Hydrology & Earth
System Sciences

A comparison of flood extent modelling approaches through constraining uncertainties on gauge data

M.G.F. Werner

Delft University of Technology, P.O. Box 5048, 2600 GA, Delft, The Netherlands / WL | Delft Hydraulics, P.O. Box 177, 2600 MH, Delft, The Netherlands

E-mail: Micha.Werner@wldelft.nl

Abstract

A comparison is made of 1D, 2D and integrated 1D-2D hydraulic models in predicting flood stages in a 17 km reach of the River Saar in Germany. The models perform comparably when calibrated against limited data available from a single gauge in the reach for three low to medium flood events. In validation against a larger event than those used in calibration, extrapolation with the 1D and particularly the integrated 1D-2D model is reliable, if uncertain, while the 2D model is unreliable. The difference stems from the way in which the models deal with flow in the main channel and in the floodplain and with turbulent momentum interchange between the two domains. The importance of using spatial calibration data for testing models giving spatial predictions is shown. Even simple binary (eye-witness) observations on the presence or absence of flooding in establishing a reliable model structure to predict flood extent can be very valuable.

Keywords: floods, hydraulic modelling, model calibration, uncertainty analysis

Introduction

A key element of flood risk analysis is quantification of flood hazard through estimation of the extent of flooding for design events of selected return periods. To derive flood extent maps for these, often of a larger magnitude than recorded events, various approaches can be applied. These models have a sound physical basis on the shallow water equations but, to represent, adequately, the flow characteristics in the reach under consideration, some form of calibration is required and two challenges have to be faced. Firstly, typically, few spatial data are available on the extent of flooding, with measurements perhaps only from reach scale gauges at intermediate distances of 10–30 km (Bates and De Roo, 2000). At best, there may be one or two maps of flood extent from recent events (Aronica *et al.*, 2002; Horritt and Bates, 2002; Romanowicz and Beven, 2003); on rare occasions distributed observations of flood levels in the floodplain may be available (Aronica *et al.*, 1998). Additional eye-witness accounts of flood extent or expert knowledge may be used (as shown in Franks *et al.*, 1998 for a distributed hydrological model). Secondly, design

events are mostly of greater magnitude than the observed events available for calibration, thus requiring extrapolation, which should always be treated carefully. This is most apparent in data-driven modelling, but holds also in application of conceptual and physical models. An example of the danger in extrapolation of a conceptual model is given in Uhlenbrook *et al.* (1999).

A simple approach to establishing the extent of flooding involves estimating a planar water level surface using observations of water levels at gauges. This planar surface is then crossed with a digital elevation model of the reach and areas with a depth greater than zero retained (Priestnall *et al.*, 2000). Although this may work well in short straight reaches where there are good height constraints on the planar surface, the extreme simplicity of the approach is reflected in its generally poor performance (Horritt and Bates, 2001). Application to events larger than observed requires reliable extrapolation of rating curves at the gauges, which in itself poses a problem. More commonly, some form of hydrodynamic modelling is used, including 1D hydrodynamic modelling and subsequent spatial

interpolation (Marche *et al.*, 1990; Péron *et al.*, 1996; Horritt and Bates, 2001; Werner, 2001; Tate *et al.*, 2002), 2D hydrodynamic modelling (Bates *et al.*, 1998; Beffa and Connell, 2001), cell inundation-based 2D modelling of the floodplain combined with kinematic or diffusive wave 1D modelling of the main channel (Cunge, 1975; Bates and De Roo, 2000) and, more recently, full hydrodynamic integrated 1D-2D codes (Verwey, 2001; Frank *et al.*, 2001).

In this paper, three of these modelling concepts were applied to predicting the extent of inundation on a 17 km reach of the River Saar in Germany. Floodplain and main channel roughness coefficients in the models were conditioned using several medium to large flood events within the generalised likelihood uncertainty estimation (GLUE, Beven and Binley, 1992) framework, and resulting posterior parameter distributions were subsequently applied in predictive mode to estimate flood stages during the large December 1993 event.

Study site and available data

The Schwemlinger Wiesen floodplain is a relatively wide floodplain on the Saar River, a tributary of the Mosel River, which is a tributary of the Rhine. The floodplain is located just upstream of the weir at Mettlach and downstream of the gauging station at Fremersdorf, an important flow gauge in the German part of the Saar (Fig. 1). The reach under consideration contains three characteristic sections. Directly downstream of the Fremersdorf gauging station, flow is constrained in the main channel with the floodplain constricted by a motorway on one side and a main road on the other. The middle section is the Schwemlinger Wiesen floodplain up to about 1 km. wide Extensive inundation on the floodplain during high flood events is constrained mostly by the motorway and other embankments. The last section is a narrow non-alluvial gorge containing an almost 180° river bend known as the Saarschleife. Here, the Saar cuts through the Rhenish slate mountains, and the section is characterised by a narrow main channel with very steep valley sides. The bed slope in this section is also relatively steep and the weir at Mettlach is used to maintain navigability. Head difference over the weir can be up to 14 m, so that downstream levels do not affect flow over the weir even at very high discharges, making the weir site ideal as a downstream boundary to a model. The rule curve for the upstream level at the weir is set to 167.5 m for low to moderate flows. For high flows this is reduced to 165.5 m to lower upstream flood levels. To increase navigability further, the main channel in the entire reach has been enlarged (WSV, 1987) and, hence, significant overbank flow occurs only for the largest events. For lower flows, backwater effects from the Mettlach weir prevail. Cross sections of the entire reach, surveyed after the enlarging of the main channel, are available at 100 m intervals. Given
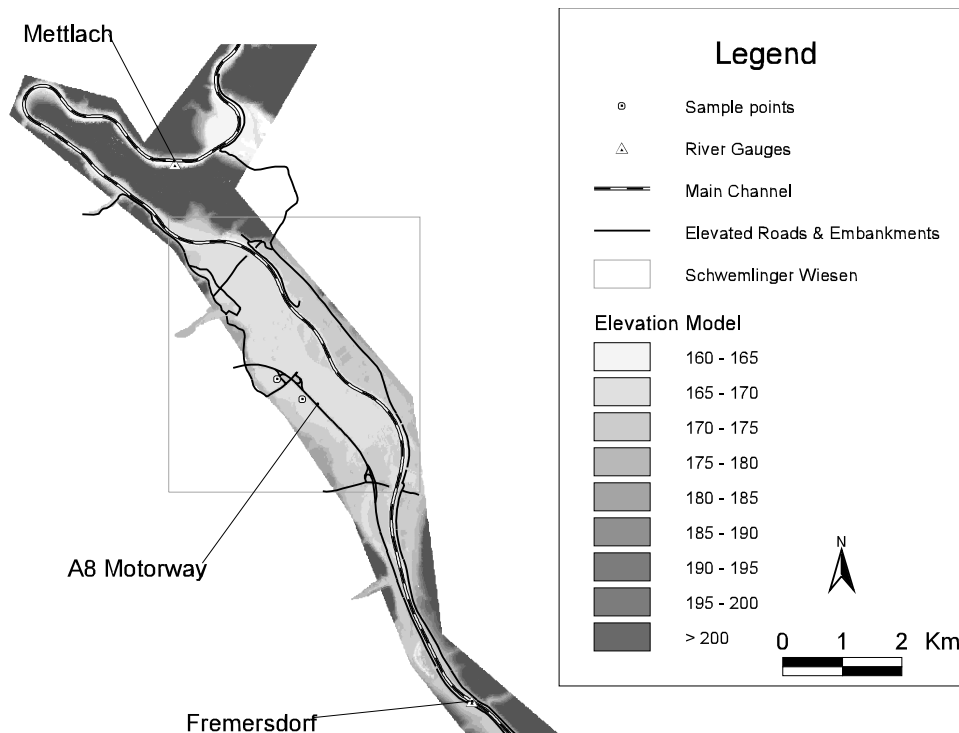

