



Application of a lower-fidelity surrogate hydraulic model for historic flood reconstruction



A. Bomers^{a,*}, R.M.J. Schielen^{a,b}, S.J.M.H. Hulscher^a

^a University of Twente, Dienstweg 1, Enschede, the Netherlands

^b Ministry of Infrastructure and Water Management-Rijkswaterstaat, Arnhem, the Netherlands

ARTICLE INFO

Keywords:

Lower-fidelity model
Sensitivity analysis
Uncertainty
Historic flood reconstruction

ABSTRACT

Two dimensional hydraulic models are useful to reconstruct maximum discharges and uncertainties of historic flood events. Since many model runs are needed to include the effects of uncertain input parameters, a sophisticated 2D model is not applicable due to computational time. Therefore, this paper studies whether a lower-fidelity model can be used instead. The presented methodological framework shows that a 1D-2D coupled model is capable of simulating maximum discharges with high accuracy in only a fraction of the calculation time needed for the high-fidelity model. Therefore, the lower-fidelity model is used to perform the sensitivity analysis. Multiple Linear Regression analysis and the computation of the Sobol' indices are used to apportion the model output variance to the most influential input parameters. We used the 1926 flood of the Rhine river as a case study and found that the roughness of grassland areas was by far the most influential parameter.

1. Introduction

Currently, the Dutch water policy is changing from a probability exceedance approach towards a risk based approach. In addition to the probabilities of floods due to multiple failure mechanisms, this new approach also considers the consequences of a flood. The risk based approach results in a significant increase in the safety levels in areas where the consequences are large (Dutch Ministry of Infrastructure and the Environment and Ministry of Economic Affairs, 2014). A maximum return period of 1250 years was defined for the river areas in the probability exceedance approach, while the risk based approach has maximum return periods of 100,000 years. The prediction of design discharges corresponding to such rare events is highly uncertain. These predictions are most often based on relatively short data sets of measured weather conditions or discharges. Therefore, the data set does not include the natural phenomena characterised by a very low frequency (Barriendos et al., 2003).

The confidence interval of large design discharges can be reduced by extending the data set of measured discharges with historical and paleo data of extreme flood events (Neppel et al., 2010; Sheffer et al., 2003). Many studies have reconstructed historic floods to expand the data set of measured discharges (e.g. Herget et al. (2015); Herget and Meurs (2010); Llasat et al. (2005); Neppel et al. (2010); O'Connell et al. (2002); Sheffer et al. (2003); Toonen et al. (2015); Zhou et al. (2002)).

Herget et al. (2015) and Herget and Meurs (2010) reconstructed historic discharges in the city of Cologne, Germany, based on historical documents. They predicted mean flow velocities at the time of the historic flood events with the use of a reconstructed river channel and floodplain bathymetry. The empirical Manning's equation was used to estimate the historic discharges of a specific cross section near the city of Cologne. Neppel et al. (2010) used hydraulic modelling of a reach of about two kilometres length to account for geomorphological changes. With this model, present and historic rating curves were constructed and applied to determine flood discharge series (Neppel et al., 2010). O'Connell et al. (2002) used Bayesian statistics to create paleohydrologic bound data for flood frequency analysis. Paleohydrologic bound data represent stages and discharges that have not been exceeded since the geomorphic surface stabilized (O'Connell et al., 2002). These bounds are not actual floods, but are limits on flood stage over a measured time interval. O'Connell et al. (2002) found that paleohydrologic bounds reduce the uncertainties of the flood distribution curve by placing large observed discharges in their proper long-term contexts. Toonen et al. (2015) reconstructed Lower Rhine historical flood magnitudes of the last 450 years with the use of grain-size measurements of flood deposits at two separate research locations. They made use of linear regression plots between various grain-size descriptors and measured discharges to determine the discharges of the historic events.

Above mentioned studies tried to gain insight in the maximum

* Corresponding author.

E-mail address: a.bomers@utwente.nl (A. Bomers).

discharge of a historic flood. However, none of these studies used hydraulic models to describe maximum discharges and its uncertainties along a long stretch of a river including possible bifurcations during the historic events. However, the use of hydraulic models may decrease the confidence intervals of the predicted maximum discharges of the reconstructed flood events. Furthermore, hydraulic models provide insight in the flow patterns and inundation extents of the historic events. For these reasons, hydraulic models will be used for historic flood reconstructions in this study.

Hydraulic models require a reconstruction of the historical geometry as input data. In addition, they require proper boundary conditions to determine the flood wave propagation along the model domain. However, the data available to reconstruct historic flood events is limited. Measured discharges or water levels are generally not available. Also, the geometry of the river, its floodplains and the hinterland may be uncertain. This uncertainty is reflected in the uncertainty of the model input parameters, affecting the maximum discharges during a flood event. For this reason, a sensitivity analysis on the maximum discharge will be necessary to find the input parameter that mostly influences the model output. This analysis will also gain insight in the confidence interval of the reconstructed maximum discharge. This insight provides us with useful information for other historical geometry reconstructions, since parameter prioritization can be used during the reconstruction.

Commonly, sophisticated two dimensional (2D) hydraulic models (in this context also called a high-fidelity model, see Section 2.1) are used for hydraulic modelling. This is because they are capable of describing maximum discharges, flood extent and inundation patterns with high accuracy. However, they have the disadvantage that a single run of a discharge wave usually takes at least several hours. Since sensitivity analyses require many model runs, 2D models are not suitable for this purpose. To reduce computational time, a surrogate model will be set up. A lower-fidelity model is developed since this type of surrogate model does not lose many physical processes of the original system. Therefore, the objective of this paper is to study whether a lower-fidelity hydraulic model can be used for historic flood reconstructions.

Lower-fidelity surrogate modelling has just recently started to gain popularity in the water resources literature (Razavi et al., 2012b). The modelling approach has been applied to groundwater models to reduce model complexity for optimization and calibration purposes (e.g. Maschler and Savic (1999); McPhee and Yeh (2008); Ulanicki et al. (1996)). It has also been applied in combination with the Monte Carlo framework for uncertainty analysis (e.g. Efendiev et al. (2005); Keating et al. (2010)). However, almost no studies have applied a lower-fidelity surrogate model for hydraulic modelling purposes. These models may have great benefits in this field since computational time can be reduced significantly while model accuracy remains sufficient. For an elaborated review on surrogate models in environmental modelling, see Razavi et al. (2012b).

Razavi et al. (2012b) argue that the response patterns of a lower-fidelity model and of a sophisticated 2D model can differ, even if both models are based on the same input data. Therefore, the results of a 2D model will be used for validation purposes. If the model output of the lower-fidelity model is close to those predicted by the 2D model, the lower-fidelity model is capable of accurately simulating the system behaviour. Hence, the lower-fidelity model can be used to perform the sensitivity analysis. For future work, the lower-fidelity model can be treated as a high-fidelity model. The proposed method (Fig. 1) will answer the following three research questions:

- Under what circumstances can a lower-fidelity model be used to simulate a historic flood event?
- How can we apply a lower-fidelity model to compute the maximum discharge and its uncertainty of a historic flood event?
- Which uncertain input parameter contributes most to the

uncertainty of the maximum discharge?

We apply the proposed method to the 1926 flood of the Rhine river. Sufficient information is available to reconstruct the 1926 geometry. In addition, water levels were measured during the event. Due to high rainfall intensities in the Lower Rhine catchment area and increased amount of melting water as a result of relatively high temperatures in Switzerland, the 1926 discharge resulted in the highest discharge at Lobith since measurements have been performed.

The outline of the paper is as follows. Firstly, the high-fidelity (2D) model is described in Section 2.1, after which the surrogate model is set up (Section 2.2). Then, the 1926 case is provided and the methodology of the sensitivity analysis is given in Section 2.3 and Section 3, respectively. Subsequently, the calibration results of the high-fidelity model (Section 4.1) and the validation results of the surrogate model (Section 4.2) are provided. Finally, the results of the sensitivity analysis are elaborated on Section 4.3. The paper ends with a discussion and the main conclusions in Section 5 and 6, respectively.

2. Methodology of surrogate modelling

In this section, the model structure of a fully 2D model is explained. This model represents the high-fidelity model in this study and is used to validate the lower-fidelity model. Thereafter, the 2D model is simplified to decrease computational time significantly. Many methods exist to simplify a high-fidelity model to create a lower-fidelity model. Why a 1D-2D coupled model is used in this study, is explained in Section 2.2.

2.1. High-fidelity model

Most often, 2D flood models are used to get insight in the consequences of high discharge stages. With 2D models, it is possible to get a high detailed and accurate representation of potential floods along a river. Up till now, the 2D Shallow Water equations are usually solved with the use of a curvilinear grid (Fig. 2). The curvilinear grid cells are aligned with the flow direction since flow variations in the channel length direction are often smaller than those in channel cross direction (Kernkamp et al., 2011). This is convenient in terms of computational time. However, a curvilinear grid has several disadvantages. Firstly, grid lines are focused and sometimes even intersect in sharp inner bends (Fig. 2, where the dashed lines indicate the focused grid lines). The focused grid lines result in unnecessarily small grid cells if the model domain is extended in the inner bend. These small grid cells significantly increase computational time. Additionally, the grid will lead to a staircase representation along closed boundaries since the grid is not capable of following the smooth boundaries of the model domain (Kernkamp et al., 2011). Finally, the grid is restrictive in representing a natural river system with different geometric features such as main channels, junction points and wide floodplains due to the curvilinear shape of the grid cells (Lai, 2010).

Due to the above mentioned shortcomings of a curvilinear grid, a hybrid grid is used to solve the 2D Shallow Water equations in this study (Fig. 2). The summer bed is discretized by curvilinear grid cells. These cells are aligned with the flow direction. The winter bed is discretized by triangular grid cells such that each triangular grid cell is connected to a single curvilinear grid cell. As a result, a smooth transition exists between the curvilinear and triangular grid cells (Fig. 2). This hybrid grid overcomes the shortcomings of a curvilinear grid. It also reduces the computational time while model accuracy stays sufficient (Bomers et al., 2019). Fig. 3 shows the hybrid grid and a typical example of model output. The open source software D-Flow Flexible Mesh (FM) is used to set up the 2D model (Deltares, 2016). In each grid cell, parameters such as water level and flow velocity can be computed for every time step. A variable time step is used based on the maximum Courant number. As a result, the model stays stable during the

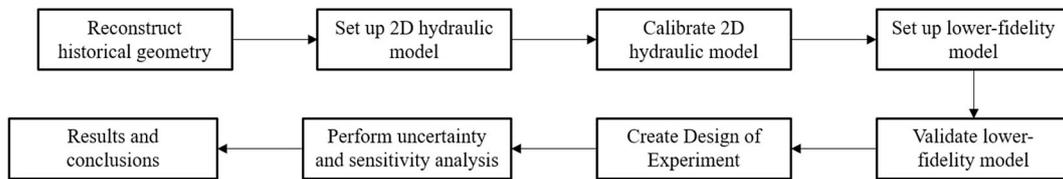


Fig. 1. Methodology for historic flood reconstruction.

simulation:

$$C = \frac{u * \Delta t}{\Delta x} \quad (1)$$

where u represents the velocity magnitude [m/s], Δt the time step [s] and Δx the grid size in x-direction [m]. A maximum Courant number of 0.95 is used and Δt is adapted accordingly.

