Check for updates

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

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

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

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

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

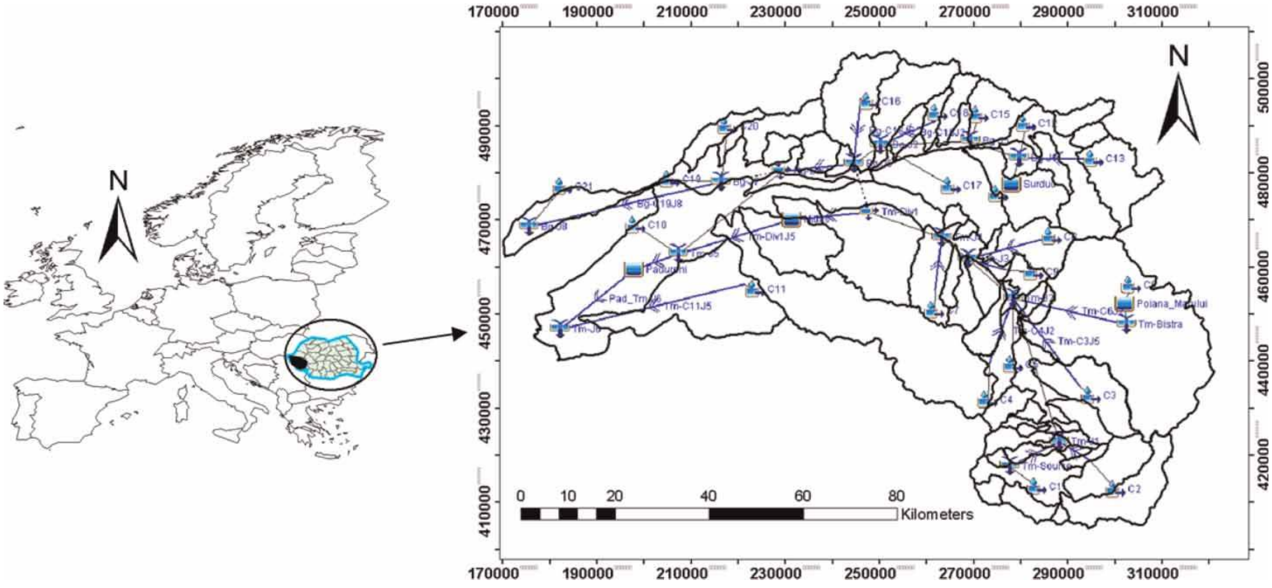

Figure 1 | Timis–Bega catchment location and HEC-HMS model.

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'

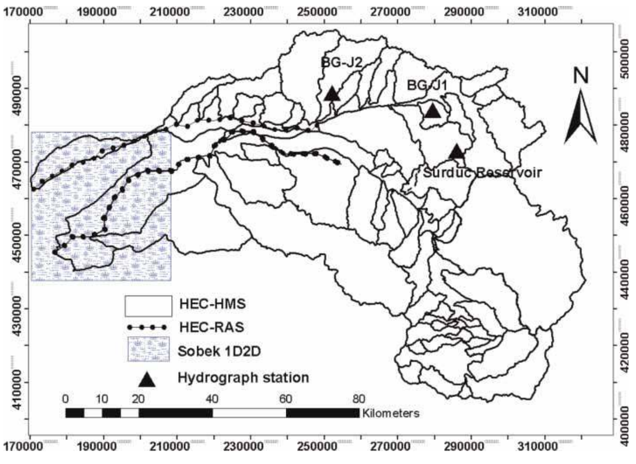

Figure 2 | Geographical coverage by hydrologic/hydraulic models.