Fig. 1. *River Saar between the Fremersdorf and Mettlach gauging stations. Flow is from South to North.*
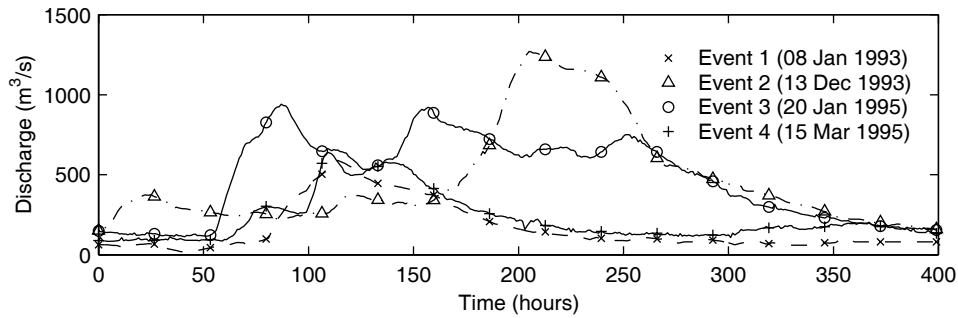
Fig. 2. *Flood hydrographs at the Fremersdorf gauging station.*

Table 1. Peak discharge and level at Fremersdorf.

	Unit	Event 1	Event 2	Event 3	Event 4
Start date		08-01-1993	13-12-1993	20-01-1995	15-03-1995
Q_{peak}	m³/s	589	1271	941	640
H_{peak}	m+N.N.	170.53	172.91	171.48	170.28

the large dimensions of the main channel with respect to the average flow in the Saar, as well as the non-alluvial sections in the lower half of the reach, the cross-section dimensions are considered to be relatively stable and representative for all flood events considered here. An elevation model derived for the floodplain through laser altimetry is also available. The original data, sampled at an average density of one observation per 2.5 m square, was processed to remove vegetation and errors.

Four flood events at the gauging station at Fremersdorf are shown in Fig. 2 and Table 1. The first and fourth of these events are of the order of the average annual flood. The January 1995 event, while being the second largest in the lower Rhine basin on record, has an estimated return period of only about 10 years at Fremersdorf. The December 1993 event is the largest in recent history in the Saar basin with an estimated return period at Fremersdorf of just under 100 years (Krahe *et al.*, 2000).

Approach

HYDRODYNAMIC MODELLING CODES

Three codes were chosen for modelling the extent of inundation in the Schwemlinger Wiesen floodplain;

- 1D hydrodynamic model code — SOBEK-RIVER, abbreviated here as *sbk*,
- 2D raster hydrodynamic code — DELFT-FLS, abbreviated here as *fls*

- Integrated 1D-2D hydrodynamic code — SOBEK-OVERLAND FLOW, abbreviated here as *s2d*

These three modelling concepts are part of the same SOBEK modelling family (see www.sobek.nl for details) and, although they share to some extent the same interface and are all capable of predicting flood extent, the difference between the codes lies primarily in the numerical approach and in the representation of floodplain and main channel flows.

One-dimensional hydrodynamic codes are applied widely for studying flood levels and discharges in river systems. The *sbk* modelling code applied here was developed predominantly for modelling river systems and estuaries. Water flow is solved using the full De Saint Venant equations, applying a Preissman-box scheme in the numerical solving routine. Main channel and floodplain geometry are given as cross-sections as a table of levels against width. Different roughness values can be applied to the main channel and in two floodplain sections. The cross-sections also distinguish between a conveyance section and a storage section to provide for large floodplain areas not contributing to channel flow but having a significant effect in flood wave attenuation. Output from the *sbk* model is in the form of levels at each cross-section calculation point along the main channel.

To obtain a flood extent map, the longitudinal profile of maximum levels is projected on to a 2D plane. On the assumption of constant water levels along each cross-
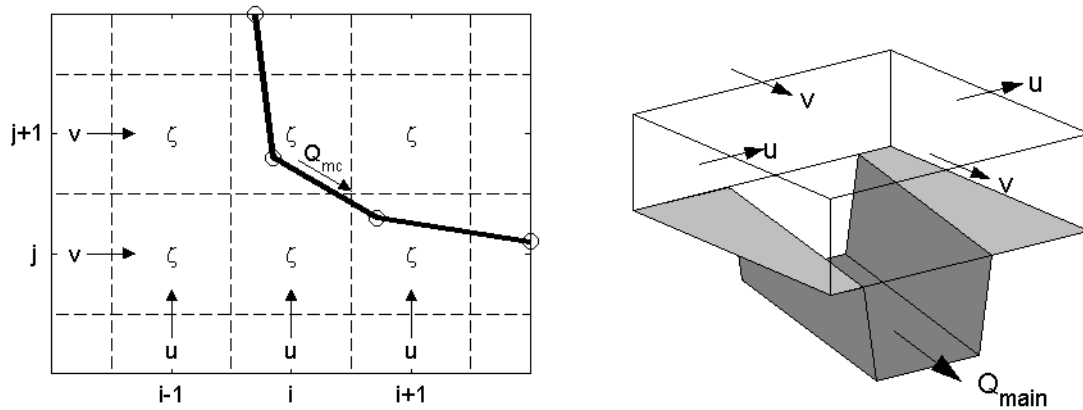
Fig. 3. *Impression of connection between 1D and 2D domains (after Frank* et al.*, 2001). Water level in grid cells i,j are indicated by z. u and v are fluxes across cell boundaries. The main channel is shown as a grey line with circles indicating 1D calculation points. These share the water level z of the cells they fall in.*

section, water levels are allocated to geo-referenced cross-section elevation points. These are then interpolated using an inverse distance weighting approach with orthogonal point selection to ensure a continuous surface which is then compared with the DEM of the floodplain and only depths greater than zero are retained. To ensure low lying areas (e.g. behind embankments) are not mistakenly considered flooded, all inundated cells unconnected to the main channel are removed from the final flood extent map. For a full description of the approach see Werner (2001).

