid very fine grids.

Finally, the conclusion of this research opens up the possibility to use this model for both hydraulic and potentially morphological problems in fluvial, coastal and hence transition areas.

Acknowledgements

The authors would like to acknowledge the Environment Agency of England and Wales and Herriot Watt University for providing the EA2D test cases, and Wallingford Software (now MWH Soft) and Halcrow for providing their comparison data. Additionally the first author is grateful to the Dutch Development Co-operation Programme and to the Dutch StuNed scholarship programme for funding his MSc studies.

References

Abbot, M.B., Ionescu, F., 1967. On the numerical computation of nearly-horizontal flows. J. Hydraul. Res. 5 (2), 97–117.
Ambrosi, D., 1995. Approximation of the shallow water equations by Roe's Riemann solver. Int. J. Num. Meth. Fluids 20, 157–168.
Begnudelli, L., Sanders, B.F., Bradford, S.F., 2008. Adaptive Godunov-based model for flood simulation. J. Hydraul. Eng. 134 (6), 714–725.
Bitzer, J., 2004. Commercial versus open source software: the role of product heterogeneity in competition. Econ. Syst. 28 (4), 369–381.
Casulli, V., 1990. Semi-implicit finite difference methods for the two-dimensional shallow water equations. J. Comp. Phys. 86, 56–74.
Casulli, V., Zanolli, P., 1998. A conservative semi-implicit scheme for open channel flows. Int. J. Appl. Sci. Comp. 5, 1–10.
Chow, V.T., 1959. Open-Channel Hydraulics. MCGraw-Hill, New York.
Cunge, J., Holly, F., Verwey, A., 1980. Practical Aspect of Computational River Hydraulics. Pitman Publishing, Massachusetts.
de Vriend, H., 1991. Mathematical modelling and large-scale coastal behaviour. J. Hydraul. Res. 29 (6), 727–740.
Ferziger, J., Peric, M., 1999. Computational Methods for Fluid Dynamics. Springer, Berlin.
Henley, M., Kemp, R., 2008. Open source software: an introduction. Comput. Law Security Rep. 24 (1), 77–85.
Heriot Watt, 2009. Specification for 2D Model Benchmarking. Environment Agency / Heriot-Watt University, Edinburgh.
Hibberd, S., Peregrine, D.H., 1979. Surf and run-up on a beach: a uniform bore. J. Fluid Mech. Digit. Archive 95 (02), 323–345.
IMPACT, 2005. Investigation of Extreme Flood Processes and Uncertainty Final Technical Report.
Lanzi, D., 2009. Competition and open source with perfect software compatibility. Inform. Econ. Pol. 21 (3), 192–200.
LeVeque, R., 1992. Numerical Methods for Conservation Laws. Birkhäuser, Bassel.
Lin, B., Wicks, J., Falconer, R., Adams, K., 2006. Integrating 1D and 2D hydrodynamic models for flood simulation. Water Manage. 159 (1), 19–25.
Muto, Y., Ishigaki, T., Rodi, W., Laurence, D., 1999. . Secondary flow in compound sinuous/meandering channels. In: Rodi, W., Laurence, D. (Eds.), Engineering Turbulence Modelling and Experiments 4. Elsevier Science Ltd, Oxford, pp. 511–520.
Neelz, S., Pender, G., 2010. Benchmarking of 2D Hydraulic Modelling Packages. SC080035/SR2. available online. Environment Agency, ISBN 978-1-84911-190-4. http://publications.environment-agency.gov.uk/pdf/SCHO0510BSNO-e-e.pdf accessed 25.08.10.
Preissmann, A., 1960. Propagation des intumescenes dans les canaux et rivières. in: 1er congress de l'Assoc. Francaise de Calcul, Grenoble, France, pp. 443–442.
Richtmeyer, R.D., 1957. Different Methods for Initial Value Problems. Interscience Publ., New York.
Roca, M., Davison, M., 2009. Two dimensional model analysis of flash-flood processes: application to the Boscastle event. J. Flood Risk Manage. 3 (1), 63–71.
Roelvink, D., Reniers, A., van Dongeren, A., van Thiel de Vries, J., McCall, R., Lescinski, J., 2008. XBeach Model Description and Manual. UNESCO-IHE

Institute for Water Education, Deltares and Delft University of Technology, Delft.
Roelvink, D., Reniers, A., van Dongeren, A., van Thiel de Vries, J., McCall, R., Lescinski, J., 2009. Modelling storm impacts on beaches, dunes and barrier islands. Coastal Eng. 66 (11–12), 1133–1152.
Soares-Frazao, S., Zech, Y., 2002. Dam-Break Flow Experiment: The Isolated Building Test Case IMPACT project technical report.
Soarez-Frazao, S., Zech, Y., 2008. Dam-break flow through an idealized city. J. Hydraul. Eng. 46 (5), 648–658.
Stelling, G.S., Duinmeijer, S.P.A., 2003. A staggered conservative scheme for every Froude number in rapidly varied shallow water flows. Int. J. Num. Meth. Fluids 43 (12), 1329–1354.
Toro, E., 1997. Riemann Solvers and Numerical Methods for Fluid Dynamics – A Practical Introduction. Springer-Verlag, Berlin.
Van Leer, B., 1979. Towards the ultimate conservative difference scheme V. A second order sequel to Godunov's method. J. Comp. Phys. 32, 101.
Villanueva, I., Wright, N.G., 2006. Linking Riemann and storage cell models for flood prediction. ICE J. Water Manage. 159 (1), 27–33.