ELSEVIER

Journal of Hydrology 268 (2002) 87–99

Journal of Hydrology

www.elsevier.com/locate/jhydrol

Evaluation of 1D and 2D numerical models for predicting river flood inundation

M.S. Horritt[a,*], P.D. Bates[b,1]

[a]*School of Geography, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK.*
[b]*School of Geographical Sciences, University of Bristol, University Road, Bristol BS8 1SS, UK*

Received 11 January 2002; revised 13 May 2002; accepted 22 May 2002

Abstract

1D and 2D models of flood hydraulics (HEC-RAS, LISFLOOD-FP and TELEMAC-2D) are tested on a 60 km reach of the river Severn, UK. Synoptic views of flood extent from radar remote sensing satellites have been acquired for flood events in 1998 and 2000. The three models are calibrated, using floodplain and channel friction as free parameters, against both the observed inundated area and records of downstream discharge. The predictive power of the models calibrated against inundation extent or discharge for one event can thus be measured using independent validation data for the second. The results show that for this reach both the HEC-RAS and TELEMAC-2D models can be calibrated against discharge or inundated area data and give good predictions of inundated area, whereas the LISFLOOD-FP needs to be calibrated against independent inundated area data to produce acceptable results. The different predictive performances of the models stem from their different responses to changes in friction parameterisation.

Keywords: Flood forecasting; Modelling; Calibration; Validation; Remote sensing

1. Introduction

Recent work on calibration and validation of 2D models of river flood inundation (Feldhaus et al., 1992; Romanowicz et al., 1996; Romanowicz and Beven, 1997; Bates et al., 1998; Horritt, 2000; Bates and De Roo 2000; Horritt and Bates, 2001a,b; Aronica et al., 2002) has demonstrated how both raster-based and finite element approaches can be used to reproduce observed inundation extent and bulk hydrometric response for river reaches 5–60 km in length. Remote sensing has also been demonstrated to be a useful tool in mapping flood extent (Horritt et al., 2001) and hence for validating numerical inundation models. These studies have, however, been limited to model calibration against a single flood event, and are therefore only a limited test of the models' predictive power. Furthermore, calibration or validation of a 1D model against 2D inundation data has yet to be carried out. There is therefore no indication of the relative performance of 1D and 2D approaches in predicting floodplain inundation. The predictive power of flood inundation models is assessed in this paper through the analysis of two flood events on the same river reach and a comparison with the predictions of 1D and 2D modelling approaches.

The use of remotely sensed maps of flood extent

* Corresponding author. Tel.: +44-113-343-6833; fax: +44-113-343-3308.

E-mail addresses: m.horritt@geog.leeds.ac.uk (M.S. Horritt), paul.bates@bristol.ac.uk (P.D. Bates).

[1] Tel.: +44-117-928-9108; fax: +44-117-928-7878.

0022-1694/02/$ - see front matter
PII: S0022-1694(02)00121-X

(Smith, 1997; Bates et al., 1997; Horritt et al., 2001) to validate flood models (Horritt, 2000) has strongly influenced the development of such codes in recent years. The 2D nature of flood maps has promoted the use of 2D models in order to promote the synergy between distributed observations and predictions, whereas point measurements of stage or discharge are more compatible with 1D models. The high resolution of remotely sensed data, especially from synthetic aperture radar (SAR) systems (typically a few tens of metres), has encouraged modelling at a higher spatial resolution than was previously practical, and has also encouraged the integration of high resolution DEMs into hydraulic models (Marks and Bates, 2000). Inundation extent data have also provided a second source of observed data independent of hydrometry, allowing models to be independently calibrated and validated. This has already exposed weaknesses in raster-based models of dynamic flood events (Horritt and Bates, 2001a), showing that model predictions calibrated against inundation extent reproduce bulk flow behaviour poorly, and vice versa. It is impossible to detect this poor performance with calibration data alone, and the combination of inundation extent data and hydrometry can therefore be used to discriminate between models in a rigorous fashion. This has also shaped the criteria for model evaluation, for example, by defining a good model as being one that can reproduce flood extent when calibrated against hydrometric data. This property will become important as inundation models are applied in operational scenarios, where models calibrated against (relatively common) hydrometric data will be used to predict flood extent for which only limited validation data sets are available.

The integration of hydrometric and flood extent data has thus been shown to be useful in discriminating between flood inundation models, and if a calibration methodology is adopted for unconstrained parameters, the optimal parameter values may well be different for models calibrated against hydrometric and inundation data. A further complication is that optimal parameter sets may well be different for different flood events, and this will reduce a model's predictive power. In particular, it is unclear as to whether a parameter set calibrated against data from an event with a certain magnitude will be valid for a more extreme event. One of the criteria for assessing model performance thus becomes the stability of the model calibration with respect to changing event magnitude. This is of practical importance, since we would like to be able to calibrate a hydraulic model against observed low magnitude (and hence relatively common) events, and use the calibrated model to predict the impact of larger magnitude events, for example, the one in 100 year flood, for planning and risk assessment purposes. Due to the rarity of inundation extent data sets, this has so far not been undertaken.

The research presented in this paper aims to assess the performance of 1D and 2D flood models, and in particular their ability to predict inundation extent for one flood event when calibrated on another. This will allow us to assess the models' suitability for practical risk and hazard assessment. The models used reflect a move in recent years from a 1D approach (represented by the US Army Corps of Engineers HEC-RAS model) towards 2D finite element (TELEMAC-2D developed by Electricité de France) and raster-based (LISFLOOD-FP) models. These models are tested on a 60 km reach of the river Severn, UK, where two flood events have been observed with satellite borne SAR sensors. The paper proceeds with a brief description of the three models, the test site and validation data, results of calibration studies for the two events, and an assessment of model performance when model calibrations are transferred between two flood events.

2. Models and test site

2.1. HEC-RAS

The HEC-RAS model solves the full 1D St Venant equations for unsteady open channel flow:

$$\frac{\partial A}{\partial t} + \frac{\partial \phi Q}{\partial x_c} + \frac{\partial (1-\phi) Q}{\partial x_f} = 0 \tag{1}$$

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x_c}\left(\frac{\phi^2 Q^2}{A_c}\right) + \frac{\partial}{\partial x_f}\left(\frac{(1-\phi)^2 Q^2}{A_f}\right) + gA_c\left(\frac{\partial z}{\partial x_c} + S_c\right) + gA_f\left(\frac{\partial z}{\partial x_f} + S_f\right) = 0 \tag{2}$$

$$\phi = \frac{K_c}{K_c + K_f}, \qquad \text{where } K = \frac{A^{5/3}}{nP^{2/3}} \tag{3}$$

$$S_c = \frac{\phi^2 Q^2 n_c^2}{R_c^{4/3} A_c^2}, \qquad S_f = \frac{(1-\phi)^2 Q^2 n_f^2}{R_f^{4/3} A_f^2} \tag{4}$$

Q is the total flow down the reach, A (A_c, A_f) the cross sectional area of the flow (in channel, floodplain), x_c and x_f are distances along the channel and floodplain (these may differ between cross sections to allow for channel sinuosity), P the wetted perimeter, R the hydraulic radius (A/P), n the Manning's roughness value and S the friction slope. ϕ determines how flow is partitioned between the floodplain and channel, according to the conveyances K_c and K_f. These equations are discretized using the finite difference method and solved using a four point implicit (box) method.