The 2D raster based hydrodynamic code used, *fls* (described in full by Stelling and Duijnmeijer, 2003), numerically solves the shallow water equations discretely using a finite volume approach with a central water level variable at each grid cell centre and a flux across cell boundaries. This staggered grid approach is used also by other similar models (Beffa and Faeh, 1994; Beffa and Connell, 2001; Molinaro *et al.*, 1994). Modelling overland flow can be problematic due to the complex floodplain topography, with (local) steep slopes, shallow depths and wetting and drying problems (Beffa and Connell, 2001) and, in *fls,* a scheme that ensures positive water levels is used to ensure stability (Stelling and Duijnmeijer, 2003). An adaptive time stepping strategy ensures stability across the full range of depths and flow conditions, including sub- and super-critical flows and transitions.

Solving water flow on a regular grid has the advantage of easy integration with digital terrain models often available in a grid format. Although the *fls* model is developed principally for modelling dam break inundation events in e.g. low lying polder areas (Hesselink *et al.*, 2003), the code can be applied equally to modelling river flooding. Care must then be taken in modelling the main channel with a grid of sufficient resolution, particularly if the channel is

not being aligned with the grid, as this can result in serious over-prediction of stages (Beffa and Connell, 2001). Application of very high resolution grids, however, results in an intractable computational problem and, as the code does not allow for adaptive grids, a balance between accurate representation and run-times must be established.

The third code applied takes a different, hybrid, approach that allows integration of the model with generic GIS grid formats such as in the *fls* code, while solving dominant main channel flow as a 1D structure as in the *sbk* code. The main channel is modelled as a 1D element using cross-sections of the main channel only, while over-bank and floodplain flows are modelled in 2D using the staggered scheme in the *fls* code. Rather than a Preissman box-scheme for solving main channel flow, the same staggered approach is applied in the 1D domain. The two domains are coupled at grid cells overlying main channel computational points through mutual volumes of the connected cell and the adjoining channel section(s). Figure 3 gives an impression of the connection between the 1D and the 2D domain. The figure on the left shows in plan view the main channel running underneath the 2D floodplain grid. The 3D view shows an impression of the horizontal division between a single cell and the main channel. While momentum exchange between the two domains is not taken into account, explicitly, additional momentum loss for overbank flow is included conceptually through a non-zero shear interface at the horizontal division between the main channel and floodplain domains. The roughness value at this interface is defined to be the same as the main channel bed roughness, and allocated to the grid cells forming the horizontal division. The code has been demonstrated in several practical applications (Frank *et al.*, 2001; Verwey, 2001).

MODEL PARAMETERISATION AND CALIBRATION

To establish predictive uncertainties of each of the model concepts, the generalised likelihood uncertainty estimation approach (GLUE) was adopted (Beven and Binley, 1992). This Monte-Carlo based technique rejects the calibration paradigm of one single optimal parameter set, considering rather an ensemble of possible parameter sets. The ensemble is constructed through sampling of constituent parameter prior distributions and, for each ensemble member, the model is run and results are compared against observations. Based on this comparison, the acceptability of each parameter set is established in terms of a likelihood measure, expressing the comparative likelihood of that parameter set and model structure giving a good description of the behaviour of the system. The procedure is set in a Bayesian framework (Romanowicz and Beven, 2003), where likelihoods derived using a given calibration data set can be updated with those of subsequent data sets (Beven and Binley, 1992),

$$L_p\left(\Theta \mid Y\right) = \frac{L_o\left(\Theta\right)\ L\left(Y \mid \Theta\right)}{C} \tag{1}$$

where $L_o(\Theta)$ is the prior likelihood of parameter set Q, $L(Y \mid \Theta)$ is the likelihood value calculated for the current evaluation given the set of observations Y and parameter set Θ, $L_p(\Theta \mid Y)$ is the posterior likelihood, and C is a normalising constant to ensure that the cumulative posterior likelihood is unity. Optionally, a behavioural threshold can be applied to divide runs with acceptable performance from those with unacceptable performance (Beven and Binley, 1992). Parameter sets that perform below the threshold are given a likelihood of zero, and remaining likelihood values are re-scaled.

APPLICATION

The three codes selected were applied to the reach between Fremersdorf and Mettlach. All three make use of the digital terrain model, based on the available laser altimetry data, the. resolution of which is so high that application in the modelling codes would result in impractical run times, without necessarily adding to model accuracy. The elevation data were, therefore, first aggregated to reduce the computational burden, particularly relevant in the *fls* and *s2d* models. Aggregation was done in two steps. In the first step, an elevation model with a resolution of 10 m × 10 m was established through averaging laser altimetry points within each grid cell and taking the maximum for grid cells crossed by line elements (e.g. the motorway embankment). In the second step, two models with a resolution of 20 m

× 20 m were derived through further aggregation of the 10 m × 10 m elevation model. In the first, no account was taken of important line elements, and the averaging rule was used for all cells. In the second, line elements were again accounted for explicitly using the maximising rule for cells crossed.

In developing the 1D model, the surveyed cross-sections were found not to extend fully across the floodplain. These were completed using sampled elevations from the digital elevation model, and combined with the observed cross sections in the main channel (Werner *et al.*, 2000). The upstream boundary for the model was set as the observed discharge at the Fremersdorf gauging station and the downstream boundary as the water level upstream of the weir at Mettlach as defined by the weir rule curve. The extent of inundation was established by projecting the 1D result on the available digital elevation model.

Models using the *fls* code were established using the 20 m × 20 m resolution elevation models, with cross-section data in the main channel interpolated to complement floodplain elevation data. Of the three sections in the reach considered, flow in the first and third reaches is predominantly in-bank. Beffa and Connell (2001) found that using a relatively coarse grid with respect to the width of the flow could result in significant overestimation of stage, particularly if the dominant flow direction is not parallel to the grid orientation. To minimise this overestimation, the extent of the 2D model was reduced to cover only the second section, the Schwemlinger Wiesen floodplain itself, where flood extent is of interest. Reducing the extent of the 2D model also reduced the computational effort. For the small first section, the upstream discharge boundary was transposed to the upstream boundary of the grid, as lag time and attenuation in the section are negligible. Head losses in the first and last section were determined using a simple steady state approximation with the average main channel geometry in the reach, the downstream level at Mettlach and the (variable) main channel roughness as inputs (French, 1986, p.229). Comparing this approximation with the 1D model results in these sections showed little differences between the approximation and the 1D model and so the former was adopted for simplicity. One difficulty with introducing a second model concept for this first and third reach is that a degree of freedom is added, whether modelled though the approximation or through the 1D approach. This is the roughness of the main channel for these sections (floodplain flow is negligible and floodplain roughness need, therefore, not be considered). In applying the combined model, the degrees of freedom were, however, reduced to main channel and floodplain roughness only by assigning the same main channel roughness in all three sections.
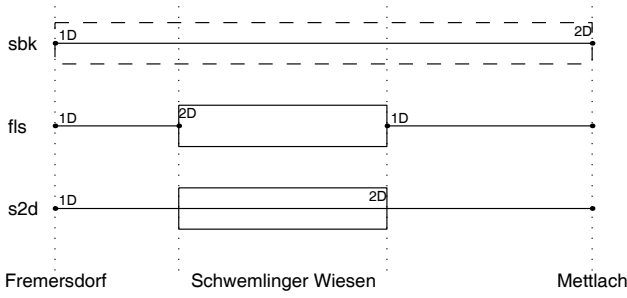
Fig. 4. *Schematic overview of three model structures showing 1D and 2D domains. The dotted line for the sbk model indicates the 2D domain is not included in the model but is obtained through projection of the 1D result.*

In the *s2d* model, the 1D modelling approach was applied to the main channel for the full reach (as in the *sbk* model) and the *2D* modelling approach was applied only for the second section of the Schwemlinger Wiesen floodplain as in the *fls* model. Cross-sections in the first and third sections were defined to cover both floodplain and main channel, while those in the second section were defined to cover only the main channel as floodplain flow is accounted for in the 2D domain. The elevation model along the main channel, being the interface between the 1D and 2D domains was set to the bankfull level. Figure 4 shows, schematically, the setup of 1D and 2D domains for each of the three codes.

