Advances in Water Resources 31 (2008) 1038–1047

Contents lists available at ScienceDirect

Advances in Water Resources

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

# Estimating uncertainty associated with water stages from a single SAR image

Guy Schumann [a,b,*], Florian Pappenberger [c], Patrick Matgen [a]

[a] Public Research Centre – Gabriel Lippmann, Belvaux L-4422, Luxembourg
[b] School of Geographical Sciences, Bristol BS81SS, UK
[c] European Centre for Medium-Range Weather Forecasts, Reading RG29AX, UK

ARTICLE INFO

*Article history:*
Received 27 January 2008
Received in revised form 31 March 2008
Accepted 11 April 2008
Available online 29 April 2008

*Keywords:*
Remote sensing
Water stages
Uncertainty
Data value
Model acceptability
Stepped GLUE

ABSTRACT

It is the goal of remote sensing to infer information about objects or a natural process from a remote location. This invokes that uncertainty in measurement should be viewed as central to remote sensing. In this study, the uncertainty associated with water stages derived from a single SAR image for the Alzette (G.D. of Luxembourg) 2003 flood is assessed using a stepped GLUE procedure. Main uncertain input factors to the SAR processing chain for estimating water stages include geolocation accuracy, spatial filter window size, image thresholding value, DEM vertical precision and the number of river cross sections at which water stages are estimated. Initial results show that even with plausible parameter values uncertainty in water stages over the entire river reach is 2.8 m on average. Adding spatially distributed field water stages to the GLUE analysis following a one-at-a-time approach helps to considerably reduce SAR water stage uncertainty (0.6 m on average) thereby identifying appropriate value ranges for each uncertain SAR water stage processing factor. For the GLUE analysis a Nash-like efficiency criterion adapted to spatial data is proposed whereby acceptable SAR model simulations are required to outperform a simpler regression model based on the field-surveyed average river bed gradient. Weighted CDFs for all factors based on the proposed efficiency criterion allow the generation of reliable uncertainty quantile ranges and 2D maps that show the uncertainty associated with SAR-derived water stages. The stepped GLUE procedure demonstrated that not all field data collected are necessary to achieve maximum constraining. A possible efficient way to decide on relevant locations at which to sample in the field is proposed. It is also suggested that the resulting uncertainty ranges and flood extent or depth maps may be used to evaluate 1D or 2D flood inundation models in terms of water stages, depths or extents. For this, the extended GLUE approach, which copes with the presence of uncertainty in the observed data, may be adopted.



## 1. Introduction

Due to the distributed nature of its data, the large spatial coverage and the temporal frequency of data acquisition, remote sensing is believed to have the potential to solve many problems in many fields of environmental science. For hydrology for example, it may solve issues of scale and data scarcity in the spatial as well as temporal domain; it may redefine our knowledge of processes due to an increase in data availability and information; it may indicate areas where to model differently or where to measure more frequently; it may disclose new information and thus lead to an improved understanding of the behaviour of a system or process. Briefly, remote sensing may be considered, and has been so by some authors [39,7,16,47,19,1], as a solution to many problems in hydrology.

Review papers on integration of remote sensing and hydrology/hydraulics exist within the literature [46,9,32,38,33,2]. However, to date, only relatively few studies have integrated remote sensing and hydraulics. Most notable studies used remote sensing, in particular synthetic aperture radar (SAR) due to its all-weather and day and night capabilities as well as the specular reflection of microwaves from open smooth water bodies, to evaluate 1D [23,17,29,42] or 2D [30,17,8,4] flood inundation models. Most of these studies use the generalised likelihood uncertainty estimation (GLUE [12]) framework. The extended GLUE approach [11] accounts for uncertainty in both the model and observation data. It has also been successfully demonstrated that remote sensing can be used to derive useful spatial 3D flood information, such as flood extent and depth maps [24,44] which may be integrated with models to improve their hydraulic functioning [43]. Other studies have utilised SAR to improve flood risk assessment through hydraulic modelling [29].