D-Flow FM allows multiple roughness definitions to be implemented in a single model run, e.g.: a Manning's value, a Nikuradse value or a Van Rijn predictor. In general, the land use classifications, and hence the roughness classes, are based on an input database. A database provided by the Dutch Ministry of Infrastructure and Water Management is used. This database includes multiple roughness definitions that coincide with the land use classification of the studied area.

Calibration of a 2D grid is required since each 2D grid has its own numerical friction caused by the resolution of the grid cells (Caviedes-Voullième et al., 2012). A coarser grid results in a somewhat dampened discharge wave. This effect can even become larger than those generated by physical friction (Caviedes-Voullième et al., 2012). During calibration, this numerical grid generated friction will be compensated such that reliable water levels are predicted. Hydraulic model calibration is most commonly done by changing the roughness of the summer bed until simulated water levels are close to measured water levels (e.g. Bomers et al. (2019) and Caviedes-Voullième et al. (2012)). In this study, the same approach was used. The calibration procedure was performed with the use of the open source software OpenDA (<http://www.openda.org/>). The basic idea of the procedures of OpenDA is to find the set of model parameters which minimizes the cost function measuring the distance between the measured water level and the model prediction (The OpenDA Association, 2016). The Quadratic Cost Function is used in combination with the Sparse DUD (Does not Use Derivate) algorithm. For N calibration parameters (in this study $N = 10$), the algorithm requires $(N + 1)$ set of parameter estimates. The cost function, based on the model predictions and measured data, is used to get a new estimate. If the cost function does not produce a better estimate, the Sparse DUD algorithm will search in opposite direction and/or decreases the searching-step until a better estimate is

found (The OpenDA Association, 2016). In this study, the calibration procedure is stopped if the average RMSE of each measurement station is smaller than 0.05 m. For more information on the calibration procedure of OpenDA, see The OpenDA Association (2016).

2.2. Lower-fidelity physically based surrogate model

A hybrid 2D grid reduces computational time compared to a curvilinear grid. However, the computational time of simulating a discharge wave of approximately three weeks is still in the order of many hours. For sensitivity analysis purposes, many model runs (120 in this study) have to be performed. Therefore, a model with a computational time in the order of minutes is desirable. For this reason, a surrogate model based on the high detailed 2D model is developed. This model is explained in more detail in the next sections.

2.2.1. Types of surrogate modelling

Surrogate models approximate the response pattern of a high detailed and computationally intensive simulation model (Razavi et al., 2012a). Many methods to construct a surrogate model exist in literature. These methods can be divided into two classes, namely (1) *response surface surrogates* which are statistical or empirical data-driven models emulating the original system, and (2) *lower-fidelity physically based surrogates* which are simplified models of the high detailed model (Razavi et al., 2012b).

Regardless of the type of response surface surrogates, usually three steps are involved (Simpson et al., 2001): (1) choosing a design of experiment for generating the training data, (2) choosing a statistical or empirical data-driven model (e.g. Artificial Neural Network, Support Vector Machine, Gaussian Process Regression model) to represent the data, and (3) fitting the surrogate model to the training data. Response surface surrogates are commonly used for automatic model calibration (Razavi et al., 2012b). To fit the response surface surrogate, training data is required. Therefore, the high-fidelity model still needs to be run multiple times. Because of the relatively long simulation time of this model, the methods based on response surface surrogates are not

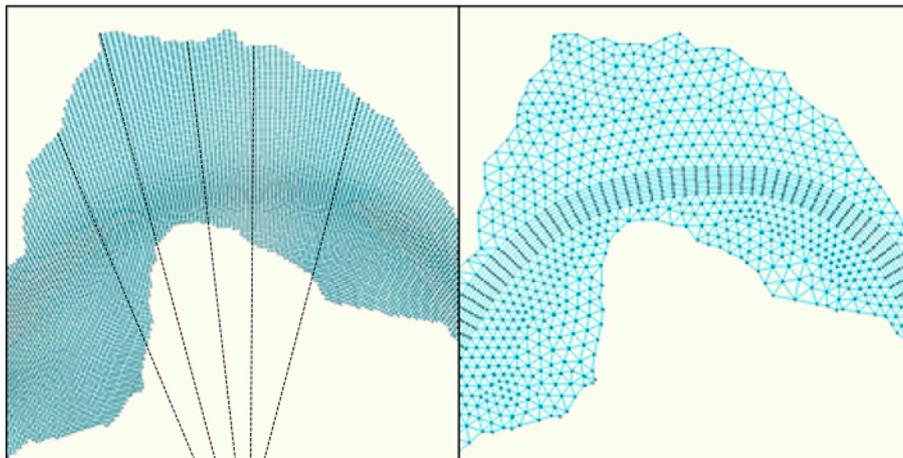


Fig. 2. An example of a curvilinear grid in which the dashed lines represent the focused grid lines (left figure) and a hybrid grid (right figure) in a sharp meander bend.

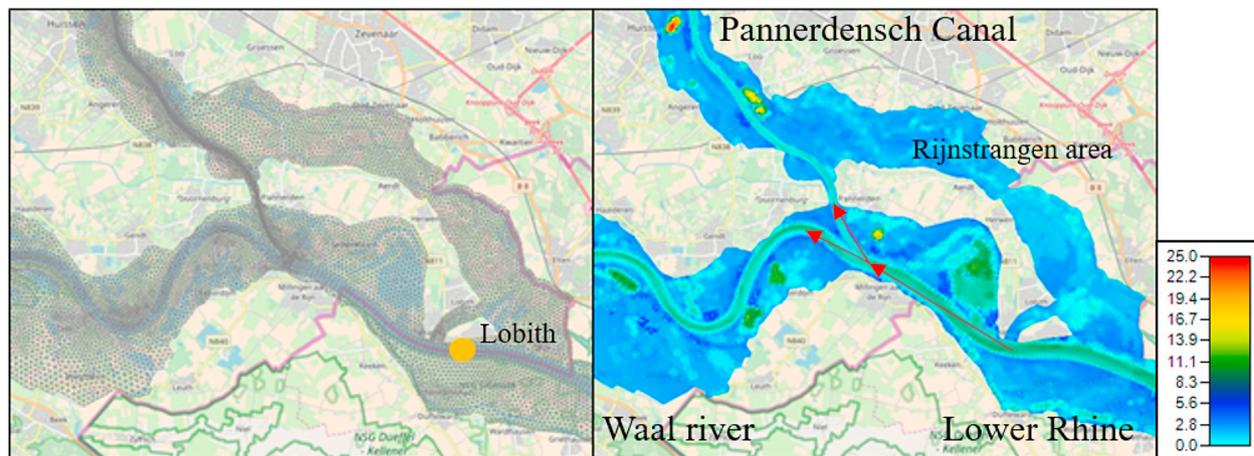


Fig. 3. Example of a hybrid grid of a 2D hydraulic model (left figure) and computed water depths as a result of an upstream discharge wave (right figure). The red arrows indicate the flow direction. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

desirable. For this reason, the high-fidelity model is simplified using method (2): creating a lower-fidelity physically based surrogate model. Lower-fidelity surrogate models are set up based on the original input data. Therefore, for lower-fidelity modelling, only a single run with the high-fidelity model is required for validation purposes. Moreover, lower-fidelity models are more reliable in predicting the output of the high-fidelity model in unexplored regions of the input space since they predict model output based on the original input data (Razavi et al., 2012b). Different methods exist to simplify the original model, e.g. larger grid size, less strict numerical convergence tolerances or, ignoring or approximation physics of the original system (Razavi et al., 2012b). Those methods were not sufficient to reduce the computational time of the high-fidelity model significantly. Therefore, it was decided to approximate several physical processes of the original system by: (1) lowering the dimension of the model, (2) increasing the computational time step, and (3) simplifying the Shallow Water equations of the fully 2D model. The set-up of the lower-fidelity model is explained in the next section.

2.2.2. Set-up lower-fidelity model

The surrogate model developed represents a 1D-2D coupled model to combine the advantages of both a fully 2D and a fully 1D model. 1D profiles give an accurate representation of flood wave propagation in case of in-channel flows (Tayefi et al., 2007). Additionally, the computational cost is relatively low compared to a fully 2D model. However, the use of 1D profiles may be insufficient for more complex flow patterns because of the simplified assumptions in the computational schemes. In the embanked areas rapidly changes in flow velocity and direction may occur. For this reason, 1D profiles are solely used for the flow between the winter dikes, i.e. the summer bed and winter bed. The 1D profiles are coupled with 2D embanked areas that are possible to inundate. The embanked areas refer to the areas protected by dikes and are therefore not part of the river system. The embanked areas are discretized with a rectangular 2D grid. Flexible grid shapes are used along the boundaries of the model domain such that the 2D grid cells follow these boundaries. The flexible grid cells along the boundaries can have a maximum of eight boundary edges. Fig. 4 shows an example in which the 1D profiles of the rivers and the 2D embanked areas are given by yellow lines and grey areas, respectively. A close-up of the 2D grid and its flexible grid shapes along the grid boundaries is also provided.

HEC-RAS (v. 5.0.3), developed by the Hydrologic Engineering Centre (HEC) of the US Army Corps of Engineers, is used for the 1D-2D flood modelling. HEC-RAS is well known for its 1D flood modelling applications. Horritt and Bates (2002) even showed that HEC-RAS produces flood extents more accurately than the 2D models of

LISFLOOD-FP and TELEMAC-2D in cases of a confined and relatively narrow river. In 2016, HEC-RAS 5.0 was officially released. With this version, it is possible to perform 1D-2D coupled computations.

Several studies have shown the applicability of 1D-2D flood modelling. Most software programs (e.g. Mike-11, HEC-RAS) that allow 1D-2D coupling are based on mass-conservation. The conservation of momentum is often neglected. Bladé et al. (2012) argue that neglecting the momentum in the coupling of a 1D profile and the 2D grid cells affects flow patterns in the floodplains in most cases. The more connected the river and the floodplains are, e.g. in case of overland flows, the more important momentum becomes since an increase in flow velocity results in an increase in momentum (Bladé et al., 2012). Conservation of momentum can only be neglected if the 1D profiles are coupled with 2D grid cells by a weir/embankment since the hypothesis of the Shallow Water equations are not fulfilled for this specific case (Bladé et al., 2012). With HEC-RAS, the weir-equation can be used to compute the flow over the embankment using the results of the 1D and 2D solution algorithms on a time step by time step basis. This allows for direct feedback at each time step between the 1D profiles and 2D grid cells (Brunner, 2014). Neglecting conservation of momentum is justified for this modelling purpose since the 1D profiles are coupled with the 2D grid cells by an embankment. Hence, the 1D-2D coupling can be treated as a weir-type connection.