In each of the three modelling codes two roughness values were defined. One was for the main channel and one for the floodplain, with the main channel roughness value also being applied to the non-zero shear interface in the *s2d* model. Although the morphology of the last section differs from that of the first two where the river runs in an alluvial bed, further disaggregation of roughness was not considered. This would limit the identifiability of parameters further, given that a low roughness in the first two sections could then compensate for a high roughness in the third and vice versa. The prior ranges from which parameter sets were sampled are described in Table 2, with ranges selected on the basis of tables in literature (Chow, 1959) and similar studies (Horritt and Bates, 2001). From these parameter prior distributions, Monte Carlo sampling was applied to construct parameter sets. The sample size for the *sbk* model, with its

Table 2: Prior parameter ranges for main channel and floodplain roughness.

Parameter	Range
Main channel roughness (n_{mc} (m$^{1/3}$s^{-1}))	0.02 - 0.06
Floodplain roughness (n_{fp} (m$^{1/3}$s^{-1}))	0.02 - 0.20

relatively short run-times, was set at 1000, while for the *fls* and *s2d* models only 100 parameter sets were used due to computational constraints in using these models (run-time in the order of 4–6 hours). These were, however, sampled using a Latin-Hypercube sampling strategy (Mckay *et al.*, 1979; Yu *et al.*, 2001) to ensure a good spread of samples in the parameter space. Comparison of runs using the *sbk* model using Latin Hypercube sampling with the full random sample of 1000 showed that parameters of posterior distributions stabilised between 50 and 100 runs, comparing well with results from Yu *et al.* (2001).

For each parameter set sampled, the model was run and resulting levels at Fremersdorf were compared. To establish likelihood values within the context of the GLUE analysis using Eqn. 1, a suitable objective function was required. The objective function selected compares the maximum level calculated at Fremersdorf with the maximum level observed:

$$L_i = \exp\left(-\frac{|\hat{h}_i^{\max} - h^{\max}|}{T} \right) \qquad (2)$$

where L_i is the likelihood measure for parameter set *i*, $\hat{h}_i^{\max}$ is the calculated maximum level at Fremersdorf and $h^{\max}$ the observed maximum as given in Table 1. The constant T determines the dispersion of the modelling error in determining parameter likelihoods. As the selected function is exponential, this constant plays an important role in determining relative likelihoods between parameter sets, and the value of T can be considered to reflect a relative belief in the model. For the comparison of model runs here the constant was selected at 0.20 m. This is of the same order of magnitude as the uncertainty in laser altimetry-derived elevation models (Marks and Bates, 2000). Although somewhat arbitrary, it serves the purpose of demonstration.

The advantage of the exponential form of the likelihood function is that updating of likelihoods from different data sets implies simple additive combination, with all observations treated as equally important (Lamb *et al.*, 1998),

$$L(\theta_i \mid Y_{1,\ldots,n}) = \frac{1}{C} \exp\left[-\left(\frac{|\hat{h}_i^{\max,1} - h^{\max,1}|}{T} + \ldots + \frac{|\hat{h}_i^{\max,n} - h^{\max,n}|}{T} \right) \right] \qquad (3)$$

where $\hat{h}_i^{\max,n}$ is the maximum level at Fremersdorf calculated for parameter set *i* for event *n and* $h^{\max,n}$ is the maximum observed level at Fremersdorf for event *n*. C is a constant to normalise likelihood values, and $L(\theta_i \mid Y_{1,\ldots,n})$ is the likelihood value of parameter set θ_i constrained on datasets $Y_{1,\ldots,n}$. Although no spatial observation data for the extent of flooding are available, the knowledge that the motorway was not overtopped in any of these events was used to

Table 3. Prior parameter ranges for main channel and floodplain roughness.

Run name	Event	Prior likelihood	Models	Comment
run a	Event 1	[-]	*sbk*	likelihoods updated
run b	Event 4	run a	*sbk*	likelihoods updated
run c	Event 3	run b	*sbk*	likelihoods updated
run d	Event 3	[-]	*sbk,fls,s2d*	likelihoods updated
run e	Event 2	[-]	*sbk,fls,s2d*	likelihoods updated
run f	Event 2	run d	*sbk,fls,s2d*	likelihoods updated
run g	Event 2	run d	*sbk,fls,s2d*	Predictive mode

constrain likelihood values further. A selection was made between parameter sets on the basis of an indication of flooding behind the motorway (sample points in Fig. 1). Parameter sets indicating flooding behind the motorway were assigned zero likelihood and discarded, while those not indicating flooding were retained, following which the likelihood values were re-normalised. Equation 2 was modified slightly by multiplying with the dimensionless operator *P*. The value of *P* is 0 for runs indicating flooding at either sample location behind the motorway and 1 for all runs where flooding behind the motorway is absent.

$$L_i = P \times \exp\left(-\frac{|\hat{h}_i^{\max} - h^{\max}|}{T}\right) \qquad (4)$$

Likelihoods for parameter sets can be used in obtaining prediction quantiles for the maximum water levels using Eqn. 5.

$$p(\hat{h}^{\max} < h) = \sum_i^N L(\theta_i) \quad \forall \; \hat{h}_i^{\max} < h \qquad (5)$$

where *p* is the probability that the maximum level calculated at Fremersdorf $\hat{h}^{\max}$ exceeds the level *h*. Likelihood values are summed only for those parameter sets where the calculated level does not exceed *h*. *N* is the number of members in the parameter set ensemble. In predictive mode, quantiles can be determined for an event by using likelihood values constrained using data only from other events, as opposed to after updating using the likelihood values with the data of the event itself. To compare predictive uncertainties of the different modelling concepts, the runs described in Table 3 were carried out with the sampled parameter sets and models. In *run a,* parameter distributions were conditioned on event 1; in *run b,* they were conditioned on events 1 and 4, while in *run c* both events 1, 4, and 3 were used. These three runs were effected using only the 1D modelling approach. In *run d,* parameter uncertainties were conditioned on event 3 only, while in *run e* only event 2 was applied. In *run f,* both the large events 2 and 3 were used to condition parameter distributions. Finally, in *run g*

the parameter distributions conditioned on event 3 were used for predicting levels for event 2 (verification stage). *Runs d* to *g* were carried out for all three of the selected modelling concepts. These last three runs for the extreme December 1993 event (event 2) were used to compare the merits of updating parameter likelihoods using different sets of calibration data. All runs using the *sbk* model applied all 1000 sampled parameter sets (run time in the order of 30 seconds) while each *s2d* and *fls* run used the 100 sampled parameter sets (run time in the order of 4 to 6 hours).