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

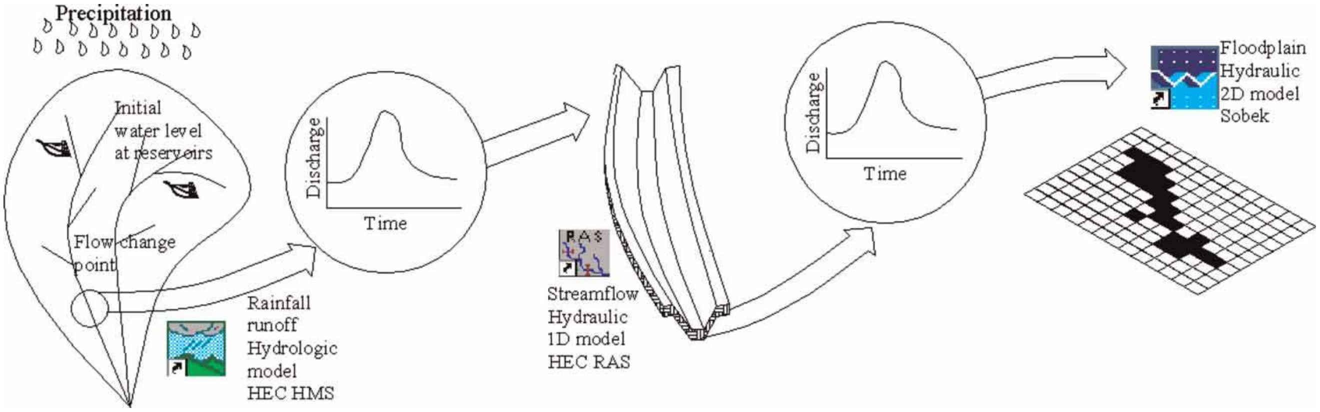
Precipitation
Initial water level at reservoirs
Flow change point
Rainfall runoff Hydrologic model HEC HMS
Discharge
Time
Streamflow Hydraulic 1D model HEC RAS
Discharge
Time
Floodplain Hydraulic 2D model Sobek

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

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

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.

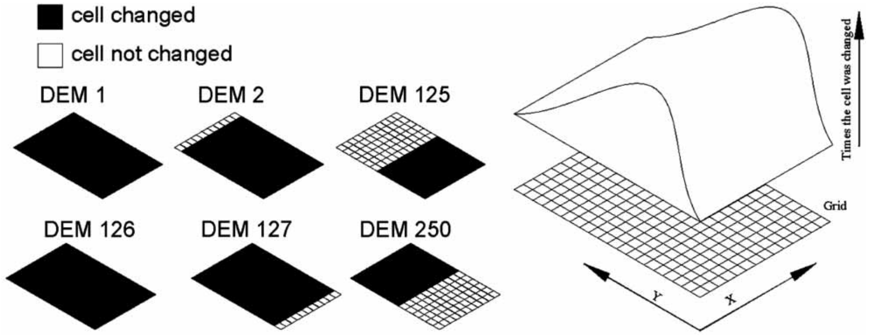

Figure 4 | Sampling DEM cells.

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

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 Amazon 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