2.2. LISFLOOD-FP

LISFLOOD-FP is a raster-based inundation model specifically developed to take advantage of high resolution topographic data sets (Bates and De Roo, 2000). Channel flow is handled using a 1D approach that is capable of capturing the downstream propagation of a floodwave and the response of flow to free surface slope, which can be described in terms of continuity and momentum equations as:

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = q \tag{5}$$

$$S_0 - \frac{n^2 P^{4/3} Q^2}{A^{10/3}} - \frac{\partial h}{\partial x} = 0 \tag{6}$$

Q is the volumetric flow rate in the channel, A the cross sectional area of the flow, q the flow into the channel from other sources (i.e. from the floodplain or possibly tributary channels), S_0 the down-slope of the bed, n Manning's coefficient of friction, P the wetted perimeter of the flow, and h the flow depth. In this case, the channel is assumed to be wide and shallow, so the wetted perimeter is approximated by the channel width. Eqs. (5) and (6) are discretized using finite differences and a fully implicit scheme for the time dependence, and the resulting non-linear system is solved using the Newton–Raphson scheme. Sufficient boundary conditions are provided by an imposed flow at the upstream end of the reach and an imposed water elevation at the downstream end. The channel parameters required to run the model are its width, bed slope, depth (for linking to floodplain flows) and Manning's n value. Width and depth are assumed to be uniform along the reach, their values assuming the average values taken from channel surveys. Channel surveys also provide the bed elevation profile, which can have a gradient which varies along the reach, and which also may become negative if the diffusive wave model is used. The Manning's n roughness is left as a calibration parameter.

Floodplain flows are similarly described in terms of continuity and momentum equations, discretized over a grid of square cells which allows the model to represent 2D dynamic flow fields on the floodplain. We assume that the flow between two cells is simply a function of the free surface height difference between those cells (Estrela and Quintas, 1994):

$$\frac{dh^{i,j}}{dt} = \frac{Q_x^{i-1,j} - Q_x^{i,j} + Q_y^{i,j-1} - Q_y^{i,j}}{\Delta x \Delta y} \tag{7}$$

$$Q_x^{i,j} = \frac{h_{flow}^{5/3}}{n} \left(\frac{h^{i-1,j} - h^{i,j}}{\Delta x} \right)^{1/2} \Delta y \tag{8}$$

where $h^{i,j}$ is the water free surface height at the node (i,j), Δx and Δy are the cell dimensions, n is the Manning's friction coefficient for the floodplain, and Q_x and Q_y describe the volumetric flow rates between floodplain cells. Q_y is defined analogously to Eq. (8). The flow depth, h_{flow}, represents the depth through which water can flow between two cells, and is defined as the difference between the highest water free surface in the two cells and the highest bed elevation (this definition has been found to give sensible results for both wetting cells and for flows linking floodplain and channel cells). While this approach does not accurately represent diffusive wave propagation on the floodplain, due to the decoupling of the x- and y-components of the flow, it is computationally simple and has been shown to give very similar results to a faithful finite difference discretisation of the diffusive wave equation (Horritt and Bates, 2001b).

Eq. (8) is also used to calculate flows between floodplain and channel cells, allowing floodplain cell depths to be updated using Eq. (7) in response to flow

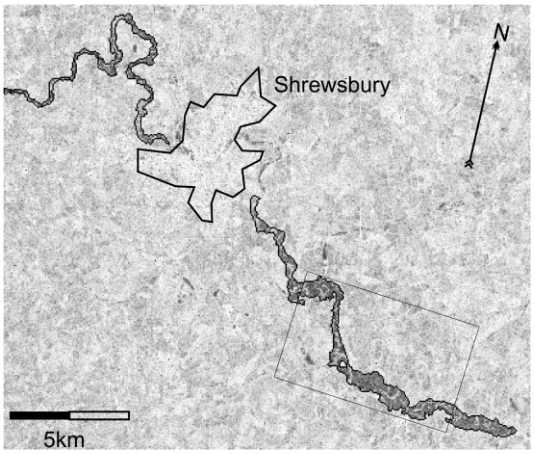

Fig. 1. RADARSAT image of the reach used to test the models. The flood boundary is delineated in black, and the urban area of Shrewsbury shown. The rectangle gives the approximate location of the detail images in Fig. 3 and subsequent maps of model predicted inundation extent.

from the channel. These flows are also used as the source term in Eq. (5), effecting the linkage of channel and floodplain flows. Thus only mass transfer between channel and floodplain is represented in the model, and this is assumed to be dependent only on relative water surface elevations. While this neglects effects such as channel-floodplain momentum transfer and the effects of advection and secondary circulation on mass transfer, it is the simplest approach to the coupling problem and should reproduce some of the behaviour of the real system.

2.3. TELEMAC-2D

The TELEMAC-2D (Galland et al., 1991; Hervouet and Van Haren, 1996) model has been applied to fluvial flooding problems for a number of river reaches and events (Bates et al., 1998). The model solves the 2D shallow water (also known as Saint-Venant or depth averaged) equations of free surface flow:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} + g\nabla(z_0 + h) + \frac{n^2 g\mathbf{v}|\mathbf{v}|}{h^{4/3}} = 0 \qquad (9)$$

$$\frac{\partial h}{\partial t} + \nabla(h\mathbf{v}) = 0 \qquad (10)$$

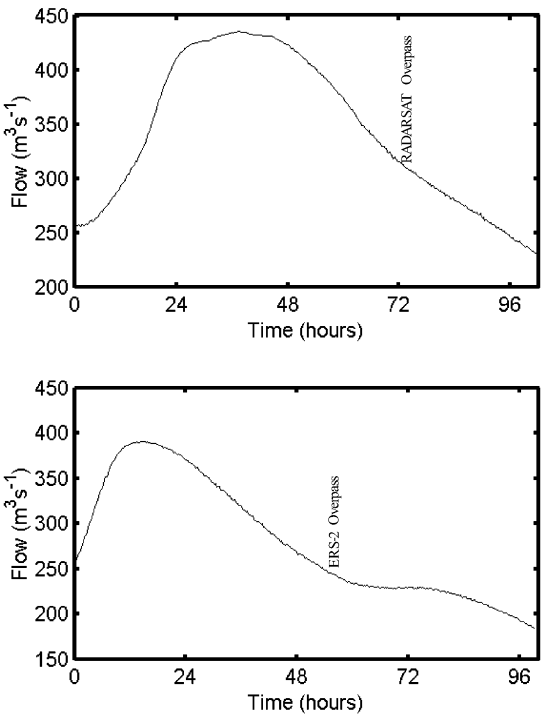

Fig. 2. Upstream hydrographs for the 1998 (top) and 2000 (bottom) flood events. The satellite overpass times are also marked.