Results and discussion

IMPACT OF ELEVATION MODEL QUALITY

The three elevation models aggregated from the original laser altimetry data were run with the *sbk* model to evaluate the quality of the aggregated elevation data. Figure 5(a) shows selected quantiles for predicted maximum level at Fremersdorf for *run e* using the 10 m × 10 m elevation model (*dem10*), the 20 m × 20 m elevation model where no account of line elements is taken (*dem20*) and the 20 m × 20 m elevation model where line elements are included explicitly (*dem20**). These quantiles were determined using both Eqns. 2 and 4 (see figure legend). The figure shows that when conditioning on the level at Fremersdorf using Eqn. 2 only, there was little difference between the elevation models. Application of Eqn. 4 changes this significantly for the model using the *dem20* elevation model as most model runs giving a reasonable estimate of the maximum level at Fremersdorf are now rejected due to flooding behind the motorway. This suggests that the model structure using this elevation model is unsuitable for estimating flood extent in this reach. Figure 8(b) shows the substantial difference between the two elevation models in the extent of flooding for pixels inundated in 95% of the runs, with the area behind the motorway being completely inundated when applying the flawed elevation model. While it is obvious a line element such as a motorway should be considered explicitly
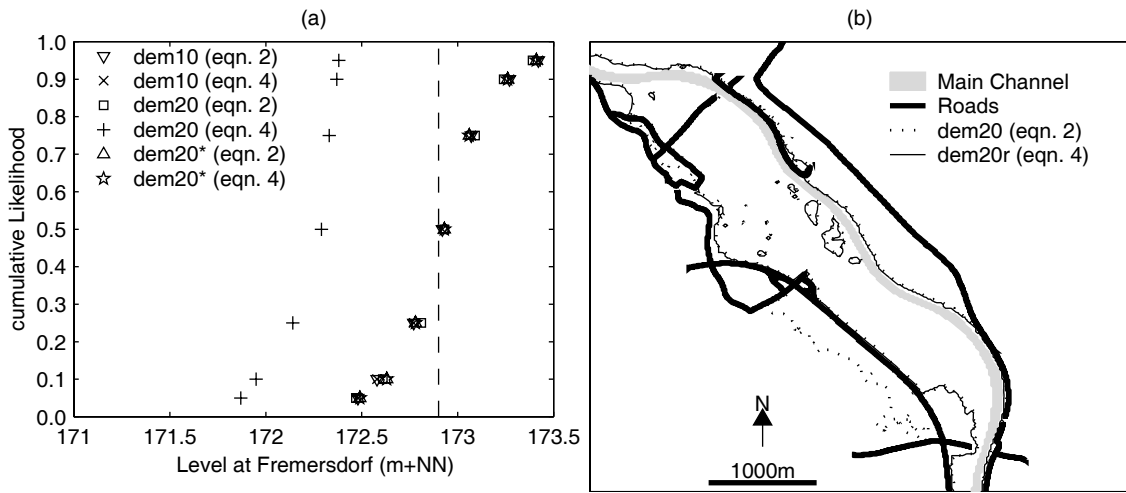
Fig. 5. *Cumulative likelihood for peak level predictions at Fremersdorf, (a) results for sbk model using DEM's of differing quality and resolution, (b) 95 % flooded area percentile.*

in the elevation data, smaller, or particularly thinner, features such as flood walls may not be well represented in the elevation data yet have a significant impact on the extent of inundation. For reliable evaluation of model structure for spatial prediction such as flood extent, it is paramount that even simple binary observations are included. Little difference was seen between the performance of models using the 10 m × 10 m and those using the coarser 20 m × 20 m elevation model (*dem20**) where line elements were considered. The latter was, therefore, used in all subsequent runs.

EFFECT OF UPDATING LIKELIHOOD VALUES
USING MULTIPLE DATA SETS

The value of updating posterior likelihood values using

multiple flood events was explored using runs a-d (Table 3) with the 1D *sbk* model. Figure 6 shows the cumulative likelihood distributions for main channel and floodplain roughness for successive updates from each of the events 1,4 and 3 (runs a to c), as well as if only Event 3 is used in updating likelihoods (run d). There is little difference between using all three events in updating likelihood values and using only the larger event 3. Both events 1 and 4 are relatively small, with limited floodplain flow. The flow conditions are, moreover, dominated by backwater effects from the weir at Mettlach and so these two events have limited value in identifying the posterior main channel roughness distribution, and serve only to reduce the dispersion of the roughness values. For event 3, where flows are higher, the backwater effects from the weir at Mettlach are less important and the posterior dispersion of main
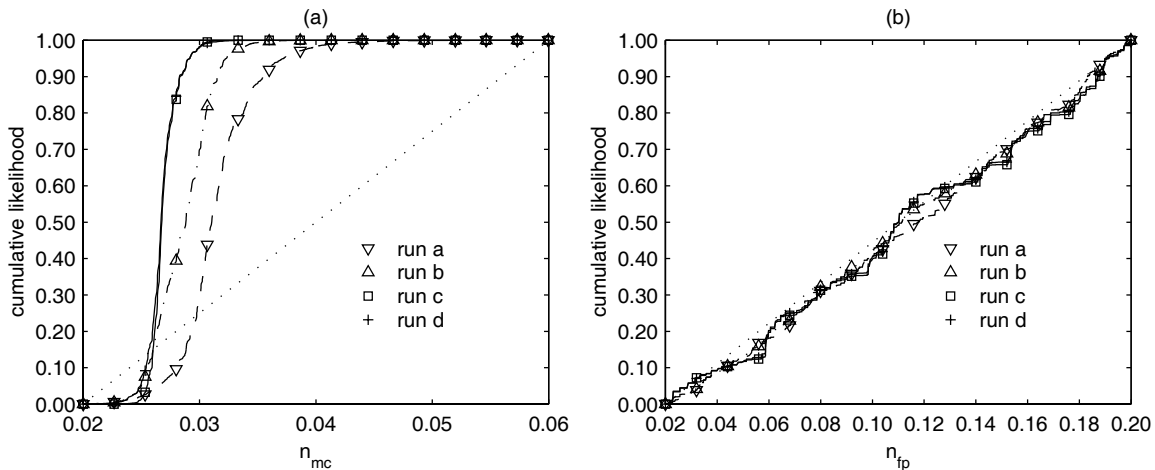

Fig. 6. *Joint posterior cumulative distribution functions for roughness values in the sbk model conditioned using different events (runs a-d). (a) shows the distribution for main channel roughness and (b) the distribution for floodplain roughness.*

channel roughness is reduced further. As events 1 and 4 were of relatively little use in determining the main channel roughness values, the information held in event 3 dominates, and the posterior distributions found for run c and run d are almost identical (Fig. 6a). In all events used here floodplain flow was very limited, with floodplain inundation being significant only for Event 3. Even then the floodplain contributes more to storage than to flow and floodplain roughness remains insensitive as a consequence. Given the limited information content in constraining main channel roughness values, events 1 and 4 were not applied in the *fls* and *s2d*, with predictive uncertainties of the different modelling concepts established using events 2 and 3 only.