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

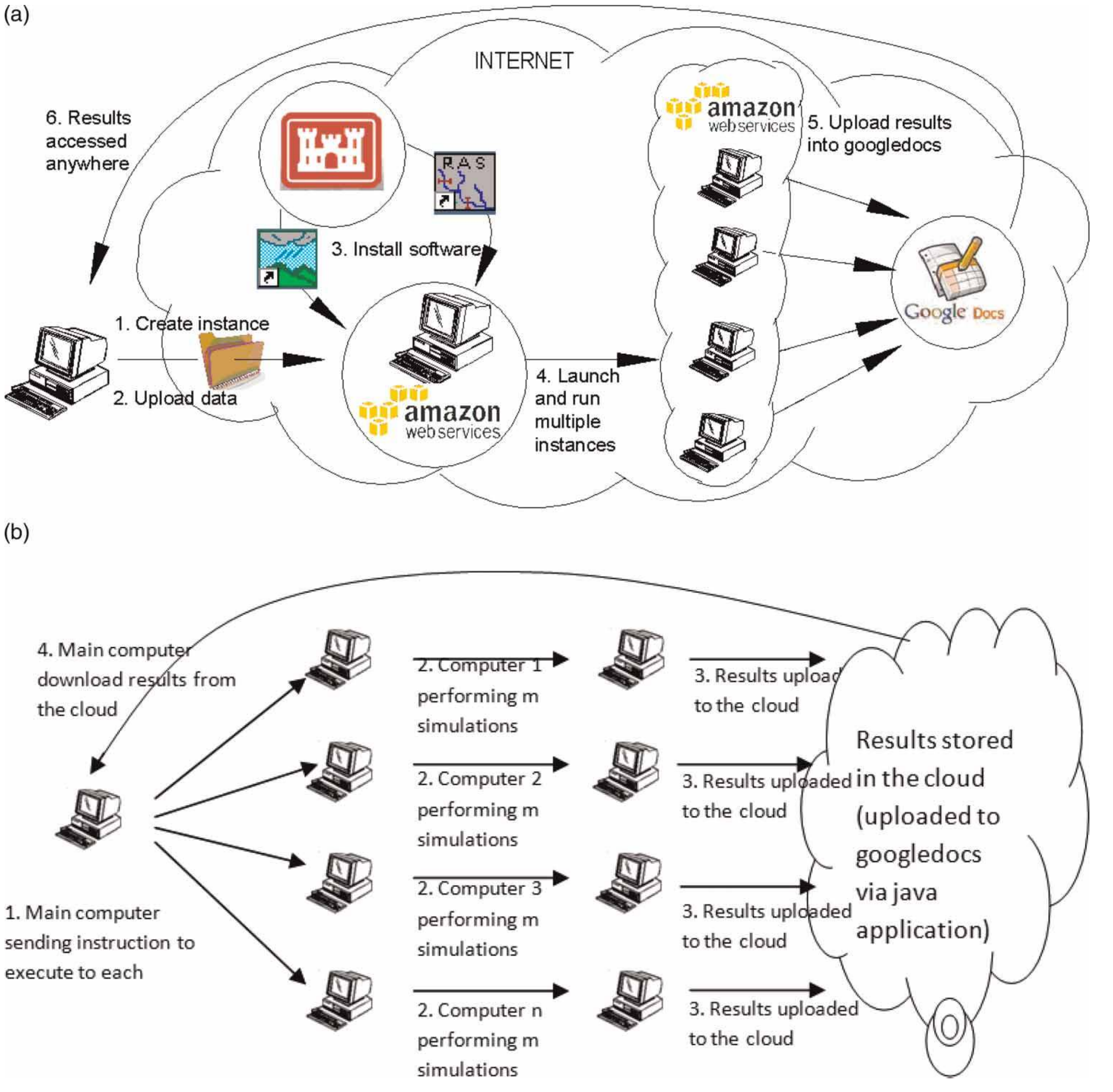
(a)
INTERNET
6. Results accessed anywhere
amazon webservices
5. Upload results into googledocs
RAS
3. Install software
1. Create instance
2. Upload data
amazon webservices
4. Launch and run multiple instances
Google Docs
(b)
4. Main computer download results from the cloud
2. Computer 1 performing m simulations
3. Results upload to the cloud
2. Computer 2 performing m simulations
3. Results uploaded to the cloud
Results stored in the cloud (uploaded to googledocs via java application)
2. Computer 3 performing m simulations
3. Results uploaded to the cloud
1. Main computer sending instruction to execute to each
2. Computer n performing m simulations
3. Results uploaded to the cloud

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

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

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.