where $\mathbf{v}$ is a 2D depth averaged velocity vector, h is the flow depth, z_0 the bed elevation, g the acceleration due to gravity, and n Manning's coefficient of friction. The TELEMAC-2D model uses Galerkin's method of weighted residuals to solve Eqs. (9) and (10) over an unstructured mesh of triangular finite elements. A streamline-upwind-Petrov–Galerkin (SUPG) technique is used for the advection of flow depth in the continuity equation to reduce the spurious spatial oscillations in depth that Galerkin's method is predisposed to, and the method of characteristics is used for the advection of velocity. The resulting linear system is solved using a gradient mean residual technique and efficient matrix assembly is ensured using element-by-element methods. The time development of the solutions is dealt with using an implicit finite difference scheme and the moving boundary nature of the problem is treated with a simple wetting and drying algorithm which eliminates spurious free surface slopes at the shoreline (Hervouet and Janin, 1994).

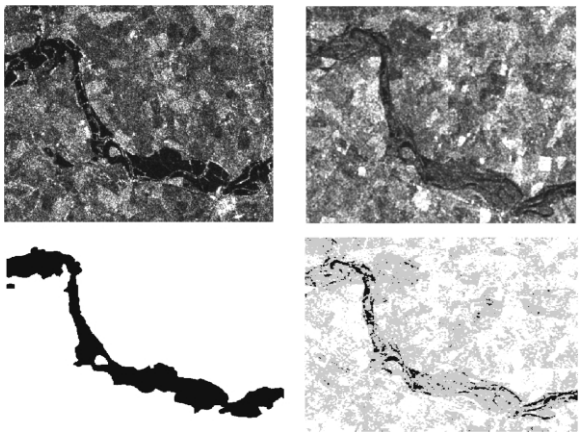

Fig. 3. 1998 RADARSAT (left) and 2000 ERS-2 (right) images (top), with their classifications (bottom). Black represents flooded regions, and grey tones represent undetermined areas. The images are ~7 km across, and cover the lower part of the test reach. © European Space Agency and Canadian Space Agency.

2.4. Test site and validation data

The three models have been set up to represent a 60 km reach of the river Severn, UK (Fig. 1). The reach is well provided with topographic data. The channel is described by a series of 19 ground surveyed cross sections, and airborne laser altimetry (Mason et al., 1999; Cobby et al., 2000) is used to derive a high resolution (50 m), high accuracy (~15 cm) floodplain DEM. Validation data are provided by two satellite SAR images. The first coincides with a flood event on 30th October 1998, with a peak flow of 435 $m^3 s^{-1}$ at the upstream end of the reach (Fig. 2, top). This is well over bankful discharge (~180 $m^3 s^{-1}$), and considerable floodplain inundation occurred. The event was observed by the RADARSAT satellite, operating at C-band (5.3 GHz, 5.6 cm), HH polarisation and an incidence angle of 36–42° (shown as backdrop to Fig. 1). The second event occurred on 11th November 2000, with a peak flow of 391 $m^3 s^{-1}$ (Fig. 2, bottom), and was captured by the ERS-2 satellite SAR, again operating at C-band, VV polarisation and an incidence angle of 20–26°. Both sensors produce imagery with a pixel size of 12.5 m and a ground resolution of approximately 25 m.

Imagery from the two sensors is shown in Fig. 3. They show that the flood is easily distinguished as a region of very low backscatter in the RADARSAT imagery (Fig. 3, top left), but detecting the waterline is much more difficult in the ERS-2 image (top right), where wind roughening of the water surface has increased backscatter to levels similar to some floodplain landcover types (Horritt, 2000; Horritt et al., 2001). This differential sensitivity to wind roughening is probably due to both the different incidence angle and polarisations used by the two sensors. The consequence is that for the greater part of the reach, it is extremely difficult to distinguish between the flooded and unflooded state in the ERS-2 imagery, but the shoreline is obvious in some places, probably where the water surface is sheltered from the wind by trees or topography. Given the different nature of the images, two processing strategies were adopted. The shoreline in the 1998 RADARSAT imagery was delineated using a statistical active contour model (Horritt, 1999; Horritt et al., 2001) which has been found to be capable of locating the shoreline to ~2 pixels, and then transformed to a raster map of the flooded/non-flooded state at a resolution of 12.5 m. The 2000 ERS-2 imagery was first smoothed using a 3×3 moving average filter to reduce speckle, then thresholded into 3 classes: flooded, unflooded and undetermined, again at 12.5 m resolution. The undetermined class can then be ignored in any calculation of model fit. The results of this classification are also given in Fig. 3 (bottom left and bottom right).

Hydrometric data can also be used for validation/-calibration. Although the downstream stage is already used as a model boundary condition, the floodwave travel time has been found to be useful in model calibration (Bates et al., 1998) even if it is not wholly independent of downstream stage measurements. A model where peak flow at the downstream boundary coincides with peak-measured flow can be viewed as predicting travel time correctly. The measured travel times were similar for the two events: 25.5 h for 1998 and 24 h for 2000.