COMPARISON OF POSTERIOR ROUGHNESS DISTRIBUTIONS OF MODELLING CONCEPTS

Figure 7(a) and (b) show selected quantiles for the three different modelling concepts as conditioned on the events 3 and 2 respectively, using uniform prior likelihood distributions for the sampled parameter sets. The 50% quantile matches the observed value very well in each case, which is not surprising as Eqn. 4, used to establish likelihood values, has a strong tendency to penalise runs where the maximum predicted level deviates from that observed. Other quantiles compare equally well, with only the *s2d* model having a slightly wider uncertainty envelope for *run d. Run e* shows similar results, suggesting the three modelling concepts are equally suitable. Combining the likelihood values for the two events in *run f*, however, shows that both the *sbk* and *fls* models tend to under-predict the maximum level at Fremersdorf, while the *s2d* model shows good

performance with the uncertainty envelope narrower than that of *run e*. The under-prediction is strengthened when the posterior parameter distributions found for Event 3 are used in verification for Event 2 (*run g*). In particular, under-prediction by the *fls* model is significant, confirming that this model is unsuitable for resolving flow where flow is dominated in the main channel described by only a few computation cells in the model (Beffa and Connell, 2001).

Figure 8 shows dotty plots for the main channel roughness values for runs d, e, and f. These are shown for the 1D model (*sbk*) in Fig. 8a and for the hybrid 1D-2D model (*s2d*) in Fig. 8b. The well constrained nature of the main channel roughness in the 1D model (Fig. 8a) would support the selection of this value as an optimal calibration. However, it can be clearly seen that on independently calibrating using Event 2 (*run e*) rather than Event 3 (*run d*), this optimum shifts to a higher roughness. This supports results from Romanowicz and Beven (2003), who found that the distribution of effective parameters may change when evaluated on flood waves of different magnitude. The shift to the right is as would be expected, as floodplain flow is more significant in Event 2. This leads to an increase in momentum interchange between main channel and floodplain. Momentum interaction is not accounted for explicitly in the 1D approach and calibrated 'effective' roughness values represent the lumped effect of bed roughness, momentum interchange between main channel and floodplain and crossing flows (Knight and Shiono, 1996). As floodplain flow becomes more significant, a higher value for 'effective' main channel roughness is required. Levels for Event 2 predicted using the posterior parameter probabilities found for Event 3 (*run d*) are under-
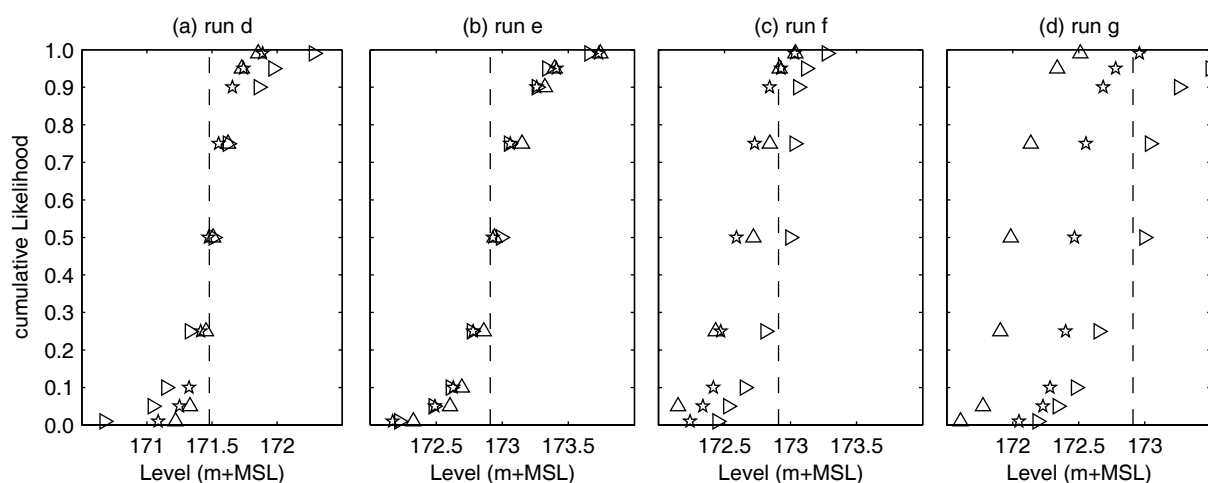

Fig. 7 *Quantiles for peak level predictions using different model concepts. (a) shows the predictions for event 3 conditioned on event 3 (run d), (b) shows the predictions for event 2 conditioned on event 2 (run e), and (c) shows predictions for event 2 conditioned on events 2 and 3 (run f). (d) shows predictions for event 2 conditioned on event 3 only*
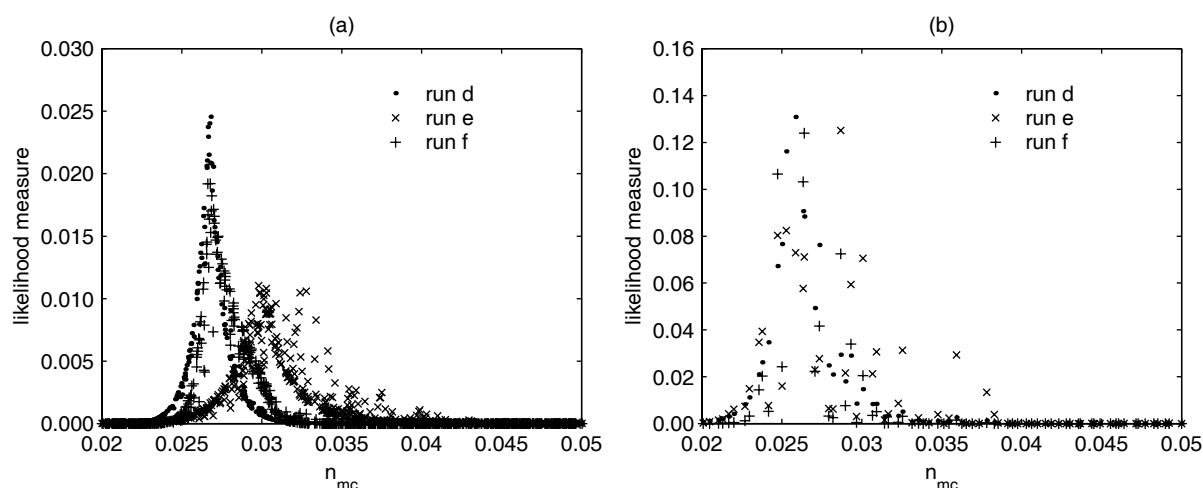
Fig. 8. *Dotty plots for main channel roughness distributions found for run d, e, and f. The x-axis shows the range of the main channel roughness values and the y-axis the likelihood value. Each dot represents the main channel roughness for one of the selected parameter sets. (a) shows the likelihood values found for the three runs using the sbk model. (b) shows these for the s2d model.*

predicted as a result (Fig. 7d). Figure 8a also shows multiple optima for the main channel roughness for Event 2 (*run e*), which are absent in Event 3 (*run d*). A possible explanation is that as the floodplain flow increases, so does the ability of the floodplain friction values to compensate for excessively high main channel roughness values. The parameter response surface is more complex as a result of the dependence between main channel and floodplain roughness values, and multiple optima can exist.