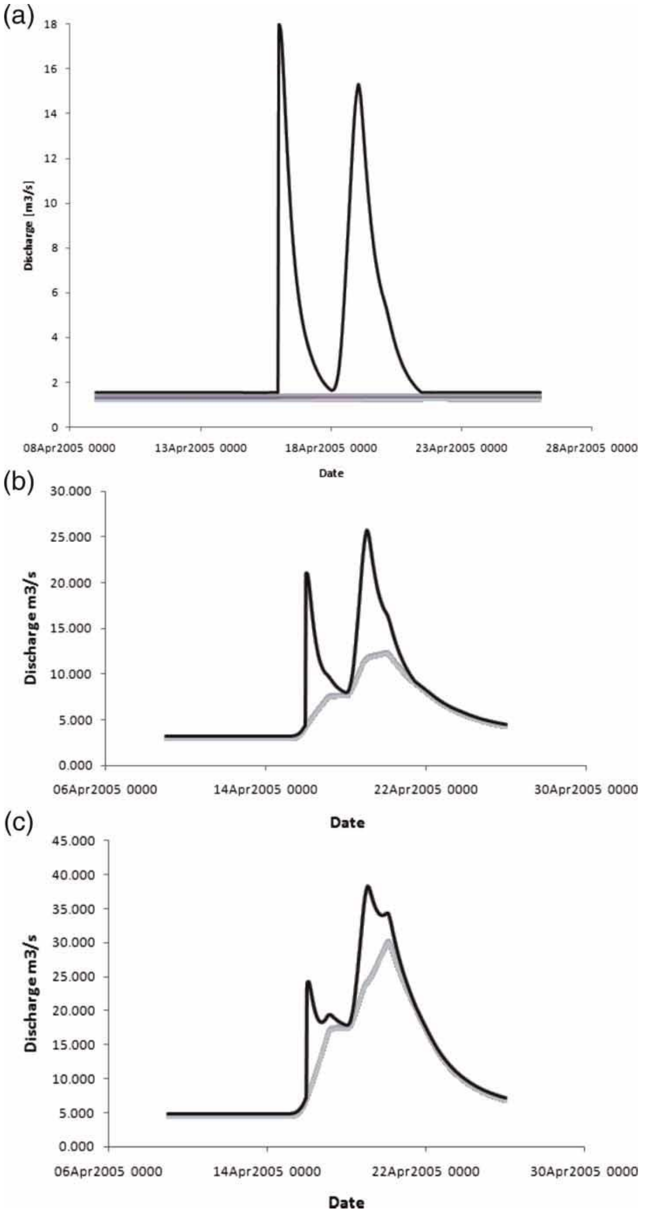

Figure 6 | (a) Hydrographs at the outlet from the Surduc reservoir. (b) Hydrographs at station BG-J1, 10 km downstream of the Surduc reservoir. (c) Hydrographs at station BG-J2, 30 km downstream of station BG-J1.

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

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

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,

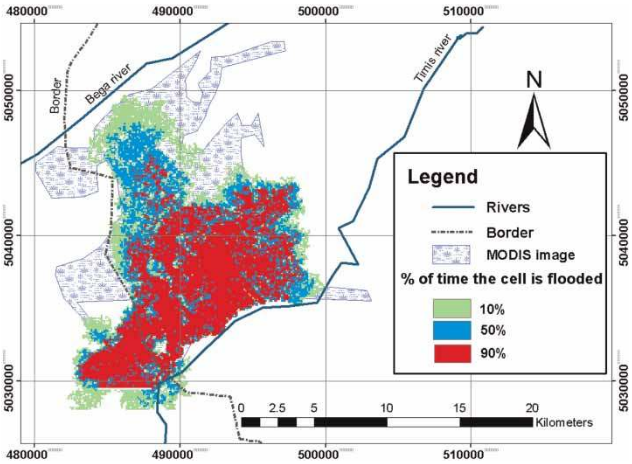

Figure 7 | MODIS flood extent and simulated flood extent.