The HEC-RAS model was set up using the 19 cross sections to provide the channel width and bed elevations. These sections were extended on both sides of the channel using the LiDAR derived DEM to provide floodplain topography. The section was then described by 5–10 points on each side of the channel coinciding with significant topographic features such as breaks of slope. The bed elevation profile and examples of the cross sections used in the HEC-RAS

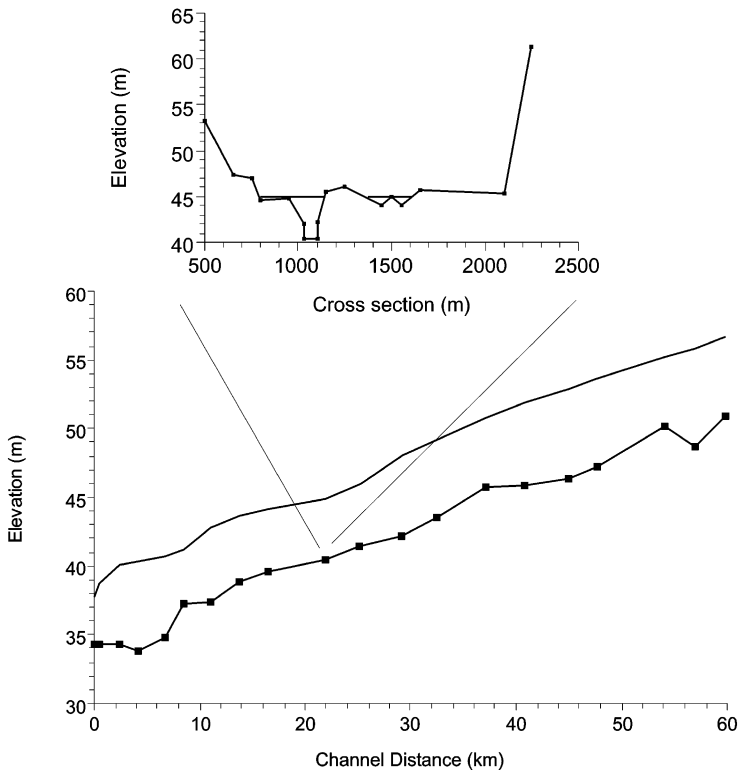

Fig. 4. Bed elevation profile, with a typical water free surface, for the HEC-RAS model (bottom). A sample cross section is also shown (top).

model are given in Fig. 4. Boundary conditions for the model are an imposed dynamic discharge at the upstream end of the reach and an imposed water surface elevation at the downstream end, both provided by stage recorders and a rated section in the case of the imposed discharge. Although the use of the measured free surface elevation at the downstream end does mean that the boundary conditions and validation data (travel times) are not fully independent, the effect of this was found to be small. The floodwave travel time remains a good source of calibration data, being strongly dependent on the model calibration and not significantly affected by the downstream boundary condition. Predicted inundation extent was then derived by re-projecting the water levels at the 19 cross sections onto the high resolution DEM.

The LISFLOOD-FP model is based directly on the 50 m resolution DEM. The location of the channel is digitised from 1:25000 scale maps of the reach. Since the channel width is of the same order as the model resolution, cells of the DEM lying over the channel can be ignored in the floodplain flow calculations (flow between channel cells being handled by the in-channel diffusive flow routing scheme), and floodplain storage near the channel does not need to be included explicitly as it does for coarse scale LISFLOOD-FP models (Horritt and Bates, 2001a). Bed elevations for the 1D flow routing scheme are taken from the 19 surveyed cross sections, and linearly interpolated in between. Channel width is taken as constant down the channel. Again the upstream boundary condition is provided by the measured discharge, and the downstream boundary condition is an imposed water surface elevation.

The TELEMAC-2D model operates on a mesh of 6485 nodes and 12107 triangular elements (Fig. 5), giving a floodplain resolution of $\sim$30 m. Floodplain topography is sampled onto the mesh using nearest neighbours from the 50 m DEM, and the channel and bank node elevations are taken from channel surveys

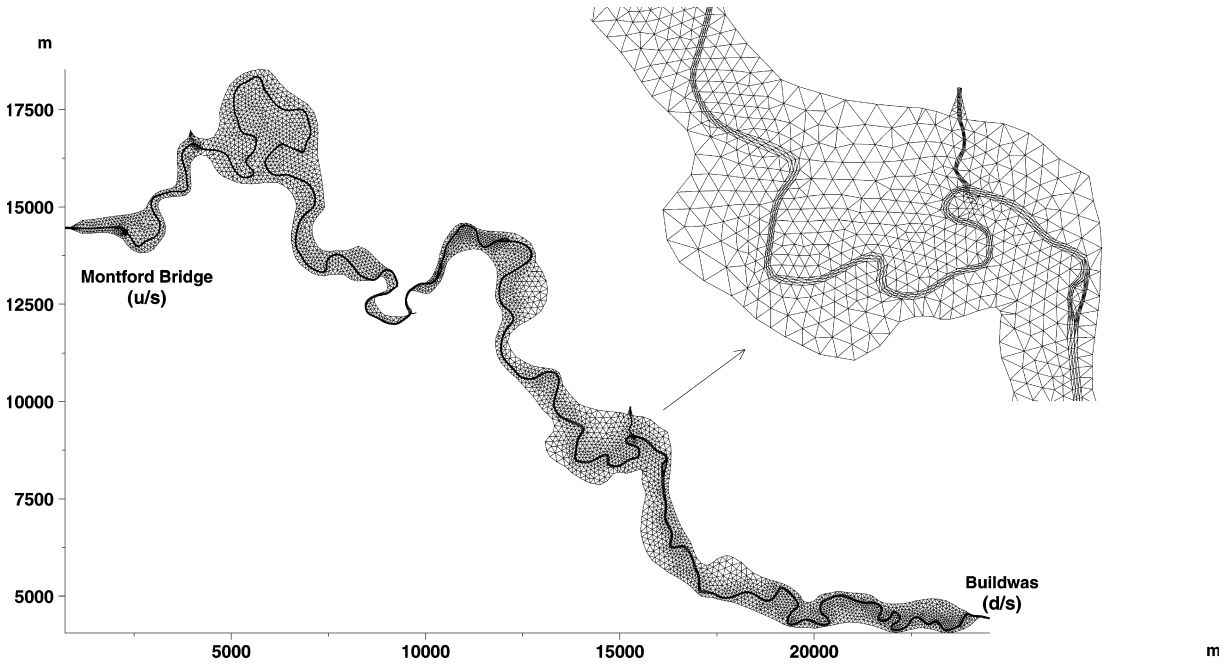

Fig. 5. Finite element mesh for the TELEMAC-2D model.

and linearly interpolated between the 19 cross sections. Channel planform and the extent of the domain are digitised from 1:25000 maps of the reach. The same boundary conditions are used as for the other two models.

3. Model calibration and validation

To make a tractable calibration problem, the potentially distributed bed roughness coefficients are limited to one value for the channel and one for the floodplain. This is also appropriate for the lumped criteria used to assess model performance in predicting inundation (described below). Channel values vary between $n = 0.01$ and $n = 0.05\ \mathrm{m}^{-1/3}$ s, with floodplain values ranging from $n = 0.02$ to $n = 0.10\ \mathrm{m}^{-1/3}$ s for the HEC-RAS and LISFLOOD-FP models. The TELEMAC calibration uses lower friction values ($n = 0.005$ to $n = 0.04\ \mathrm{m}^{-1/3}$ s for the channel, $n = 0.01$ to $n = 0.08\ \mathrm{m}^{-1/3}$ s on the floodplain), which were found necessary to ensure the optimum Manning's n values are included within the parameter space. It should be noted that the different process inclusion in each model would mean that the friction value has a different physical meaning and is drawn from a different distribution. Hence in a 1D model, the friction term accounts for the energy loss due to planform variations, whereas, for a 2D finite element model these losses are represented directly in the domain geometry at the element scale and only subsumed within the friction term at the subgrid scale. The friction coefficients for each model can thus not be absolutely compared. A simulation for each of the models, friction parameterisations and events was performed.