2.2.3. Differences between the high-fidelity and lower-fidelity model

A 1D-2D coupled model requires the same input data as a fully 2D model. Therefore, we use the same input data of the high-fidelity model to set up the 1D-2D coupled model. The Digital Elevation model (DEM) of the 2D model is used to establish the 1D profiles and 2D grid cells of the 1D-2D coupled model. Also, the boundary conditions consisting of measured discharges and water levels, as well as the land use classification for both models are identical. Therefore, we can conclude that the differences in the representations of the input parameters of the high-fidelity and the lower-fidelity model are solely caused by the level of detail of the two models itself and the different settings of D-Flow FM and HEC-RAS. These differences are explained in more detail below and are summarized in Table 1.

Firstly, the 2D Shallow Water equations of the high-fidelity model are simplified to the Diffusive Wave equations. The Diffusive Wave equations are applicable if flow separation and turbulence eddies can be neglected. This is the case if the inertial terms are much smaller than the gravity, friction and pressure terms. Test runs showed that neglecting the inertial terms of the momentum equations did not result in a change in model results. On the other hand, the use of the Diffusive Wave equations resulted in a significant reduction of the computational time. Therefore, the Diffusive Wave equations are used to compute the

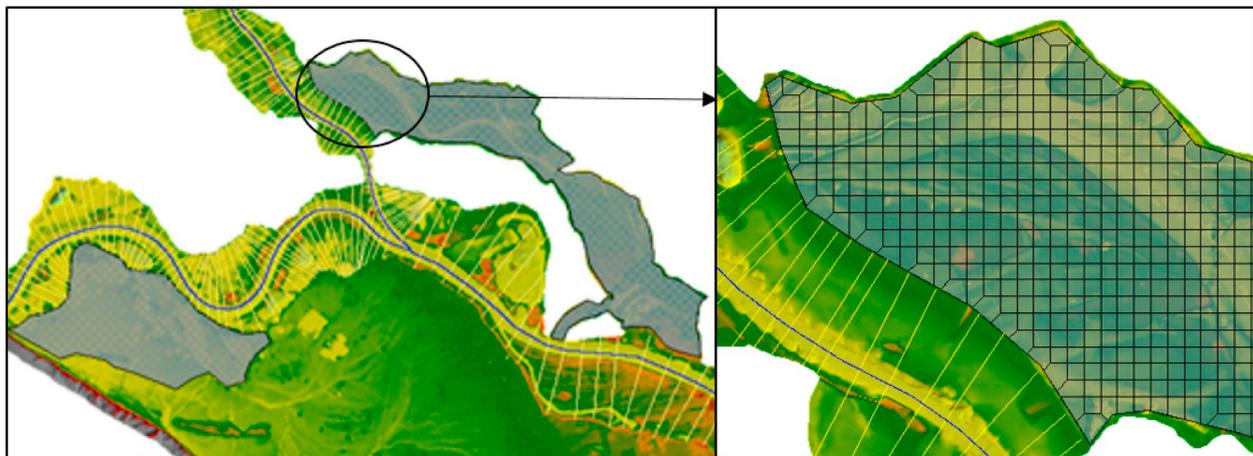


Fig. 4. Set-up of the 1D-2D coupled model (left figure) in which the yellow lines indicate the 1D profiles and the grey areas the 2D embanked areas, and a close-up of the 2D grid which clearly shows the flexible grid shapes along the boundaries of the model domain (right figure). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 1

Overview of the differences between the high and lower-fidelity model.

	High-fidelity model	Lower-fidelity model
Software	D-Flow FM	HEC-RAS
Dimension	Fully 2D	1D cross sections in summer bed and winter bed, 2D in the embanked areas
Nature	Shallow Water equations	Diffusive Wave equations
Time step	Variable time step based on maximum Courant number	Fixed time step
Roughness	Different roughness definitions	Manning's roughness coefficient
Calibration	Calibrated summer bed roughness	Uncalibrated

flow characteristics at each 1D profile and 2D grid cell. The applicability of the Diffusive Wave equations for flood modelling purposes has also been shown by e.g. Moya Quiroga et al. (2016), Moussa and Bocquillon (2009) and Leandro et al. (2014).

Secondly, the computational time step of the lower-fidelity model is increased compared to the fully 2D model to speed-up computational time. In a 2D model, the river is usually the time step limiting factor since the depths and velocities in the main channel are larger than in the embanked areas (Bladé et al., 2012) (see equation (1)). The high-fidelity model had an average time step of 3.9 s, based on the maximum Courant number. A fixed time step of five minutes can be used for the lower-fidelity model. This time step is based on a convergence argument: reducing the time step further did not result in a reasonable improvement of the model accuracy.

The land use classification of the high-fidelity model is used as input for the lower-fidelity model. D-Flow FM allows multiple roughness definitions to be implemented in a single model. However, HEC-RAS only allows a Manning's roughness coefficient for the various land use classes. Therefore, the roughness classes as used in the high-fidelity model were transformed towards Manning's roughness values based on Tables 5–6: “Values of the roughness coefficient n ” of Chow (1959).

We recall that it is necessary to calibrate the summer bed roughness of the high-fidelity model, since each 2D grid has its own numerical friction. On the other hand, it is decided to not calibrate the lower-fidelity model. As a result, the summer bed roughness can be included in the sensitivity analysis as a random parameter. This is justified since no inundations along the Lower Rhine occurred during the 1926 flood event. Therefore, correct prediction of the water levels becomes irrelevant. The lower-fidelity model is set up to accurately predict maximum discharges at Lobith during flood events instead. During the

simulation, the entire discharge wave flows in downstream direction independent of simulated water levels, since inundations are not possible to occur. Consequently, it is expected that simulated maximum discharges of the uncalibrated surrogate model are close to those predicted by the calibrated high-fidelity model. However, validation is recommended to study whether the lower-fidelity model is capable of simulating the system behaviour sufficiently.

2.2.4. Validation lower-fidelity model

Razavi et al. (2012b) argue that, even though the lower-fidelity model may be based on the same input parameters as the high-fidelity model, the response pattern can differ somewhat. This was also shown by Thokala and Martins (2007). They neglected the fluid viscosity in the Navier-Stokes equations to set up a lower-fidelity model. This resulted in less accurate results compared to the high-fidelity model. The discrepancies between the response patterns of the lower-fidelity and high-fidelity models mostly influence the local and global minimum and maximum of the system (Razavi et al., 2012b). Since this study tries to predict maximum discharges during a historic flood event, it is of high importance that the global maximum of the system is correctly modelled by the lower-fidelity model. If this is not the case, the discrepancies between the lower-fidelity and high-fidelity model can be addressed with a correction function (Razavi et al., 2012b). These kind of functions correct the response of the lower-fidelity model and align it with the response pattern of the high-fidelity model. It is thus of high importance to validate the lower-fidelity model to study whether a correction function is required to tune the model results.

If the response pattern of the lower-fidelity model is close to that of the high-fidelity model, the lower-fidelity model can be treated as the high-fidelity representation of the underlying system. Consequently, the lower-fidelity model can replace the sophisticated 2D model (Razavi et al., 2012b). The sensitivity analysis can then be safely performed with the lower-fidelity model since the input parameters of the lower-fidelity model are based on the input parameters of the high-fidelity model.

2.3. The 1926 casus

The 1926 flood event of the Rhine river is used to examine the methodology of developing a lower-fidelity model for historic flood reconstruction. The study area stretches from the areas downstream of Andernach in Germany to the three Rhine river branches in the Netherlands (Fig. 5). In this paper, the German part of the river is referred to as the Lower Rhine. The river enters the Netherlands at Lobith, where it bifurcates into the Waal river and Pannerdensch Canal.

Subsequently, the Pannerdensch Canal bifurcates into the Nederrijn and IJssel rivers. Only the summer bed, its floodplains and two embanked areas that are connected by an inlet (Ooijpolder and Rijnstrangen area (Fig. 5)), are captured in the model domain. The term inlet is used for a dike section with a relatively low crest level. Due to this low crest level, a part of the discharge wave will enter the lower-lying area behind the inlet as soon as a certain water level is exceeded. As a result, the maximum discharge further downstream decreases. The dikes represent the boundaries of the model domain and are assumed not to overflow.

2.3.1. Geographical situation

To reconstruct a historical geometry, the changes in the river system between the current geometry and the historical period of interest must be defined. An existing data set representing the 1995 geometry is made available by the Dutch Ministry of Infrastructure and Water Management. This data set is used as starting point and is adapted such that it represents the historical geometry. The following measures were taken to create the 1926 situation (Fig. 5):

- **Increase summer bed level due to erosion.** Measurements of the summer bed levels were available for the entire model domain. The changes in summer bed level between the 1995 measurements and the oldest measurements available at each location were used to estimate the 1926 summer bed level by linear extrapolation.
- **Decrease winter bed level due to sedimentation.** No measured sedimentation rates along the study area were available. Therefore, the following sedimentation rates were used to predict the 1926 winter bed level: 1 mm/year along the IJssel river, Pannerdensch Canal and Lower Rhine, 3 mm/year along the Waal river and 0.5 mm/year along the Nederrijn river (Silva et al., 2001). A linear decrease of the sedimentation rate in channel cross direction was assumed. As a result, the sedimentation near the summer bed equals the predicted sedimentation rates according to Silva et al. (2001). The sedimentation near the outside border of the floodplain equals zero.
- **Dike relocation.** On the left side of the Lower Rhine, close to the city of Emmerich, Germany, the floodplains of the river were much larger in 1926 than they are nowadays. The 1926 dike locations and hence the 1926 winter bed were based on old maps dating back to

1895 (Fig. 5, Dike relocation), provided by the German Deichverband Xanten-Kleve Der Oberdeichinspektor Dusseldorf (1895). The current summer dikes along the Pannerdensch Canal close to the Pannerdensche Kop were the 1926 winter dikes. Therefore, the present floodplains were not part of the 1926 river system. The area outside the 1995 summer dikes were removed from the geometry (Fig. 5, Pannerdensche Kop).

- **Restoration of inlets.** In 1926, two retention areas were possible to inundate at high discharge stages as a result of inlets. The Spijke inlet caused inundation of the Rijnstrangen area when the water level exceeded 15 m + NAP, equal to the crest level of the inlet (Fig. 5, Rijnstrangen area). In the Ooijpolder, three inlets were active. The total length of the inlets was 150 m. The Ooijpolder started to inundate at a water level of 12.5 m + NAP, equalling the height of the three inlets. The location of the inlets was based on historical 1926 maps (Fig. 5, Ooijpolder).
- **Restoration of meander cut offs.** In 1955 and 1969 two meanders near Doesburg and Rheden were cut off (Fig. 5, Meander cut offs). Due to these meander cut offs the total length of the IJssel river decreased with almost nine kilometres. The location of the meander bends are based on historical 1926 maps.