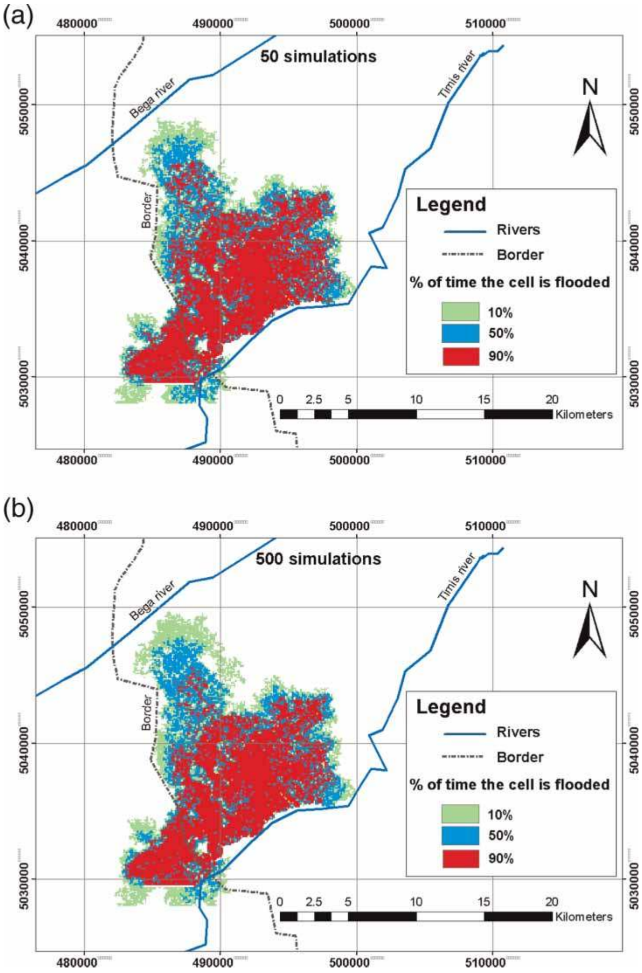

Figure 8 | (a) CPF map for 50 simulations. (b) CPF map for 500 simulations.

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

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.

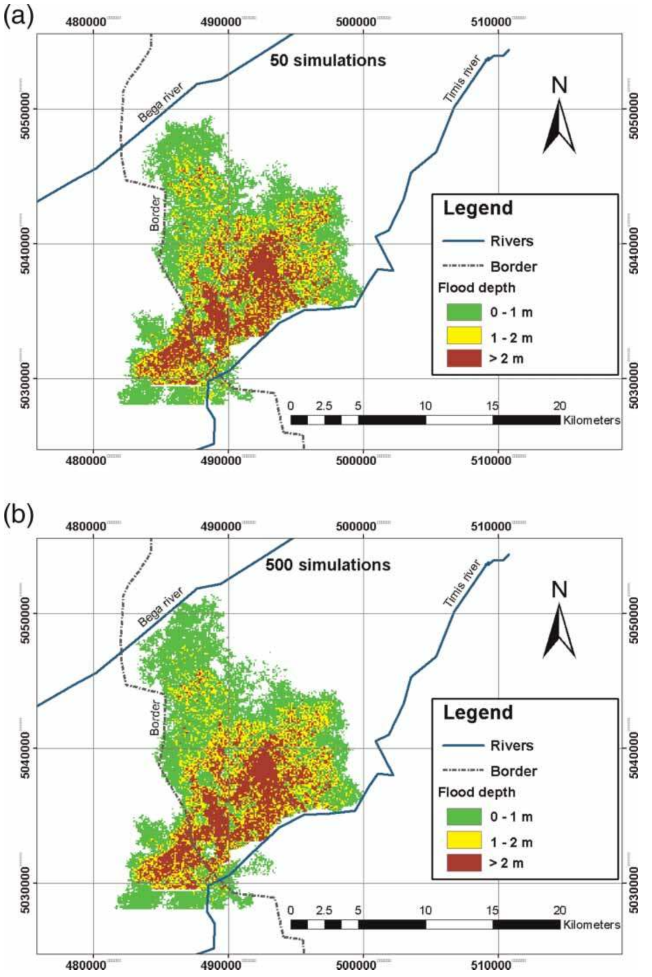

Figure 9 | (a) Mean flood depth for 50 simulations. (b) Mean flood depth for 500 simulations.