Model predictions of inundation extent are compared with the satellite data using the measure of fit:

$$F = \frac{\mathrm{Num}(S_{\mathrm{mod}} \cap S_{\mathrm{obs}})}{\mathrm{Num}(S_{\mathrm{mod}} \cup S_{\mathrm{obs}})} \times 100 \tag{11}$$

S_{mod} and S_{obs} are the sets of domain sub-regions (pixels, elements or cells) predicted as flooded by the model and observed to be flooded in the satellite imagery, and Num(·) denotes the number of members of the set. F therefore varies between 0 for model with no overlap between predicted and observed inundated areas and 100 for a model where these coincide perfectly. This measure has

Table 1
Summary of optimum friction coefficients calibrated on inundated area for the 3 models and two events

Event	Model	n_{ch}	n_{fl}	F
1998	HEC-RAS	0.05	0.06	64.83
	LISFLOOD-FP	0.03	0.06–0.08	63.81
	TELEMAC	0.02	0.02	65.28
2000	HEC-RAS	0.04	0.10	41.79
	LISFLOOD-FP	0.03	0.08	41.38
	TELEMAC	0.02	0.04	37.41

Table 2
Summary of optimum friction coefficients calibrated on floodwave travel time for the 3 models and two events. The observed travel times were 25.5 h for the 1998 event and 24 h for the 2000 event

Event	Model	n_{ch}	n_{fl}	Best travel time
1998	HEC-RAS	0.04–0.05	0.08–0.10	20
	LISFLOOD-FP	0.02	0.02–0.10	27
	TELEMAC	0.04	0.08	25
2000	HEC-RAS	0.03–0.05	0.06–0.10	26
	LISFLOOD-FP	0.02	0.02–0.10	29
	TELEMAC	0.02	0.06	24

been found to give good results in other calibration studies (Aronica et al., 2002; Horritt and Bates, 2001a), and allows a useful comparison between model performance for different modelled reaches and flood events. The models also generate a downstream discharge, which is used to predict the floodwave travel time.

Results of the calibration are given in Tables 1 and 2 for the three models applied to the 1998 and 2000 events. On first inspection, the tables appear to show

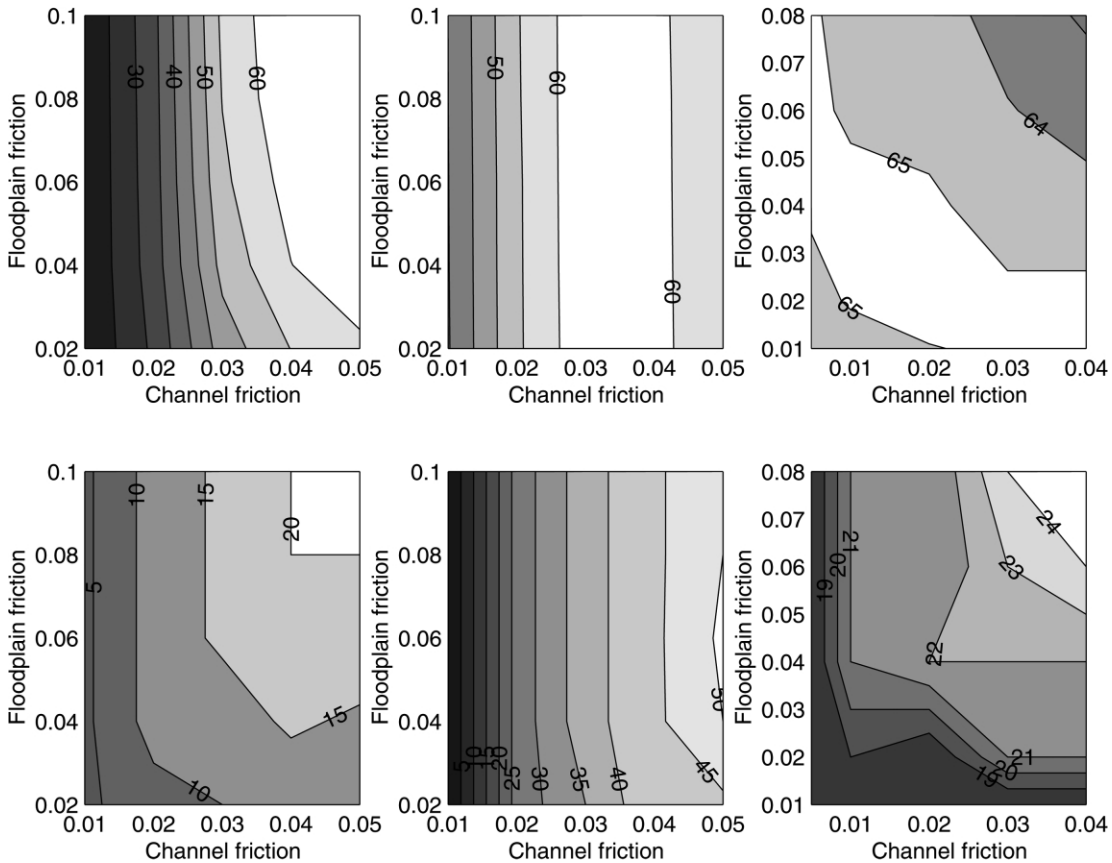

Fig. 6. Model sensitivity: the surfaces show how model performance changes with friction coefficients. Top row: measure of fit F for the 1998 event. Bottom row: floodwave travel time for the 1998 event. Left: HEC-RAS, middle: LISFLOOD-FP, right: TELEMAC-2D.