2.3.2. Boundary conditions

The 1926 flood event is simulated for a period of approximately three weeks, starting on the 22nd of December 1925 till the 8th of January 1926. From the 26th of December onwards, the weather conditions changed drastically. High rainfall intensities occurred in almost the entire catchment area of the Rhine river (Dutch Ministry of Infrastructure and the Environment, 1926). This resulted in a rapid rise of the discharge wave, starting on the 27th of December.

Fig. 6 shows the discharge wave at Andernach, representing the upstream boundary condition (Data source: German Federal Waterways and Shipping Administration (WSV), communicated by the German Federal Institute of Hydrology (BfG)). The downstream boundary conditions consist of h(t)-relations based on daily measured water levels available at <http://waterinfo.rws.nl> and provided by the Dutch Ministry of Infrastructure and Water Management. Three streams enter the Lower Rhine, namely the Lippe, Ruhr and Sieg rivers. These streams

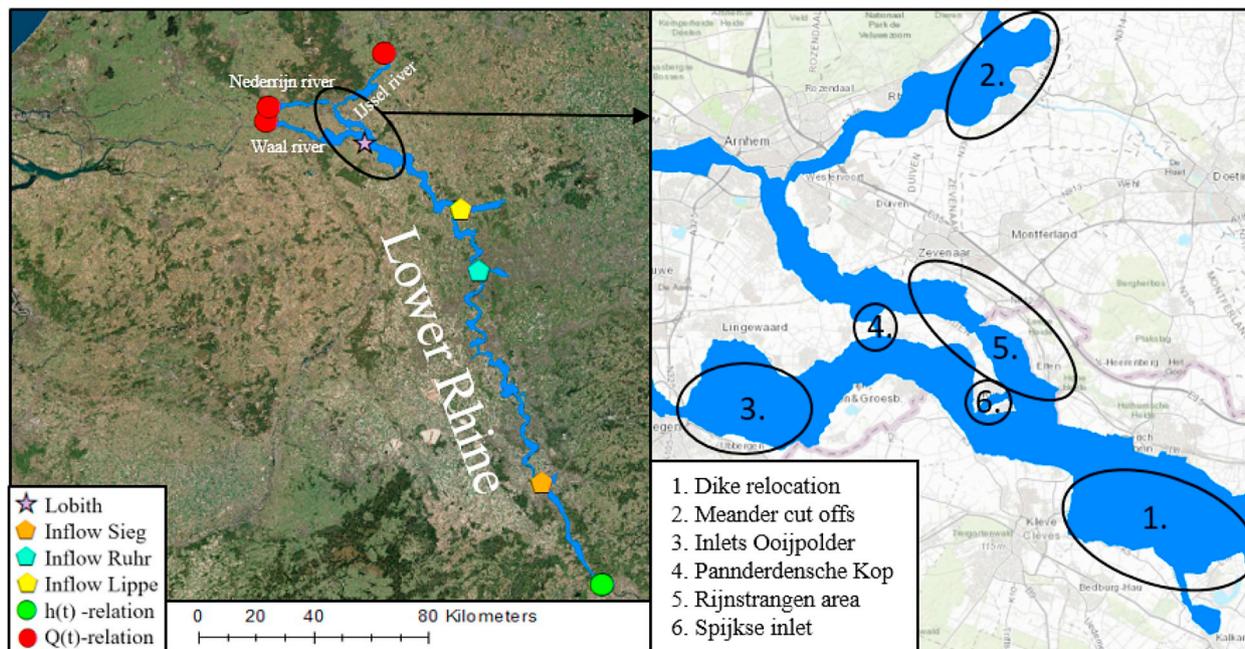


Fig. 5. Boundaries of the study area and location of inflow (left figure) and location of artificial measures taken to change the 1995 geometry into the 1926 situation (right figure).

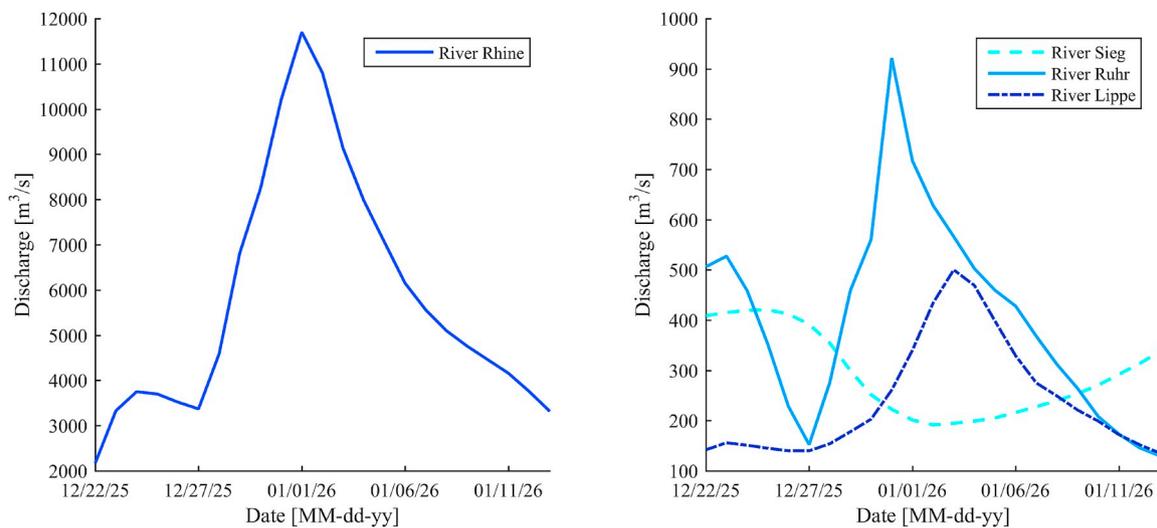


Fig. 6. Discharge waves of the Rhine river at Andernach (left figure) and the three tributaries Sieg, Ruhr and Lippe (right figure). Note: only daily discharge measurements are available resulting in the sharp peaks of the different discharge waves.

were included in the model domain by source points (discharge inflow, Figs. 5 and 6). The presented boundary conditions and source points are used in both the high-fidelity as well as the lower-fidelity model to set up the models.

3. Methodology of sensitivity analysis

In this study uncertainty and sensitivity analyses are performed. An uncertainty analysis is executed to compute the maximum discharge at Lobith with its standard deviation as a result of the uncertain input parameters. Next, a sensitivity analysis is performed to study which parameter mostly influence the uncertainty of the model output. The main objective of the sensitivity analysis is the so called factor prioritization. With this prioritization, it becomes clear on which parameter to focus during historical geometry reconstruction for flood modelling purposes in order to reduce the potential uncertainty in the model output.

During the analyses, we only focus on the parameters that influence the maximum discharge at Lobith. A test run was performed in which all roughness parameters along the Dutch river branches were increased with 20%. In this run, the roughness values are close to the upper bound of the truncated normal distributions. The run showed that the increase in roughness resulted in only a minor decrease of the maximum discharge at Lobith of approximately 0.2%, from 12,402 to 12,373 m³/s. This minor decrease suggests that the Dutch river branches are sufficiently downstream such that the effects of different summer bed roughness on the maximum discharge are negligible. Therefore, the study only focuses on the uncertainties of the input parameters in the most upstream part of the model domain: the city of Andernach until the location where the Rhine river bifurcates into the Waal river and Pannderdench Canal. The Dutch Rhine river branches are seen as fixed boundary conditions of the model since they do not influence model response. Therefore, they can be excluded from the global sensitivity analysis.

3.1. Input parameters

The lower-fidelity model is used to establish the uncertainty and sensitivity of the 1926 discharge at Lobith. Only the input parameters that are based on an estimation, i.e. those that are uncertain, are included in the analysis. In addition, parameters that require the development of a new surrogate model when changed (e.g. a planometric change) are excluded from the analysis for pragmatic reasons. The

following parameters are considered during the sensitivity analysis: (1) roughness parameters of the various types of land use classes and (2) the bed levels of the summer bed and winter bed. In general, two kinds of uncertainties exist. The first uncertainty is as a result of the randomness of variations in nature (inherent uncertainty). The second uncertainty is caused by limited knowledge (epistemic uncertainty) (Warmink et al., 2013). The uncertainty of the different roughness classes is mainly caused by inherent uncertainty since it depends amongst others on the season (e.g. grass grows faster during summer periods resulting in a larger roughness) as well as on maintenance (e.g. the frequency of mowing grass fields). The uncertainty of the summer bed and winter bed levels are caused by epistemic uncertainty. No measured 1926 bed levels are present. Therefore, the bed levels are based on extrapolation techniques and estimated sedimentation rates.

For all roughness parameters, we link the value with the largest probability of occurrence as well as its minimum and maximum bounds to the tables of Chow (1959). Truncated normal distributions are used in this study since a normal distribution better fits the data if some information about the input parameters is available (tails of the distribution and the expected value). Contrarily, a uniform distribution assumes that there is no knowledge about the value with the largest probability of occurrence. Only a range of input values is known. Therefore, we can conclude that for older historic events, the distributions of the uncertain input parameters will shift towards uniform distributions since less and less information is available.

The roughness parameters are divided into five land use classes: summer bed, lakes, grasslands, forest and urban areas. A smooth channel with no vegetation is assumed for the entire summer bed, having a minimum Mannings roughness of 0.025, a normal value of 0.028 which is used as the expected value, and a maximum value of 0.033 (Chow, 1959). These numbers are used to set up the truncated normal distribution. The same method was used to define the truncated normal distributions of the other roughness classes (Table 2).

A comparable method is used to set up the truncated normal distributions of the summer bed levels and winter bed levels. The 1926 summer bed levels were computed based on extrapolation of measured bed level changes (see Section 2.3). The uncertainty ranges of the summer bed levels were based on these extrapolation values. The minimum change in bed level corresponds to no change compared to the oldest measured bed value. Consequently, the 1926 bed level equals the oldest measured bed level. The maximum change in bed level equals the extrapolation of the trend between 1995 and the latest measured bed level multiplied with a factor two. A factor of two is chosen to

Table 2
Minimum, maximum and standard deviation of the different input parameters.

Input parameter	Minimum value	Maximum value	Standard deviation
Summer bed	0.025	0.033	0.002
Lakes	0.024	0.034	0.003
Graslands	0.037	0.075	0.009
Forest	0.098	0.178	0.020
Urban areas	0.029	0.039	0.003
St. Winter bed level	−0.070 m	0.070 m	0.035 m
St. Summer bed level (1)	−0.150 m	0.150 m	0.075 m
St. Summer bed level (2)	−0.520 m	0.520 m	0.260 m
St. Summer bed level (3)	−0.090 m	0.090 m	0.045 m

include a large uncertainty range. The summer bed is divided into three classes:

1. From the most upstream location Andernach (river km 614) until Walsum (river km 789). Here, almost no erosion has occurred between 1995 and 1926. Additionally, the bed level has been compensated for bed level decrease due to mining activities at several locations.
2. From Walsum until the German-Dutch border (river km 857). Here, there is relatively much uncertainty in the amount of erosion since the oldest measured bed level dates back to only 1960.
3. From the German-Dutch border till the first bifurcation point of the Rhine river (river km 867). Here, there is little uncertainty in the 1926 bed level since the oldest measurements date back to 1934.