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

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

Table 1 | Mean flood depth standard deviation

No. of simulations	$\sigma = 0–1$ m	$\sigma = 1–2$ m	$\sigma =$ 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 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,

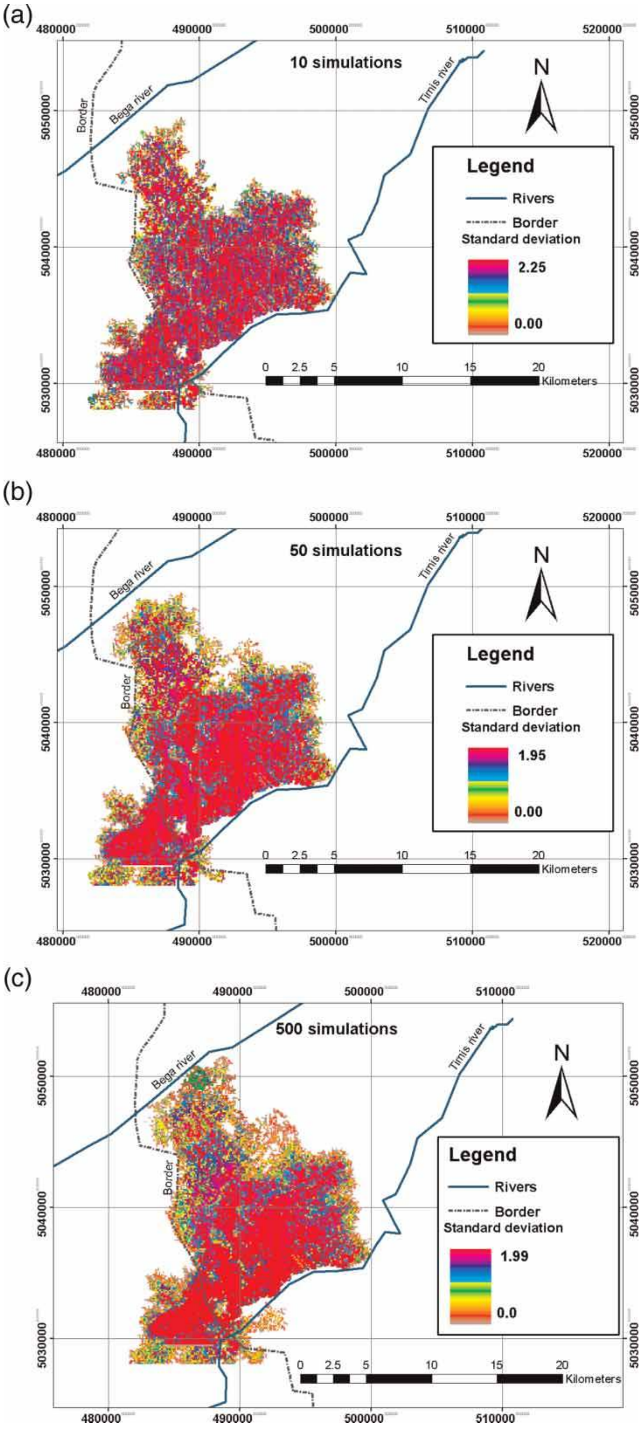

Figure 10 | (a) Standard deviation of depth for 10 simulations. (b) Standard deviation of depth for 50 simulations. (c) Standard deviation of depth for 500 simulations.

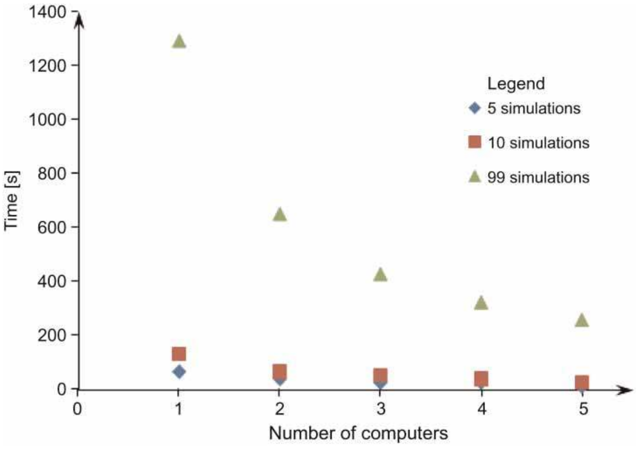

Figure 11 | Time saved in HEC-HMS model using distributed computing.