Despite the relatively low number of runs for the *s2d* model, the posterior distribution of main channel roughness values shows the same tendency as in the *sbk* model, with a slight shift towards rougher values for posterior distributions updated using event 2 (*run e*). This shift is, however, less marked than for the 1D model, as confirmed in Fig. 7d where predicted levels for Event 2 using the posterior parameter probabilities found for Event 3 (*run g*) show much less deviation from the observed value. The smaller change in effective roughness parameters is attributed mainly to the structure of the model. Although not explicitly considered as a turbulence closure model, the non-zero shear at the interface between the 1D and 2D domains, allows for momentum exchange due to interaction between flow in the main channel and floodplain, thus giving a better description of the non-linear stage-discharge relationship.

To account for change in 'effective' roughness values in the 1D approach, a stage dependent main channel roughness parameterisation could be introduced (such as the discharge adjustment factor proposed in Ackers (1992)). Although this would indeed improve the reliability of the 1D model, it also results in an additional parameter, making the identification problem even more difficult. Moreover, for

the reach in question, main channel roughness is less sensitive at lower flows (Fig. 8) due to backwater effects, leaving very little data to estimate reliably parameters of a stage-dependent main channel roughness value. For even higher events, the impact of momentum interchange of the total channel flow becomes theoretically less as the floodplain and main channel start to act as a single channel (Ackers, 1992). The optimum for the 1D model would be expected then to shift to the left. Similar results for applicability of 1D modelling from Ali (2002) show that, for the case they studied, single channel behaviour was attained at a depth in the order of twice the bankfull depth. This factor is, however, dependent on channel and floodplain geometry and relative roughness values, making identification in a natural channel very difficult. Extrapolation in predictive mode for high magnitude design events of some empirical relationship derived on limited data then becomes very uncertain.

Conclusions

This paper has presented results of a comparison between different modelling approaches through constraining parameter uncertainties on gauge data from events of differing magnitude. In setting up these models, it was shown that even for the most widely applied 1D hydrodynamic approach, the establishment of a suitable representation of floodplain topography is of prime importance. In particular, it is important to ensure representation of line elements in the floodplain that can constrain the extent of flooding, even for larger events. Despite this, models using low quality elevation data perform as well as models using high quality

elevation data when calibrated against only bulk flow data from river gauges. As the required model output is the spatially represented flood extent, use of spatial observations in establishing model suitability is paramount. Even simple (eye witness) observations such as presence or lack of flooding behind embankments can be invaluable in identifying flaws in the elevation model. Applying these in calibrating flood extent models rather than relying only on gauge data could greatly improve the reliability of flood extent modelling.

Three codes for modelling flood extent were selected; a 1D hydrodynamic code with subsequent projection of water levels to the 2D domain through interpolation, a 2D raster hydrodynamic code, and an integrated 1D-2D code. Given the availability of a suitable digital elevation model, the three codes are shown to be equally suitable in predicting levels at the upstream gauging location for the events against which the codes are calibrated. However, setting the calibration within the GLUE uncertainty framework shows that extrapolation of models calibrated using medium and low events can result in unreliable results. When validating the 2D model against maximum stages for the December 1993 event in the River Saar at Fremersdorf, these are severely underestimated if main channel and floodplain roughness parameters were constrained using the January 1995 event only. Even for the integrated 1D-2D approach, the 90% confidence range for the maximum level at the Fremersdorf gauging station in the 1993 event increases from 58 cm when predicted using roughness values constrained on both the 1995 and 1993 events to 115 cm when these are constrained on the 1995 event alone. The December event was used here as a proxy to a design event, as these are often of a magnitude higher than events available for calibration, and this range would no doubt increase for design events of even higher magnitude.

The differences between the codes come primarily from how these resolve floodplain and main channel flow and the interaction between these two domains. Despite its easy integration with available GIS data formats, the 2D raster-based hydrodynamic model is unsuitable for the type of river reach considered here as it is unable, properly, to capture the dominant main channel flow without resorting to a grid resolution that would result in an intractable computational problem. For the 1D code, the simplicity with which the main channel–floodplain interaction is described results in the over-estimation of conveyance for low over-bank flow. Roughness parameters used in the model describe not only the momentum loss due to bed shear, but include a lumped representation of losses due to momentum interchange in the main channel due to secondary currents and crossing flows. Although depth-dependent roughness relationships

can be introduced to account for this such as described by Ackers (1992), identifying coefficients of this relationship can be difficult in a real river given the lack of data. Useful stage data in determining the relationship may be limited further if, in low and medium flood events, the flow regime in the river is influenced by backwater effects from structures. The ability of the integrated 1D-2D code to extrapolate reliably in this case is due to its conceptual consideration of the momentum exchange between floodplain and main channel, thus allowing a better representation of the actual stage-discharge curve and, hence, more reliable extrapolation to larger events.

A limitation of this study in assessing uncertainties in roughness parameters is that data were used from only a single gauging station at the upstream end of the reach for all but the models using the flawed elevation model. As a result, good simulations in calibration are obtained easily using crude topography and spatially lumped roughness parameters. If more independent level observations were available and used to calculate likelihoods, then more uncertainty in calibrated roughness parameters could be expected.

The results do show that extrapolation to design events must be considered carefully not only when applying data-driven or conceptual models to estimate the response to design events (Uhlenbrook *et al.*, 1999) but also when using models where the basic equations are physically-based. Lack of calibration data, particularly for events of the magnitude of design events, combined with the conceptual approach with which complex flow dynamics such as turbulent momentum interchange are resolved, results in the parameters used in these equations being uncertain and difficult to identify. Given the three models compared here, the most reliable in extrapolating to design events greater in magnitude than those available for calibration would be the integrated 1D-2D code, but even then predictive uncertainties are substantial.

Acknowledgments

References

Ackers, P., 1992. Hydraulic design of two-stage channels. *Proc. Inst. Civil Eng.- Water Mar.,* **96,** 247–257.