Table 3
Predictive performance of the 3 models using independent calibration/validation data

Calibration data	Validation data	HEC-RAS	LISFLOOD-FP	TELEMAC-2D
1998 Hydro	1998 SAR	64.24–64.74	54.10–54.57	62.88
2000 Hydro	2000 SAR	37.79–41.79	33.06–33.72	37.13
1998 Hydro	2000 SAR	40.80–41.79	33.06–33.72	35.90
2000 Hydro	1998 SAR	55.21–64.83	54.10–54.57	64.46
1998 SAR	2000 SAR	41.48	41.29–41.38	36.85
2000 SAR	1998 SAR	64.65	63.81	65.21

that in terms of calibration, the LISFLOOD-FP models is the best, having approximately the same optimal friction values for both events and giving similar optima when calibrated against the radar and hydrometric data. The model with the most sophisticated process representation, TELEMAC-2D, performs more poorly in this respect than the other models. However, all three models give similar levels of performance in terms of F at their optimum calibrations, $\sim$65% for the 1998 event and $\sim$40% for the 2000 event. The difference in maximum performance for the two events is due to the different sources of validation data. The ERS-2 image contains large areas of uncertain flood state, and in particular lacks data for large areas in the middle of the domain which are easily and correctly predicted as flooded by the models of the 1998 event. If the uncertain areas are due to wind roughening of the water surface, then we might expect the more sheltered areas at the sides of the valley to provide most of the validation data, just where we would expect the biggest differences in performance to show up.

Further differences in model performance are exposed if the full calibration surface (measure of fit F or floodwave travel time as a function of roughness values) is examined, as shown in Fig. 6 for the 1998 event. Results for the 2000 event are similar, but with lower values of F due to the poor quality of the ERS-2 data. As in Horritt and Bates (2001b), we see that the different model respond differently to changing friction parameters. HEC-RAS shows the most consistent response of inundation extent and travel times: both are optimised for high friction values. The good performance may, however, be due in part to the constraint on the maximum friction values, and using very high values in order to reproduce the 25.5 h travel time may result in poorer inundation predictions. The different sensitivities to channel and floodplain friction are also what might be expected. For low channel friction, water levels are low, there is little inundation and so sensitivity to floodplain friction is minimal. Higher channel frictions increase water depth in the channel, more water is forced onto the floodplain, and so floodplain friction exerts a greater influence on both travel times and inundation extent. LISFLOOD-FP gives a smooth response surface, with very little sensitivity to floodplain friction when the inundated area is considered. The model is slightly more sensitive to floodplain friction when floodwave travel times are compared. The surfaces for TELEMAC-2D are similar in shape to those for HEC-RAS, but showing less overall sensitivity, varying only between 63 and 65%, compared with 20–65% for HEC-RAS.

How does this calibration performance affect the predictive performance of the three models? The answer depends on how we use the models in the predictive mode. Records of time varying discharge, and hence floodwave travel times, are relatively common, but inundation extent data are still relatively rare due to the limited number of radar satellites in operation. It would therefore make sense to define a 'good' model as one which can be calibrated against discharge measurements and then provides the most accurate predictions of flood extent. This is also in accord with the role of inundation models in flood risk assessment, where it is the extent of the flood, rather than the discharge, which is of interest.

The performance is therefore assessed by calibrating on one flood event and measuring the performance in predicting inundated area for the other, calibrating using the hydrometric data and measuring F (Eq. (11)) for the same event, etc. These combinations give 6 performance evaluations using

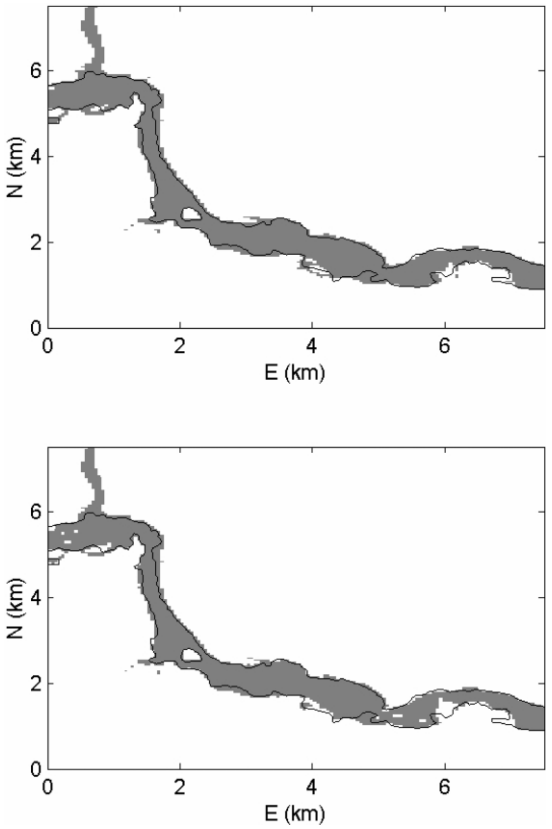

Fig. 7. Flooded area predicted by HEC-RAS model for the 1998 event when calibrated on hydrometric data for the 1998 event (top), and when calibrated on the flood extent data from the 2000 event (bottom).

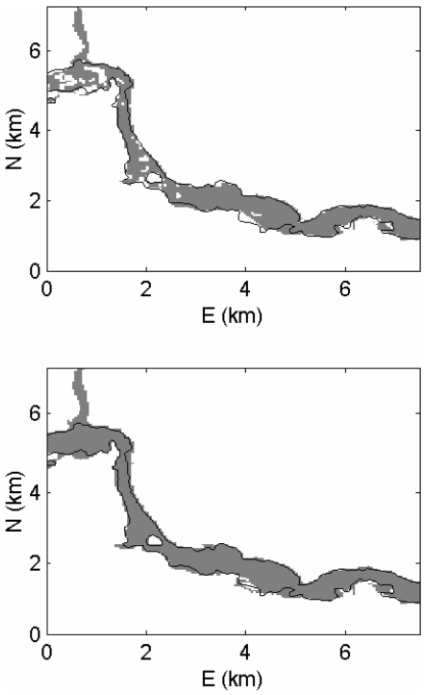

Fig. 8. Flooded area predicted by LISFLOOD-FP model for the 1998 event when calibrated on hydrometric data for the 1998 event (top), and when calibrated on the flood extent data from the 2000 event (bottom).

different stage/inundation data combinations, which are given in Table 3. Where ranges of the performance are given, this is because the calibration fails to pick out a single friction value, several parameter sets giving the same performance when compared with floodwave travel times. Some conclusions can be drawn from the results:

1. In terms of simulations where the model is calibrated against the data it is trying to predict (Tables 1 and 2), the TELEMAC-2D marginally outperforms the other models apart from for the 2000 inundation data. In terms of predictive performance (Table 3), the HEC-RAS model performs marginally better (on 4 out of 6 measures) than TELEMAC-2D, which shows the best performance on 2 measures and which in turn is better than LISFLOOD-FP. LISFLOOD-FP does not produce the best predictive performance for any of the measures and data sets examined. It is interesting to note that TELEMAC-2D is better when predicting inundation for the 1998 than HEC-RAS, the situation being reversed for the 2000 event.
2. The HEC-RAS model performs as well in terms of predicting inundated area when calibrated on inundated area for the other event as when calibrated on floodwave travel time.
3. Predictions of flood extent from the LISFLOOD-FP model are significantly poorer when the model is calibrated against floodwave travel time compared with when this model is calibrated against inundated area, although this is consistent with the code design aims (predicting inundation extent rather than flow routing properties) and previous sensitivity analysis studies of this class of model (Romanowicz et al., 1996).
4. The TELEMAC model results do not depend on