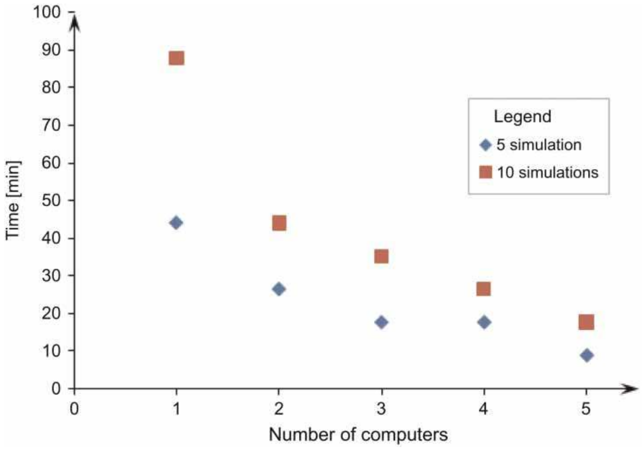

Figure 12 | Time saved in HEC-RAS model using distributed computing.

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

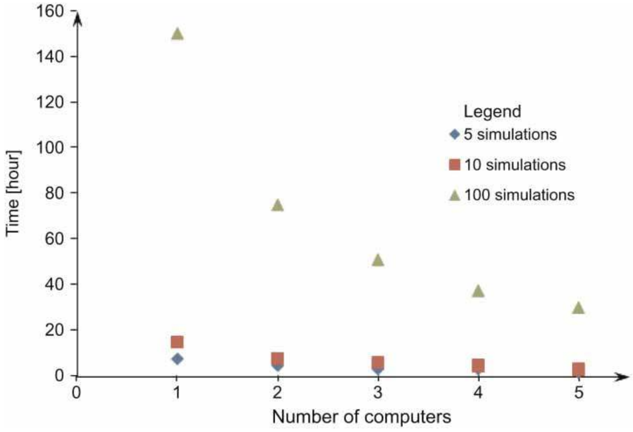

Figure 13 | Time saved in Sobek model using distributed computing.

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

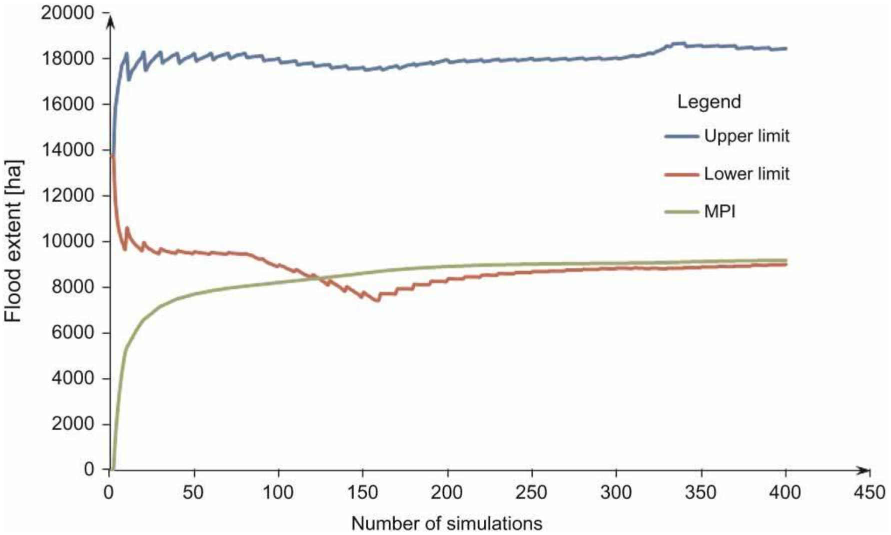

Figure 14 | Mean prediction interval.

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.

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

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.

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest

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

Downloaded from http://iwaponline.com/jh/article-pdf/15/1/55/386919/55.pdf
by guest