The winter bed level consists of just one class since no deviations in uncertainty along the Lower Rhine exist. The estimated sedimentation rate of 1 mm/year is used to define the ranges of the winter bed level in the Lower Rhine (Silva et al., 2001). The minimum value equals no change in bed level compared to the 1995 situation. The maximum range equals the sedimentation rate of 1 mm/year multiplied with a factor of two. Again a factor of two is chosen to include a large uncertainty range since the 1 mm/year sedimentation rate is relatively speculative. Since the summer bed and winter bed levels vary along the study area, their truncated normal distributions and corresponding minimum and maximum values are given as change from its 1926 reference value (Table 2). These values will be referred to as Standardized (St.) bed levels from now on. A value equal to zero correspond with the reconstructed 1926 geometry.

3.2. Design of experiment

Before a sensitivity analysis can be performed, a Design of Experiment (DoE) has to be defined. DoEs employ different space filling strategies to capture the behaviour of the underlying system over limited ranges of the input parameters (Razavi et al., 2012b). A DoE results in a sample in which the boundary values of the input parameters are based on physical conditions. This sample can be used in a Monte Carlo analysis. Most commonly used DoE methods in literature appear to be full factorial design, fractional factorial design, central composite design and latin hypercube sampling (LHS) (Razavi et al., 2012b). In general, a full factorial design, a fractional factorial design and a central composite design require a relatively large number of simulations to generate all combinations to represent the corners of the input space (Razavi et al., 2012b; Saltelli et al., 2008). Contrarily, LHS can easily scale to different numbers of input parameters without the need for extra simulation runs (Razavi et al., 2012b). Thus, a stratified LHS sample has as advantage that less model runs are required since a stratified sample achieves a better coverage of the sample space of the

input parameters (Saltelli et al., 2000). For this reason, a LHS design is used in this study.

The nine input parameters are divided into eight levels. Each level has an equal probability of occurrence of 12.5%, based on the determined truncated normal distributions in Section 3.1. For each run, each level is randomly selected, constraining that if a level is already selected it cannot be selected again. This results in a set of eight simulations in which all eight levels of the nine input parameters are present.

No clear guidelines exist concerning the minimal number of runs required in a Monte Carlo analysis. This number depends on the number and range of the input parameters and on the shape of the response surface. These features are largely unknown in advance (Pappenberger et al., 2005). In this study convergence of the uncertainty of the discharge at Lobith, expressed as standard deviation, is used as stopping-criteria, following the method of Pappenberger et al. (2005). If an additional run results in a change of the standard deviation smaller than 0.05 m³/s, it is assumed that the sample sufficiently represents the input space of the different input parameters. This criteria resulted in 120 model runs, corresponding with 15 latin hypercube sets.

To check whether the input space is sufficiently captured by the sample, two additional model runs were performed with the most extreme situations. These scenarios represent the limits of the probability distribution functions of the input parameters. Table 3 and Fig. 7 show the range of maximum discharges at Lobith modelled in the 120 Monte Carlo runs and the range found with the two most extreme cases. Note that all runs are performed with the lower-fidelity surrogate model. The minimum and maximum values of the sample are close to the predicted values of the two most extreme runs. Therefore, we can conclude that the input space is sufficiently captured by the sampling data set.

3.3. Stratified Monte Carlo analysis

The results of the Monte Carlo analysis are used to determine the uncertainty in model predictions. Additionally, the results are used to apportion this uncertainty to the contribution of the individual input parameters. Two sensitivity analysis methods are used, namely Multiple Linear Regression analysis and Sobol' indices explained in Sections 3.3.1 and 3.3.2 respectively.

3.3.1. Multiple linear regression analysis

If the number of simulations is much larger than the number of input parameters, a LHS can be very effective in revealing the influence of each parameter using a regression analysis (Saltelli et al., 2008). If the model does not contain any interactions between the input parameters (i.e. the model is additive), the linear regression function can be given as (Scheidt et al., 2018):

$$y = \beta_0 + \sum_{i=1}^N \beta_i x_i \quad (2)$$

where y represents the model output (in this study the maximum discharge at Lobith) and x_i the different input parameters. The coefficients β_0 and β_i are determined by the least-square computation, based on the squared differences between the model output produced by the regression model and the actual model output produced by the surrogate model (Saltelli et al., 2008).

Table 3

Minimum and Maximum discharge at Lobith (Qmin, Lobith/Qmax, Lbith) as a result of the two most extreme model runs and the 120 runs within the Monte Carlo (MC) analysis.

	Extreme case	MC runs	Difference
Qmin, Lobith [m ³ /s]	12,285	12,293	17
Qmax, Lobith [m ³ /s]	12,548	12,531	8

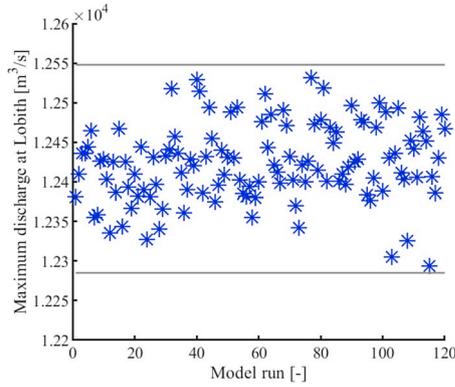


Fig. 7. Input space of the LHS representing the maximum discharges at Lobith modelled during each model run. The grey lines indicate the results of the two most extreme model runs.

The coefficient β_i is used to determine the importance of each parameter x_i with respect to the model output. If the input parameters are independent, the absolute standardized regression coefficient $\hat{\beta}_i$ can be used as a measure of sensitivity (Scheidt et al., 2018):

$$\hat{\beta}_i = \left| \beta_i \frac{\sigma_i}{\sigma_y} \right| \quad (3)$$

where $\hat{\beta}_i$ represents the standardized regression coefficient, and σ_i and σ_y represent the standard deviations for the input parameter x_i and the model output respectively.

However, the applicability of a linear regression analysis depends on the degree of linearity of the model (Saltelli et al., 2008). A measure for linearity is expressed by (Saltelli et al., 2008):

$$R^2 = \sum_{i=1}^N (\hat{\beta}_i)^2 \quad (4)$$

where R^2 represents the model coefficient of determination. This value is equal to the fraction of the variance of the original data that is explained by the regression model. A value of R^2 equal to one indicates that the model is linear (Saltelli et al., 2008) and that the multiple linear regression model is capable of expressing all variance of the original data.

3.3.2. Sobol' indices

If the model is not linear, Sobol' indices can be used to determine the sensitivity of the input parameters. Sobol' indices are widely used as global sensitivity analysis method in literature. We are specifically interested in the first-order indices, i.e. the effect without interactions of input parameters, since the sensitivity analysis is used for factor prioritization purposes (Saltelli et al., 2008). Li and Mahadevan (2016) present an effective method to estimate the first-order Sobol' indices analytically. This method can be applied to any kind of data set and is not restricted to a specific sampling strategy. Furthermore, the method can be applied to models with correlated input parameters. Li and Mahadevan (2016) found that the method is highly efficient and that it is especially useful in ranking and identifying important parameters. The formula used is as follow (Li and Mahadevan, 2016):

$$S_i = 1 - \frac{E_{x_i}(V_{x_{-i}}(y|x_i))}{V_y} \quad (5)$$

where S_i represents the Sobol' first-order index, $V_{x_{-i}}(y|x_i)$ indicates the conditional variance of y caused by all input parameters other than x_i , E_{x_i} represents the expected value as a result of fixing input parameter x_i , and V_y represents the variance of y .

The Monte Carlo sample has a relatively small size. Therefore, the 95% confidence intervals of the Sobol' indices are computed based on a

resampling strategy. The MATLAB Statistics Toolbox is used to perform the computation. The method to compute the 95% confidence intervals is based on the work of Dubreuil et al. (2014) in which a bootstrap resampling strategy is used. Computation of confidence intervals by bootstrap resampling is widely used in global sensitivity analysis and has been used in combination with surrogate models by Gayton et al. (2003) and Janon et al. (2011). Bootstrap resampling aims at determining confidence intervals of a parameter of interest using only one design of experiment (Efron and Tibshirani, 1993). The method consists of the creation of new designs of experiment by drawing with replacement in the original design.

The method used is presented in Fig. 8. The LHS sample consisting of 120 model runs is resampled, after which the confidence intervals of the first-order Sobol' indices are computed. If these confidence intervals have not reached a specific convergence criterion yet, more bootstrap resamples are drawn. The computation is repeated until the convergence criterion is met. The criterion as suggested by Dubreuil et al. (2014) is used. They suggested to stop the procedure at the iteration for which all confidence interval sizes have reached a range which is less than x percent of the maximum bootstrap mean of the sensitivity indices. The choice of parameter x depends on the goal of the sensitivity analysis. If the goal is only determining the most dominant input parameter, a relatively large value of x in the order of 30% can be used. However, if the model has many variables of equal sensitivity indices, it is better to look at the convergence graph at each bootstrap iteration and decide manually when to stop the procedure (Dubreuil et al., 2014). The first convergence criteria (30%) is used which will be evaluated by checking the convergence graphs of the Sobol' indices as suggested by Dubreuil et al. (2014).

4. Results

4.1. Calibration high-fidelity model

The river branches Lower Rhine, Waal river and Pannerdensch Canal were calibrated with the use of measured water levels. The discharge partitioning along the Dutch river branches was based on the report of the Dutch Ministry of Infrastructure and the Environment (1952). During the calibration procedure, this discharge partitioning had to be met. The IJssel and Nederrijn rivers were excluded from the calibration procedure since many inundations along the IJssel river have occurred during the 1926 flood event. These inundations influence the water levels at both river branches. Even a very low summer bed roughness near the locations of the inundations did not result in the correct water levels. For this study purpose, it is accepted that the water levels along the IJssel and Nederrijn rivers were not calibrated

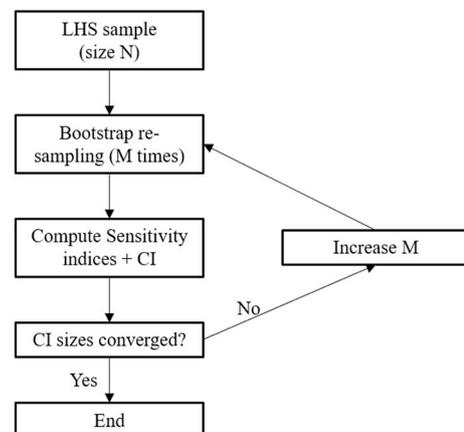


Fig. 8. Bootstrap method for resampling in which CI stands for confidence intervals.

Table 4
Maximum discharges along the Lower Rhine and discharge partitioning along the Dutch Rhine river branches predicted by the high-fidelity and lower-fidelity model, where Qmax represents the maximum discharge at the specific location.