Ali, S. and Goodwin, P., 2002. The predictive ability of {1-D} models to simulate flood-plain processes. In: *Hydroinformatics 2002,* R.A. Falconer, B. Bin, E.L. Harris and C.A.M.E. Wilson (Eds.), IWA publishing. 247–252

Aronica, G., Bates, P.D. and Horritt, M.S., 2002. Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. *Hydrol. Process.,* **16,** 2001–2016.

Aronica, G., Hankin, B.G. and Beven, K.J., 1998. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. *Adv. Water Resour.,* **22,** 349–365.

Bates, P.D. and De Roo, A.P.J., 2000. A simple raster-based model for flood inundation simulation. *J. Hydrol.,* **236,** 54–77.

Bates, P.D., Stewart, M.D., Siggers, G.B., Smith, C.N., Hervouet, J.-M. and Sellin, R.H.J., 1998. Internal and external validation of a two dimensional finite element model for river flood simulation. *Proc. Inst. Civil Eng.- Water Mar.,* **130,** 127–141.

Beffa, C. and Faeh R., 1994. Flood Propagation on Mobile Beds under Mountainous Flow Conditions. In: *Proceedings of the Specialty Conference on Modelling of Flood Propagation over Initially Dry Areas,* P. Molinaro and L. Natale (Eds.), ASCE, NY, USA. 327–341.

Beffa, C. and Connell, R.J., 2001. Two-Dimensional Flood Plain Flow. I: Model Description, *J. Hydrol. Eng.,* **6,** 397–405.

Beven, K.J. and Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Process.,* **6,** 279–298.

Chow, V.T., 1959. *Open Channel Hydraulics.* McGraw-Hill, USA, 680pp.

Cunge, J.A., 1975, Two dimensional modeling of floodplains. In: *Unsteady Flow in Open Channels,* K. Mahmood and V. Yevjevich (Eds.), vol. II, Water Resources Publications, P.O.Box 303, Fort Collins, Co. USA. 705–762.

Frank, E., Ostan, A., Coccato, M. and Stelling, G.S., 2001. Use of an integrated one dimensional two dimensional hydraulic modelling approach for flood hazard and risk mapping. In: *Proceedings of the 1st Conference on River Basin Management,* R.A. Falconer and W.R. Blain (Eds.), WIT Press, Southampton, UK. 99–108.

Franks, S.W., Gineste, P., Beven, K.J. and Merot, P., 1998. On constraining the predictions of a distributed model: The incorporation of fuzzy estimates of saturated areas in the calibration process. *Water Resour. Res.,* **34,** 787–797.

French, R.H., 1986. *Open Channel Hydraulics,* McGraw-Hill, USA. 680pp.

Hesselink, A.W., Stelling, G.S., Kwadijk, J.C.J. and Middelkoop, H., 2003. Inundation of a Dutch river polder, sensitivity analysis of a physically based inundation model using historic data. *Water Resour. Res.,* **39,** 1234.

Horritt, M.S. and Bates, P.D., 2001. Predicting floodplain inundation: raster-based modelling versus the finite-element approach. *Hydrol. Process.,* **15,** 825–842.

Horritt, M.S. and Bates, P.D., 2002. Evaluation of 1-D and 2-D numerical models for predicting river flood inundation. *J. Hydrol.,* **268,** 87–99.

Knight, D.W. and Shiono, K., 1996. River Channel and Floodplain Hydraulics. In: *Floodplain Processes,* Chapter 5. Wiley, Chichester, UK. 139–181.

Krahe, P., Busch, N., Daamen, K., van Haselen, C., Hils, M. and Werner, M., 2000. Assessment of flood risk for the river Saar with respect to environmental changes - results of a case study within the EUROTAS project .2000. In: *European Conference on Advances in Flood Research,* A. Bronstert, C. Bismuth and L. Menzel (Eds.), Potsdam, Germany. 99–110.

Lamb, R., Beven, K. and Myrabø, S., 1998. Use of spatially distributed water table observations to constrain uncertainty in a rainfall-runoff model. *Adv. Water Resour.,* **22,** 305–317.

Marche, C., Lessard, G. and el Gharbi, B., 1990. Kriging technique for river flood representation. *J. Hydraul. Res.,* **28,** 629–643.

Marks, K. and Bates, P., 2000. Integration of high-resolution topographic data with floodplain flow models. *Hydrol. Process.,* **14,** 2109–2122.

McKay, M.D., Beckman, R.J. and Conover, W.J., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics,* **21,** 239–245.

Molinaro, P., Di Filippo, A. and Ferrari, F., 1994. Modelling of Flood Wave Propagation over Flat Areas of Complex Topography in Presence of Different Infrastructures. In: *Proceedings of the Specialty Conference on Modelling of Flood Propagation over Intitially Dry Areas,* P. Molinaro and L. Natale (Eds.) ASCE, New York, NY, USA. 209–228.

Péron, S., Quiquempoix, O. and Coureau, C., 1996. CAZIMIR: Project for a representation of the results of 1D calculations of floods resulting from dam breaks on 2D maps. In: *Hydroinformatics 1996,* A. Müller (Ed.). 123–127.

Priestnall, G., Jaafar, J. and Duncan, A., 2000. Extracting urban features from LiDAR digital surface models. *Comput., Environ. Urban Syst.,* **24,** 65–78.

Romanowicz, R.J. and Beven, K.J., 2003. Estimation of flood inundation probabilities as conditioned on event inundation maps. *Water Resour. Res.,* **39**(3), 1073, doi:10.1029/2001WR001056.

Stelling, G.S. and Duinmeijer, S.P.A., 2003. A staggered conservative scheme for every Froude number in rapidly varies shallow water flows. *Int. J. Numer. Method. Fl. F.,* **43,** 1329–1354.

Tate, E.C, Maidment, D.R., Olivera, F. and Anderson, D.J., 2002, Creating a Terrain Model for Floodplain Mapping. *J. Hydrol. Eng.,* **7,** 100–108.

Uhlenbrook, S., Seibert, J., Leibendgut, C.and Rodhe, A., 1999. Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrolog. Sci. J.,* **4,** 779–797.

Verwey, A., 2001. Latest Developments in Floodplain Modelling - 1D/2D Integration, *Conference on Civil Engineering Hydraulics,* Institution of Engineers, Hobart, Australia.

Werner, M.G.F., 2001. Impact of grid size in GIS based flood extent mapping using a 1D flow model. *Phys. Chem. Earth, B: Hydrol. Oceans Atmos.,* **26,** 517–522.

Werner, M.G.F., Ververs, M., Van Haselen, C., Daamen, K. and Muerlebach, M., 2000. A comparison of methods for generating cross sections for flood modelling using detailed floodplain elevation models. In: *European Conference on Advances in Flood Research,* A. Bronstert, C. Bismuth and L. Menzel (Eds.), Potsdam, Germany. 73–86.

WSV, 1987. *Ausbau der Saar - eine Dokumentation,* Wasser- und Schifffahrtsverwaltung des Bundes, Saarbrücken, Germany (*in German*).

Yu, P-S., Yang, T-C. and Chen, S-J., 2001. Comparison of uncertainty analysis methods for a distributed rainfall-runoff model. *J. Hydrol.,* **244,** 43–59.