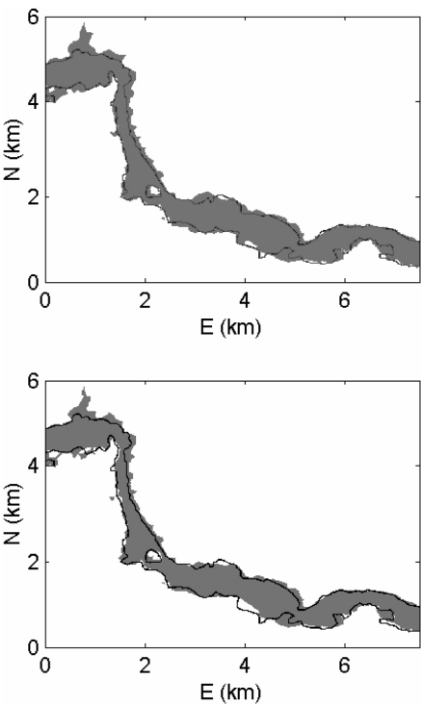

Fig. 9. Flooded area predicted by the TELEMAC-2D model for the 1998 event when calibrated on hydrometric data for the 1998 event (top), and when calibrated on the flood extent data from the 2000 event (bottom).

the calibration data, giving similar predictive results when calibrated against travel times and inundated areas.

The differences in performance are illustrated in Fig. 7 for the HEC-RAS model, in Fig. 8 for the LISFLOOD-FP model and in Fig. 9 for the TELEMAC-2D model. The results are shown for the 1998 event modelled using the calibrations found using the 1998 hydrometric data and the 2000 SAR data. The HEC-RAS and TELEMAC-2D models perform equally well in predicting the inundated area whether calibrated on the 2000 inundated area or the floodwave travel time, whereas the performance of the LISFLOOD-FP model is dependent on the calibration data used.

4. Discussion

The results presented here confirm some of the assumptions usually made about flood model calibration and validation. The 1D HEC-RAS model can be adequately calibrated on hydrometric data, and can then be used to make adequate predictions of flood extent when water free surfaces are extrapolated onto a high resolution DEM. This is a potentially useful property of the model, as it will permit flood extent predictions (for which few validation data sets are available) to be made from models calibrated on relatively common hydrometric data. The TELEMAC-2D model performs similarly, while the LISFLOOD-FP has difficulty in predicting flood extent when calibrated against hydrometric data, although this is consistent with its design philosophy and physical basis.

The differences in predictive performance for the three models arises from their different response to the calibration process. HEC-RAS has the optimum friction coefficients in the same region of the parameter space, and thus inundated area is predicted well when the model is calibrated against hydrometric data. The reason for LISFLOOD-FP's poor predictive performance is evident, since the optima occur for different channel friction values when calibrated against the inundated area and floodwave travel times. The model is also quite sensitive to channel friction (for both predicted area and travel time), and so the LISFLOOD-FP model when calibrated against floodwave travel time reproduces the inundation extent poorly. The optima for TELEMAC-2D lie at diametrically opposite sides of the parameter space, and TELEMAC-2D's good predictive performance comes in this case from its lower sensitivity to either friction parameter: any choice will give a good prediction of inundation extent, and calibration is less important than for the other two models. The insensitivity of the TELEMAC-2D model to friction parameterisation has been noted when applied to another reach (Horritt, 2000) and may reflect the physical basis of friction in full 2D unstructured models as here more components of the 1D friction term are explicitly represented in the model mesh. The different sensitivities may thus tell us something about the relative significance of various components of frictional resistance during floods.

The similarities between the two 2D approaches would be expected to be greater, whereas the results show the HEC-RAS and TELEMAC-2D models giving similar levels of performance despite their different dimensionalities. Both TELEMAC-2D and HEC-RAS

use the full shallow water equations, in 1D and 2D forms, so the inclusion of inertia and advection terms might contribute to the difference. LISFLOOD-FP's insensitivity to floodplain friction indicates that floodplain flow processes (a momentum effect) are unimportant in the model, despite floodplain storage (a mass conservation effect) having a significant effect on floodwave travel time (Horritt and Bates, 2001a). The other two models show more sensitivity to floodplain friction (as well as both including floodplain storage effects), indicating that floodplain flow may be having an effect on both inundation extent and floodwave travel time. Thus floodplain flow processes or increased channel process representation may be causing the different calibration response. In any case, the different responses indicate that for some or all of the models, friction parameters are being used to compensate for different process representations, and are not simply parameterising the bed friction terms, as we would wish.

The somewhat surprising result is how well the 1D HEC-RAS model performs compared to the more sophisticated 2D approaches adopted by LISFLOOD-FP and TELEMAC-2D. The methodology used here does possess some features, which will tend to handicap the 2D models and reduce the process representation advantage they have over HEC-RAS. Two-dimensional models can use a more distributed friction parameterisation, which may improve results, although this would make for a more complex (verging on the intractable) calibration problem. All three models are also using the same LiDAR-derived DEM, and this will certainly help in producing accurate shorelines. The results indicate too that it is the topography, which is the major factor determining flood inundation patterns, and it is a relatively simple task to model water levels and bulk flow dynamics. The key step, which then determines the quality of the inundation predictions is thus the re-projection of these results back on to a DEM. Where this is of high resolution and accuracy this study indicates, it is likely, that even quite simple models will perform adequately. This accords with previous results from (Horritt and Bates, 2001a) from a study of the scaling behaviour of the LISFLOOD-FP model when applied to the 1998 Severn flood event where it was demonstrated that even very coarse resolution models (250–500 m grid cells) performed as well as high resolution models if a re-projection step was included.