	High-fidelity model	Lower-fidelity model	Difference [%]
Qmax Bonn [m ³ /s]	11,509	11,580	0.6
Qmax Cologne [m ³ /s]	11,632	11,715	0.7
Qmax Dusseldorf [m ³ /s]	11,365	11,598	2.1
Qmax Rees [m ³ /s]	12,351	12,572	1.8
Qmax Emmerich [m ³ /s]	12,297	12,453	1.3
Qmax Lobith [m ³ /s]	12,282	12,402	1.0
Waal river [%]	70.3	71.9	1.5
Pannerdensch Canal [%]	29.7	28.0	1.7
Nederrijn river [%]	58.7	56.2	2.4
IJssel river [%]	41.4	43.8	2.4

correctly. These branches are located more than 15 km downstream of Lobith such that backwater effects has vanished at Lobith. The IJssel and Nederrijn rivers have thus no effect on the maximum discharge at this location.

In the data set, only daily measured water levels are available. Hence, the maximum measured water level may be lower than the occurred maximum water level. Therefore, we calibrated on the three days with the highest water levels for each measurement station present along the river branches. If the model is capable of predicting the correct shape and correct water levels at three moments in time near the peak discharge, it is likely that also the correct maximum water level is predicted by the model.

The 1926 discharge wave was simulated. Maximum water levels at 10 measurement stations were validated after model calibration. It was found that simulated maximum water levels only deviated 2 cm on average compared to the measurements. Therefore, it can be concluded that the high-fidelity model is capable of simulating maximum water levels with high accuracy after calibration of the summer bed roughness.

4.2. Validation and uncertainty of the lower-fidelity model

The model output was compared with the model output of the high-fidelity model to study whether it is justified to use the lower-fidelity model to perform the sensitivity analysis. We found that the high-fidelity model simulates a maximum discharge at Lobith of 12,282 m³/s with the 1926 measured discharge wave at Andernach as upstream boundary condition. The lower-fidelity model, with all random input parameters set to their expected value, predicts a maximum discharge of 12,402 m³/s. This deviates less than 1.0% compared to the high-fidelity model. Although, correct prediction of the maximum discharge at Lobith has the focus in this study, it is also desirable that the lower-

Table 5

Results Multiple Linear Regression analysis in which the most influential parameter has a ranking equal to 1 and the most non-influential parameter a ranking equal to 9.

Input parameter	β_i	σ_i [m ³ /s]	$\hat{\beta}_i$	Ranking	Surface area [%]	
Roughness class	Summer bed	-3.65×10^3	1.97×10^{-3}	0.15	2	13.3
	Lakes	-1.81×10^3	2.68×10^{-3}	0.10	5	13.2
	Grasslands	-4.81×10^3	8.71×10^{-3}	0.86	1	55.6
	Forest	-2.83×10^2	1.95×10^{-2}	0.11	4	6.4
	Urban areas	-2.29×10^2	2.73×10^{-3}	0.01	7	11.4
Bed level	Winter bed	-70.3	3.18×10^{-2}	0.05	6	
	Summer bed (1)	1.2	7.00×10^{-2}	0.00	9	
	Summer bed (2)	27.5	0.25	0.13	3	
	Summer bed (3)	8.3	0.04	0.01	8	

fidelity predicts correct discharge stages at other locations. Table 4 shows that the lower-fidelity model predicts maximum discharges along the Lower Rhine with high accuracy, having a maximum deviation of 2.1% compared to the high-fidelity model. In addition, the lower-fidelity model is capable of accurately predicting the discharge partitioning along the Dutch Rhine river branches (Table 4). These values indicate that the surrogate model is capable of representing the system behaviour of the high-fidelity model. Therefore, no correction-function is needed to tune the model results of the lower-fidelity model. We can thus conclude that the lower-fidelity model can be treated as a high-fidelity model from now on. Hence, the sensitivity analysis can be performed with the 1D-2D coupled model.

The results of the uncertainty analysis show that the average maximum discharge at Lobith as a result of the Monte Carlo sample equals 12,424 m³/s. This value has a standard deviation of 49 m³/s caused by the uncertainty in the input parameters. This relatively low standard deviation shows that uncertainties in the input parameters only have a limited effect on the maximum discharge at Lobith during the 1926 flood event.

4.3. Sensitivity analysis

Table 6

Computed Sobol' indices with the method of Li and Mahadevan (2016) in which the most influential parameter has a ranking equal to 1 and the most non-influential parameter a ranking equal to 9.

Input parameter	S_i	Ranking	Surface area [%]	
Roughness class	Summer bed	0.10	2	13.3
	Lakes	0.01	7	13.2
	Grasslands	0.77	1	55.6
	Forest	0.05	5	6.4
	Urban areas	-0.03	9	11.4
Bed level	Winter bed	0.09	3	
	Summer bed (1)	0.01	8	
	Summer bed (2)	0.06	4	
	Summer bed (3)	0.03	6	

4.3.1. Multiple linear regression analysis

A multiple linear regression analysis was performed in which it was assumed that the model response as a result of the varying input parameters was linear. This is not the case since the model coefficient of determination R^2 (equation (4)) equals 0.81. This value means that the regression model is capable of explaining 81% of the variance of the

surrogate output. The remaining 19% is ignored by the regression model. However, Table 5 clearly shows that the roughness of grasslands highly influences the maximum discharge at Lobith because of its high sensitivity measure $\hat{\beta}_i$ (equation (3)). The high standardized regression coefficient of the roughness of grasslands can be explained by the fact that grassland is the most dominant land cover in the model domain with a surface area of 55.6% (Table 5). In addition, the uncertainty within the class itself is relatively large (Table 2) since grasslands most often have a higher roughness during summer periods due to growing season compared to the winter periods. Only the roughness of forest has a larger uncertainty range. However, the surface area covered by forest is much less (6.4%).

4.3.2. Sobol' indices

In the previous section it was shown that with the Multiple Linear Regression analysis only 81% of the variance of the surrogate model output could be explained. In order to check the results of the Multiple Linear Regression analysis, the Sobol' indices are computed. These indices are independent of model linearity. The results show that the roughness of grasslands is dominant with respect to influencing the uncertainty of the maximum discharge at Lobith (Table 6). This is in line with the results of the Multiple Linear Regression analysis.

If $\sum_{i=1}^r S_i = 1$, the variance of the model output is solely caused by the variance of the input parameters itself. In that case, there are no interactions between the different input parameters resulting in an increase in the variance of the model output. In other words, the model is additive. The results show that the first-order Sobol' indices are approximately 1 indicating that the model does not include any interactions of the input parameters.

In principle $\sum_{i=1}^r S_i$ cannot be larger than 1. In addition, the first-order Sobol' index computed for each uncertain input parameter cannot be lower than 0 (Saltelli et al., 2008). In this study, the computed $\sum_{i=1}^r S_i$ is slightly larger than 1 and the Sobol' index for the roughness of urban areas is smaller than 0. This is caused by the relatively little sample size of only 120 runs. To overcome this problem, we resampled the 120 runs as explained in Section 3.3. With this resampled data set, the 95% confidence intervals of the first-order Sobol' indices are computed (Fig. 9). Fig. 10 shows that the first-order Sobol' indices have converged after approximately 700 bootstrap resamples. This results in a data set of 700×120 model runs. The outcomes then show that the roughness of grasslands remains the most dominant input parameter.

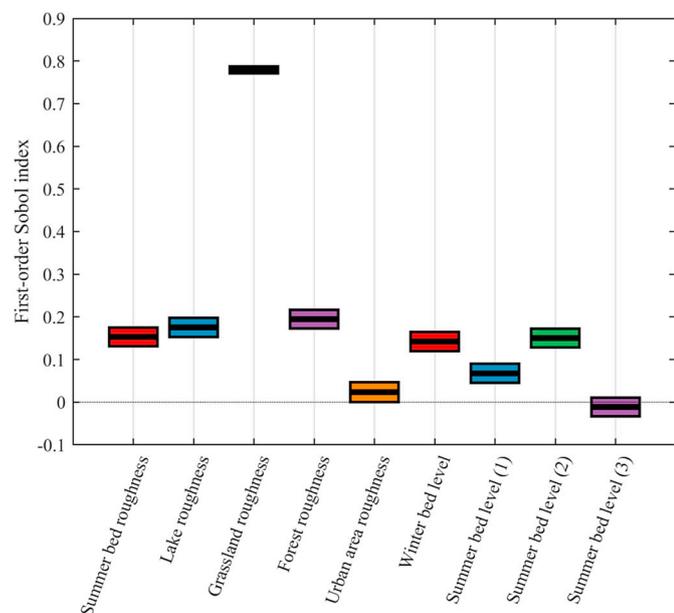


Fig. 9. First-order Sobol' indices and its 95% confidence intervals based on the bootstrap resamples.

The lower bound of its confidence interval is under any condition larger than the sensitivity index of the other input parameters. Therefore, we can conclude that for this specific case, most attention must be paid to the roughness class with the largest surface area and which has a relatively large uncertainty range. Correct prediction of this parameter will result in a significant reduction of the output variance. It must be noted that the uncertainty of the model output was small in this study. In general, the output variance depends on the probability distribution functions of the uncertain input parameters. It can be expected that the output variance will increase for older historic events. Hence, a significant reduction in model output variance can be reached if the most influential input parameter is correctly predicted. This influential input parameter can be found by applying the method for factor prioritization as presented in this study.

5. Discussion

In this study, a methodology was developed to reconstruct historic flood events with the use of a lower-fidelity model. The maximum discharge is predicted as well as its uncertainty as a result of the uncertain input parameter. General problems that arose were mostly related to the choice of the surrogate model type and the characteristics of the flood event. Therefore, another historic event may ask for a different approach since the assumptions made for the 1926 event may not apply. To put things into perspective, an overview and discussion are presented of the problems that may arise during historic flood reconstruction and resulting sensitivity analysis.

1. To predict a historic discharge, an associated geometry should be reconstructed. The geometry during the 1926 event was well known since maps of this time period are available. However, for events further in the past the geometry might be more uncertain. These spatial uncertainties must be included in the analyses. A major drawback is that for each (uncertain) geometric situation a separate model must be set up. Consequently, for each model, the sensitivity analysis must be performed separately. This significantly increases the total number of simulations. Furthermore, for older events the uncertainties in the input parameters may become larger. Hence, the shape of their probability distributions may change. We assumed that the uncertain input parameters of the 1926 flood event could be described by truncated normal distributions. These distributions will shift towards uniform distributions for older events if less information is available.
2. A lower-fidelity based surrogate model was developed to reduce computational time. Many other methods exist to set up a surrogate model, each with their own benefits and drawbacks. A different study approach may lead to the need of another type of surrogate model. In general, a 1D-2D coupled model is capable of simulating any kind of flood event. The 1D profiles enable correct prediction of discharge stages below bankfull conditions (Horritt and Bates, 2002). These 1D profiles can be coupled by 2D grid cells to include the possibility of simulating overland flows if the discharge exceeds the bankfull discharge, referring to the situation in which the discharge is larger than the main channel and floodplain capacity. Therefore, this type of lower-fidelity model can be used to accurately simulate flood wave propagation for both discharges below as well as above bankfull conditions.
3. The 1D-2D coupled model was not calibrated on maximum water levels. The objective of the surrogate model was accurate prediction of maximum discharges at Lobith. However, calibration on maximum water levels is required if dike breaches and/or overtopping have evolved during the flood event. For such a case, correct prediction of maximum water levels becomes important since this value indicates whether overtopping occurs. This influences the maximum discharge further downstream. Therefore, it is recommended to use the summer bed roughness of the lower-fidelity model as calibration

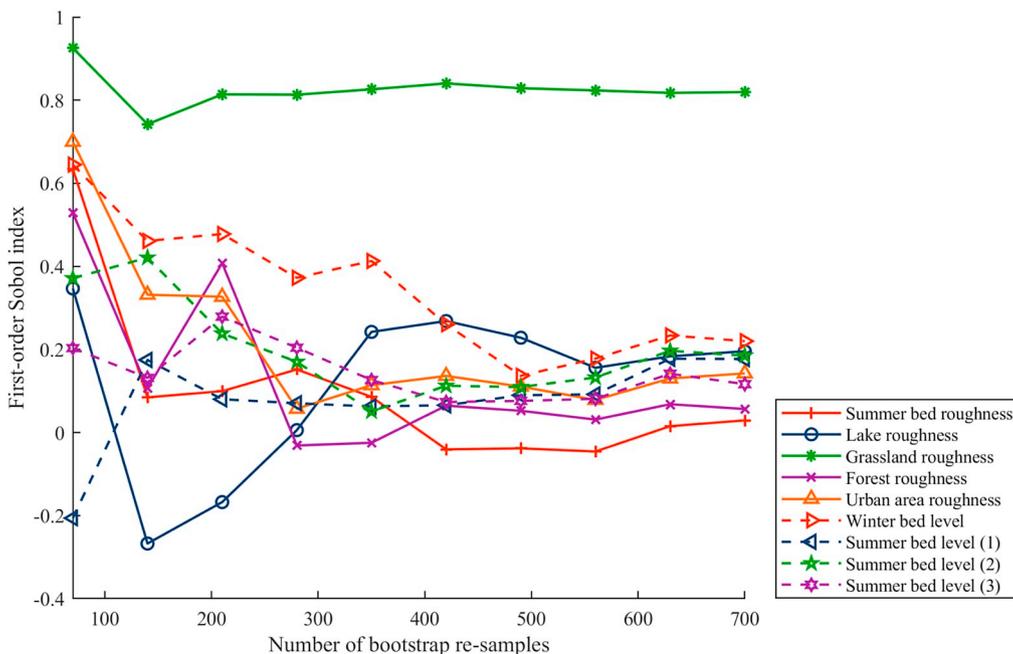


Fig. 10. Convergence of the first-order Sobol' indices based on the bootstrap resamples.

parameter to correctly predict water levels in case of discharges exceeding bankfull conditions.

4. To perform the sensitivity analysis, a decision had to be made about the range of the truncated normal distributions of the input parameters. The ranges of the roughness parameters were based on the tables of Chow (1959). A smooth channel with no vegetation was assumed to determine the roughness of the summer bed. This results in a relatively low expected Mannings roughness value of 0.028, with a total range of between 0.023 and 0.033. It is expected that the dimensions of sand dunes during flood events are highly uncertain. This uncertainty may influence summer bed roughness significantly. The measured Mannings roughness of the summer bed during the 1998 event with a maximum discharge of 9464 m³/s at Lobith ranges of between 0.030 and 0.035 (Julien et al., 2002). These values are higher than the values that we used. Paarlberg et al. (2010) found a clear dependency between increase in the discharge and increase in the dune heights. However, it is still unclear to what extent dune heights increase during flood events. Some literature even suggest that the dunes are washed out under extreme conditions (e.g. Best (2005) and Naqshband et al. (2014)), resulting in much lower values of the roughness parameter. It is not the roughness value itself that influences the uncertainty of the maximum discharge, but rather the uncertainty range of the summer bed roughness. Therefore, the relatively broad roughness range for the summer bed used in this study is considered appropriate for the 1926 flood event.
5. In this study, only geometrical uncertainties in the input parameters are included in the sensitivity analyses. These parameters are the bed levels of the summer bed and winter bed and the roughness of the various land use classes. However, much more uncertainties exist which can be related to the model structure, model parameters and boundary conditions. These inherent uncertainties can be considered in the sensitivity analysis by including them as random input parameters in the LHS. This will result in more insight in the most dominant type of uncertainty, i.e. uncertainty as a result of the input parameters, model parameters or model set-up. This study is recommended for future work since here, we only focused on the uncertainties of the geometrical input parameters to illustrate our method.

6. Conclusions

The objective of this paper was to study whether a lower-fidelity hydraulic model can be used for historic flood reconstruction. In this paper, a general framework is presented that shows which problems have to be tackled in order to enable historic flood reconstruction with the use of a surrogate model.

A 1D-2D coupled model was developed as lower-fidelity model that is capable of simulating flood wave propagation with high accuracy. It was found that model results predicted by the lower-fidelity model were close to those predicted by the high-fidelity model. The lower-fidelity model is thus capable of accurately predicting system behaviour. In addition, the proposed 1D-2D coupled model can be applied to any type of historic flood event. This is because it is capable of accurately simulating flood wave propagation for both discharges below as above bankfull conditions. However, if the simulated discharges exceed the bankfull discharge, model calibration is recommended since correct prediction of water levels becomes highly relevant for these cases.

A sensitivity analysis is required to determine the parameters that mostly influence the uncertainty in the model output. The lower-fidelity model could be used to perform this analysis. This significantly decreased computational time compared to the use of a fully 2D model. For future work, we propose that a 1D-2D coupled model can be treated as a high-fidelity model in general. Therefore, setting up a sophisticated 2D model for validation will not be needed.

The proposed methodology was tested with the use of the 1926 flood event of the Rhine river. The lower-fidelity model predicts a maximum discharge at Lobith of 12,402 m³/s for this historic event, deviating only 1.0% compared to the high-fidelity model (12,282 m³/s). The uncertainty of this maximum discharge at Lobith equals 49 m³/s. The uncertainty in model output is relatively small because a large amount of data of the 1926 flood event was available. Reconstruction of an older flood event will probably result in larger uncertainties of the input parameters since less information is available. As a result, the truncated normal distributions used to describe the uncertainty of the various input parameters will shift towards uniform distributions. This will have a negative effect on the model output uncertainty.

The sensitivity analysis showed that the model output was most sensitive to the roughness class with the largest share in surface area (in this case the roughness of the grassland areas). Moreover, the location

of the roughness class was important since areas close to the river have a relatively large impact on model results. These two aspects in combination with the uncertainty range of the input parameter itself determined the influence on model response.

Acknowledgement

This research is supported by the Netherlands Organisation for Scientific Research (NWO, project 14506) which is partly funded by the Ministry of Economic Affairs and Climate Policy. Furthermore, the research is supported by the Ministry of Infrastructure and Water Management and Deltares. This research has benefited from cooperation within the network of the Netherlands Centre for River studies NCR (www.ncr-web.org).

The authors would like to thank the Dutch Ministry of Infrastructure and Water Management and the German Federal Institute of Hydrology for providing the data. Besides, the authors would like to thank the following persons for their suggestions and valuable insights: Prof. Dr. Herget (University of Bonn), Dr. Aguilar-Lopez (Technical University Delft), Dr. Lammersen (Dutch Ministry of Infrastructure and Water Management), Van Doornik Msc. (Lieveense CSO), Berends Msc. (University of Twente) and Dr. Mara (University of La Réunion). In addition, the authors would like to thank the anonymous reviewer, Sheikholeslami Msc. (University of Saskatchewan) and the associate editor Dr. Razavi for their suggestions during the review process, which greatly improved the quality of the paper.