The results presented here should, however, be treated with caution, and the extension of the methodology to other reaches and flood events may reveal different behaviour. Flow on this reach is confined to a relatively narrow valley, and we might expect more complex overbank flow patterns in wider floodplains, and in that case the 2D approach may prove more effective than 1D. Indeed, the TELEMAC-2D model applied to a reach of the river Thames has shown more sensitivity to floodplain friction than to channel friction (Horritt, 2000), indicating that floodplain flow processes are more important for this reach. The reach has a much wider floodplain ($\sim$1.5 km wide, compared to a channel width of 20 m) than the Severn (floodplains 200–1000 m wide, compared to a 50 m wide channel), and so model response to calibration is seen to be a function of the channel-floodplain geometry. In the case where floodplain flow processes are important, a wide floodplain of complex morphology may therefore require a 2D approach. Two-dimensional modelling may also be more appropriate if other hydraulic processes, such as turbulent momentum exchange between channel and floodplain waters (Knight, 1989; Knight and Shiono, 1996), are significant. These results cannot therefore be taken as representative of all river–floodplain systems. They do show, however, that in some cases a 1D approach may be very effective in predicting flood extent, as long as the results can be projected onto a high resolution DEM.

5. Conclusions

The ability of flood flow models to predict inundation extent and floodwave travel times has been tested using independent calibration data from hydrometric and satellite sources. All three models are capable of predicting flood extent and travel times to similar levels of accuracy at optimum calibration. However, differences emerge according to the calibration data used when the models are used in predictive mode. Both HEC-RAS and TELEMAC-2D are capable of making equally good predictions of inundated area, whether calibrated against floodwave travel times or inundated area data from another event. The LISFLOOD-FP model, however, requires independent inundated area data for calibration in

order to make good predictions of inundation extent—calibration against discharge data gives relatively poor results. The differences in predictive performance are due to the different model responses to friction parameterisations. If we are forced to use a calibration methodology due to lack of parameterisation data, the best model will be the one with the most useful response surface. The work presented here indicates that the HEC-RAS model is the best model for this reach when assessed in these terms.

Acknowledgments

This work was supported by the UK Natural Environment Research Council grant GR3 CO030 and the EC FP5 project European Flood Forecasting System, and MSH is supported by a NERC research fellowship. Satellite, LiDAR and hydrometric data were provided by the Environment Agency. LiDAR processing was carried out by David Cobby and David Mason at the Environmental Systems Science Centre at the University of Reading, UK.

References

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

Bates, P.D., De Roo, A.P.J., 2000. A simple raster-based model for flood inundation simulation. Journal of Hydrology 236, 54–77.

Bates, P.D., Horritt, M.S., Smith, C.N., Mason, D.C., 1997. Integrating remote sensing observations of flood hydrology and hydraulic modelling. Hydrological Processes 11, 1777–1795.

Bates, P.D., Stewart, M.D., Siggers, G.B., Smith, C.N., Hervouet, J.-M., Sellin, R.J.H., 1998. Internal and external validation of a two-dimensional finite element code for river flood simulations. Proceedings of the Institute of Civil Engineers, Water, Maritime and Energy 130, 127–141.

Cobby, D.M., Mason, D.C., Davenport, I.J., Horritt, M.S., 2000. Obtaining accurate maps of topography and vegetation to improve 2D hydraulic flood models. Proceedings of EOS/SPIE Symposium on Remote Sensing for Agriculture, Ecosystems and Hydrology II, Barcelona, 25–29 September, SPIE, vol. 4171. International Society for Optical Engineering, Bellingham, WA, pp. 125–136.

Estrela, T., Quintas, L., 1994. Use of GIS in the modelling of flows on floodplains. In: White, H.R., Watts, J. (Eds.), Second International Conference on River Flood Hydraulics, Wiley, Chichester, pp. 177–189.

Feldhaus, R., Höttges, J., Brockhaus, T., Rouvé, G., 1992. Finite element simulation of flow and pollution transport applied to a part of the river Rhine. In: Falconer, R.A., Shiono, K., Matthews, R.G.S. (Eds.), Hydraulic and Environmental Modelling: Estuarine and River Waters, Ashgate publishing, Aldershot, pp. 323–334.

Galland, J.C., Goutal, N., Hervouet, J.-M., 1991. TELEMAC—a new numerical-model for solving shallow-water equations. Advances in Water Resources 14 (3), 138–148.

Hervouet, J.-M., Janin, J.-M., 1994. Finite element algorithms for modelling flood propagation. In: Molinaro, P., Natale, L. (Eds.), Modelling flood propagation over initially dry areas, American Society of Civil Engineers, New York, pp. 243–256.

Hervouet, J.-M., Van Haren, L., 1996. Recent advances in numerical methods for fluid flows. In: Anderson, M.G., Walling, D.E., Bates, P.D. (Eds.), Floodplain Processes, Wiley, Chichester, pp. 183–214.

Horritt, M.S., 1999. A statistical active contour model for SAR image segmentation. Image and Vision Computing 17, 213–224.

Horritt, M.S., 2000. Calibration and validation of a 2-dimensional finite element flood flow model using satellite radar imagery. Water Resources Research 36 (11), 3279–3291.

Horritt, M.S., Bates, P.D., 2001a. Effects of spatial resolution on a raster-based model of flood flow. Journal of Hydrology 253, 239–249.

Horritt, M.S., Bates, P.D., 2001b. Predicting floodplain inundation: raster-based modelling versus the finite element approach. Hydrological Processes 15, 825–842.

Horritt, M.S., Mason, D.C., Luckman, A.J., 2001. Flood boundary delineation from synthetic aperture radar imagery using a statistical active contour model. International Journal of Remote Sensing 22 (13), 2489–2507.

Knight, D.W., 1989. Hydraulics of flood channels. In: Beven, K., Carling, P. (Eds.), Floods: Hydrological, Sedimentological and Geomorphological Imolication, Wiley, Chichester, pp. 83–105.

Knight, D.W., Shiono, K., 1996. River channel and floodplain hydraulics. In: Anderson, M.G., Walling, D.E., Bates, P.D. (Eds.), Floodplain Processes, Wiley, Chichester, pp. 139–181.

Marks, K., Bates, P.D., 2000. Integration of high resolution topographic data with floodplain flow models. Hydrological Processes 14, 2109–2122.

Mason, D.C., Cobby, D.M., Davenport, I.J., 1999. Image processing of airborne scanning laser altimetry for some environmental applications. Image and Signal Processing for Remote Sensing V, SPIE, vol. 3871. International Society for Optical Engineering, Bellingham, WA, pp. 55–62.

Romanowicz, R., Beven, K.J., 1997. Dynamic real-time prediction of flood inundation probabilities. Hydrological Sciences Journal 43, 181–196.

Romanowicz, R., Beven, K.J., Tawn, J., 1996. Bayesian calibration of flood inundation models. In: Anderson, M.G., Walling, D.E., Bates, P.D. (Eds.), Floodplain Processes, Wiley, Chichester, pp. 333–360.

Smith, L.C., 1997. Satellite remote sensing of river inundation area, stage and discharge: a review. Hydrological Processes 11 (10), 1427–1439.