References

- Barriendos, M., Coeur, D., Lang, M., Llasat, M.C., Naullet, R., Lemaître, F., Barrera, A., 2003. Stationarity analysis of historical flood series in France and Spain. *Nat. Hazards Earth Syst. Sci.* 3, 583–592. <https://doi.org/10.5194/nhess-3-583-2003>.
- Best, J., 2005. The fluid dynamics of river dunes: a review and some future research directions. *J. Geophys. Res.* 110, 1–21. <https://doi.org/10.1029/2004JF000218>.
- Bladé, E., Gómez-Valentín, M., Dolz, J., Aragón-Hernández, J.L., Corestein, G., Sánchez-Juny, M., 2012. Integration of 1D and 2D finite volume schemes for computations of water flow in natural channels. *Adv. Water Resour.* 42, 17–29. <https://doi.org/10.1016/j.advwatres.2012.03.021>.
- Bomers, A., Schielen, R.M.J., Hulscher, S.J.M.H., 2019. The influence of grid shape and grid size on river modelling performance. In: *Environmental Fluid Mechanics*, <https://doi.org/10.1007/s10652-019-09670-4>.
- Brunner, G.W., 2014. Combined 1D and 2D Modeling with HEC-RAS. Technical Report. Hydrologic Engineering Centre (HEC) of the US Army Corps of Engineers. URL: <https://www.hec.gov/Modeling%20with%20HEC-RAS.pdf>.
- Caviedes-Voullième, D., García-Navarro, P., Murillo, J., 2012. Influence of mesh structure on 2D full shallow water equations and SCS Curve Number simulation of rainfall/runoff events. *J. Hydrol.* 448, 39–59. <https://doi.org/10.1016/j.jhydrol.2012.04.006>.
- Chow, V.T., 1959. *Open Channel Hydraulics*. McGraw-Hill Book Company, Inc, New York, USA.
- Deltares, 2016. *D-flow Flexible Mesh, User Manual. Version 1.2.1*. Technical Report. Deltares, Delft.
- Der Oberdeichinspektor Dusseldorf, 1895. *Deichbuch der Schau Huisberden. Deichbau im Niederrhein No. VII.10, Königliche Oberdeichinspektion Dusseldorf, Book provided by Deichverband Xanten-Kleve*.
- Dubreuil, S., Berveiller, M., Petitjean, F., Salaün, M., 2014. Construction of bootstrap confidence intervals on sensitivity indices computed by polynomial chaos expansion. *Reliab. Eng. Syst. Saf.* 121, 263–275. <https://doi.org/10.1016/j.ress.2013.09.011>.
- Dutch Ministry of Infrastructure and the Environment, 1926. *Verslag van het voorgavellene tijdens het hoge opperwater op de Nederlandsche rivieren in den winter van 1925 op 1926*. Technical Report Ministry of Infrastructure and the Environment. 's-Gravenhage URL: <https://repository.tudelft.nl/islandora/object/uuid:Sae4476-3245-42da-a9aa-5807f2b77b6e>.
- Dutch Ministry of Infrastructure and the Environment, 1952. *Afvoerkrommen 1948-1951 van de Bovenrijn en zijn Takken. Nota 51.23*. Technical Report. Rijkswaterstaat, Dir. Bovenrivieren, afd. Studiedienst. Anrhem, the Netherlands.
- Dutch Ministry of Infrastructure and the Environment, Ministry of Economic Affairs, 2014. *Delta Programme 2015 - Working on the Delta - the Decisions to Keep the Netherlands Safe and Liveable*. Technical Report. Ministry of Infrastructure and the Environment and Ministry of Economic Affairs. URL: <https://english.deltacommissaris.nl/delta-programme/documents/publications/2014/09/16/delta-programme-2015>.
- Efendiev, Y., Datta-Gupta, A., Ginting, V., Ma, X., Mallick, B., 2005. An efficient two-stage Markov chain Monte Carlo method for dynamic data integration. *Water Resour. Res.* 41, 1–6. <https://doi.org/10.1029/2004WR003764>.
- Efron, B., Tibshirani, R.J., 1993. *An Introduction to the Bootstrap*. Chapman & Hall/CRC, New York, USA. <http://projecteuclid.org/euclid.ss/1063994971>. 1011.1669v3.
- Gayton, N., Bourinet, J., Lemaire, M., 2003. CQ2RS: a new statistical approach to the response surface method for reliability analysis. *Struct. Saf.* 25, 99–121. [https://doi.org/10.1016/S0167-4730\(02\)00045-0](https://doi.org/10.1016/S0167-4730(02)00045-0).
- Herget, J., Kapala, A., Krell, M., Rustemeier, E., Simmer, C., Wyss, A., 2015. The millennium flood of July 1342 revisited. *Catena* 130, 82–94. <https://doi.org/10.1016/j.catena.2014.12.010>.
- Herget, J., Meurs, H., 2010. Reconstructing peak discharges for historic flood levels in the city of Cologne, Germany. *Glob. Planet. Chang.* 70, 108–116. <https://doi.org/10.1016/j.gloplacha.2009.11.011>.
- Horritt, M.S., Bates, P.D., 2002. Evaluation of 1D and 2D numerical models for predicting river flood inundation. *J. Hydrol.* 268, 87–99. [https://doi.org/10.1016/S0022-1694\(02\)00121-X](https://doi.org/10.1016/S0022-1694(02)00121-X).
- Janon, A., Nodet, M., Prieur, C., Janon, A., Nodet, M., Prieur, C., 2011. *Confidence Intervals for Sensitivity Indices Using Reduced-Basis Metamodels*. Technical Report. Technical Report. INRIA.
- Julien, P.Y., ASCE, M., Klaassen, G.J., Brinke, W.B.M.T., Wilbers, A.W.E., 2002. Case study: bed resistance of Rhine river during 1998 flood. *J. Hydraul. Eng.* 128. [https://doi.org/10.1061/\(ASCE\)/0733-9429\(2002\)128:12\(1042\)](https://doi.org/10.1061/(ASCE)/0733-9429(2002)128:12(1042)).
- Keating, E.H., Doherty, J., Vrugt, J.A., Kang, Q., 2010. Optimization and uncertainty assessment of strongly nonlinear groundwater models with high parameter dimensionality. *Water Resour. Res.* 46, 1–18. <https://doi.org/10.1029/2009WR008584>.
- Kernkamp, H.W.J., Van Dam, A., Stelling, G.S., De Goede, E.D., 2011. Efficient scheme for the shallow water equations on unstructured grids with application to the Continental Shelf. *Ocean Dynam.* 61, 1175–1188. <https://doi.org/10.1007/s10236-011-0423-6>.
- Lai, Y.G., 2010. Two-dimensional depth-averaged flow modeling with an unstructured hybrid mesh. *J. Hydraul. Eng.* 136, 12–23. [https://doi.org/10.1061/\(ASCE\)HY.1943-7900.0000134](https://doi.org/10.1061/(ASCE)HY.1943-7900.0000134).
- Leandro, J., Chen, A.S., Schumann, A., 2014. A 2D parallel diffusive wave model for floodplain inundation with variable time step (P-DWave). *J. Hydrol.* 517, 250–259. <https://doi.org/10.1016/j.jhydrol.2014.05.020>.
- Li, C., Mahadevan, S., 2016. An efficient modularized sample-based method to estimate the first-order Sobol' index. *Reliab. Eng. Syst. Saf.* 153, 110–121. <https://doi.org/10.1016/j.ress.2016.04.012>.
- Llasat, M.C., Barriendos, M., Barrera, A., Rigo, T., 2005. Floods in Catalonia (NE Spain) since the 14th century. Climatological and meteorological aspects from historical documentary sources and old instrumental records. *J. Hydrol.* 313, 32–47. <https://doi.org/10.1016/j.jhydrol.2005.02.004>.
- Maschler, T., Savic, D.A., 1999. *Simplification of Water Supply Network Models through Linearisation*. Technical Report. Centre of Water Systems, Report No.99/01. School of Engineering, University of Exeter, Exeter, United Kingdom.
- McPhee, J., Yeh, W.W.G., 2008. Groundwater management using model reduction via empirical orthogonal functions. *J. Water Resour. Plan. Manag.* 134, 161–170. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2008\)134:2\(161\)](https://doi.org/10.1061/(ASCE)0733-9496(2008)134:2(161)).
- Moussa, R., Bocquillon, C., 2009. On the use of the diffusive wave for modelling extreme flood events with overbank flow in the floodplain. *J. Hydrol.* 374, 116–135. <https://doi.org/10.1016/j.jhydrol.2009.06.006>.
- Moya Quiroga, V., Kure, S., Udo, K., Mano, A., 2016. Application of 2D numerical simulation for the analysis of the February 2014 Bolivian Amazonia flood: application of the new HEC-RAS version 5. *RIBAGUA - Revista Iberoamericana del Agua* 3, 25–33. <https://doi.org/10.1016/j.riba.2015.12.001>.
- Naqshband, S., Ribberink, J.S., Hulscher, S.J.M.H., 2014. Using both free surface effect and sediment transport mode parameters in defining the morphology of river dunes and their evolution to upper stage plane beds. *J. Hydraul. Eng.* 140. [https://doi.org/10.1061/\(ASCE\)HY.1943-7900.0000873](https://doi.org/10.1061/(ASCE)HY.1943-7900.0000873).
- Neppel, L., Renard, B., Lang, M., Ayril, P.a., Coeur, D., Gaume, E., Jacob, N., Payrastre, O., Pobanz, K., Vinet, F., 2010. Flood frequency analysis using historical data: accounting for random and systematic errors and systematic errors. *Hydrol. Sci. J.* 55, 192–208. <https://doi.org/10.1080/02626660903546092>.
- O'Connell, D.R.H., Ostenaar, D.A., Levisch, D.R., Klingner, R.E., 2002. Bayesian flood frequency analysis with paleohydrologic bound data. *Water Resour. Res.* 38, 1058–1071. <https://doi.org/10.1029/2000WR000028>.
- Paarlberg, A.J., Dohmen-janssen, C.M., Hulscher, S.J.M.H., Termes, P., Schielen, R., 2010. Modelling the effect of time-dependent river dune evolution on bed roughness and stage. *Earth Surf. Process. Landforms*. <https://doi.org/10.1002/esp.2074>.
- Pappenberger, F., Beven, K., Horritt, M., Blazkova, S., 2005. Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations. *J. Hydrol.* 302, 46–69. <https://doi.org/10.1016/j.jhydrol.2004.06.036>.
- Razavi, S., Tolson, B.A., Burn, D.H., 2012a. Numerical assessment of metamodelling strategies in computationally intensive optimization. *Environ. Model. Softw.* 34, 67–86. <https://doi.org/10.1016/j.envsoft.2011.09.010>.
- Razavi, S., Tolson, B.A., Burn, D.H., 2012b. Review of surrogate modeling in water resources. *Water Resour. Res.* 48. <https://doi.org/10.1029/2011WR011527>.
- Saltelli, A., Chan, K., Scott, E.M., 2000. *Sensitivity Analysis*. John Wiley & Sons, Ltd.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. *Global Sensitivity Analysis: the Primer*. John Wiley & Sons, Ltd.
- Scheidt, C., Li, L., Caers, J., 2018. *Quantifying Uncertainty in Subsurface Systems*, vol. 236. John Wiley & Sons, Ltd of Geophysical Monograph Series.
- Sheffer, N.A., Enzel, Y., Benito, G., Grodek, T., Poart, N., Lang, M., Naullet, R., Coeur, D., 2003. Paleofloods and historical floods of the ardeche river, France. *Water Resour. Res.* 39. <https://doi.org/10.1029/2003WR002468>.
- Silva, W., Klijn, F., Dijkman, J., 2001. *Room for the Rhine Branches in The Netherlands. What the Research Has Taught Us*. Technical Report. WL — Delft Hydraulics report R3294, Delft, The Netherlands.
- Simpson, T.W., Peplinski, J., Koch, P.N., Allen, J.K., 2001. Metamodels for computer-

- based engineering design: survey and recommendations. *Eng. Comput.* 17, 129–150. <https://doi.org/10.1007/PL00007198>.
- Tayefi, V., Lane, S.N., Hardy, R.J., Yu, D., 2007. A comparison of one- and two-dimensional approaches to modelling flood inundation over complex upland floodplains. *Hydrol. Process.* 21, 3190–3202. <https://doi.org/10.1002/hyp.6523>.
- The OpenDA Association, 2016. OpenDA User Documentation. Technical Report. . info@openda.org.
- Thokala, P., Martins, J.R., 2007. Variable-complexity optimization applied to airfoil design. *Eng. Optim.* 39, 271–286. <https://doi.org/10.1080/03052150601107976>.
- Toonen, W.H.J., Winkels, T.G., Cohen, K.M., Prins, M.A., Middelkoop, H., 2015. Lower Rhine historical flood magnitudes of the last 450 years reproduced from grain-size measurements of flood deposits using End Member Modelling. *Catena* 130, 69–81. <https://doi.org/10.1016/j.catena.2014.12.004>.
- Ulanicki, B., Zehnpfund, A., Martinez, F., 1996. Simplification of water distribution network models. In: *Proceedings of the 2nd International Conference on Hydroinformatics*, Rotterdam, The Netherlands, pp. 493–500. <https://doi.org/10.13140/RG.2.1.4340.8404>.
- Warmink, J.J., Straatsma, M.W., Huthoff, F., Booij, M.J., Hulscher, S.J., 2013. Uncertainty of design water levels due to combined bed form and vegetation roughness in the Dutch River Waal. *J. Flood Risk Manag.* 6, 302–318. <https://doi.org/10.1111/jfr3.12014>.
- Zhou, Y., Ma, Z., Wang, L., 2002. Chaotic dynamics of the flood series in the Huaihe River Basin for the last 500 years. *J. Hydraul. Eng.* 258, 100–110. [https://doi.org/10.1016/S0022-1694\(01\)00561-3](https://doi.org/10.1016/S0022-1694(01)00561-3).