Major issues, such as incommensurability in data and scale, difficulties in data and image availability, a long and sometimes exhaustive processing chain of remote sensing data products prior

* Corresponding author. Address: School of Geographical Sciences, Bristol BS81SS, UK. Tel.: +44 117 928 9954; fax: +44 117 928 7878.
*E-mail address:* guy.schumann@bristol.ac.uk (G. Schumann).

0309-1708/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.advwatres.2008.04.008

to information retrieval as well as the large amount of uncertainty associated with estimated variables from remote sensing, hinder the integration of remote sensing with hydrology or hydraulics. In brief, with remote sensing being a proxy for so many environmental variables, different knowledge and skills are required to make most use of it. The associated uncertainty with such proxy is probably one of the most important factors in the field of remote sensing and geographical information science and has been a major topic in both these fields for over a decade [5,6,49]. Uncertainties may arise from the quality of appropriateness of the data used to describe the system, from aggregation (temporal and spatial) and simplification as well as from lack of data and approximation [36].

In this context, it is important to establish the information content for constraining models if all sources of uncertainty are acknowledged and incorporated. As a prerequisite of this, it is necessary to estimate the uncertainty and try to constrain it most effectively. According to Rotmans and Asselt [36], new knowledge on complex processes may reveal the presence of uncertainties that were previously unknown or were understated. In this sense, it is important to associate some level of uncertainty with the new information derived from remote sensing. This has been attempted in a flood inundation study by Schumann et al. [42] who have tried to estimate the uncertainty in water stages from SAR imagery resulting from inaccuracies in flood boundary position. They have used that information to evaluate an uncertain flood inundation model within an extended GLUE framework [11], which copes with the presence of uncertainty in the observed data.

As an extension of the first part of the study by Schumann et al. [42], it is the aim of this study to propose a methodology for assessing and estimating the uncertainty associated with water stages from a simple flood model based on remote sensing processing [44]. Uncertainty is defined, expressed and viewed in many different ways throughout the literature. This study focuses on the uncertainty resulting from a lack of knowledge about model parameters and their values. This kind of uncertainty may be particularly important when models of any kind are derived from remote sensing [6,15].

Model uncertainties may be estimated in many different ways (see [37]). The GLUE procedure mentioned earlier accounts for errors in both the model and observation data. It is a paradigm for evaluating complex environmental models that favours the concept of 'equifinal' parameter sets over the 'optimum' parameter set [13]. From a prior distribution of feasible model parameters multiple simulations are performed and evaluated with some data according to some likelihood or performance measure. A reasonable threshold definition based on the modelled system in question decides on the acceptability of a model simulation. From this 'behavioural' posterior distribution of parameters, cumulative distribution functions (CDFs) may be generated for the model output based on the likelihood of acceptability. The outcome of GLUE is an estimation of the model output uncertainty as a result of parameter uncertainty.

It is worth noting that GLUE is by no means perfect (e.g., the often subjective choice of a threshold value of some likelihood measure, i.e., the 'cutting-off' of the model parameter space to generate the posterior distribution; the absence of a formal statistical likelihood measure; all uncertainty is related to parameter uncertainty; not really giving the predictive uncertainty of the model) and has been subject to discussion recently [20,14,21]. Nonetheless, it is a widely used and easy-to-implement technique which allows constraining and estimating model uncertainty related to uncertain model parameters and seems therefore suitable for this study. Moreover, Beven et al. [14] show that the arguments on which the discussion of Mantovan et al. [20] is based are ill-defined. In most real world cases, particularly in flood inundation studies, the definition of formal likelihoods is hardly possible, as it requires assumptions about the nature of errors that are rarely known or justified.

During a first stage, appropriate bounding [26] of important SAR flood model processing parameters is undertaken (i.e., setting an appropriate interval for parameter values) and the model output uncertainty is estimated within a GLUE framework. Thereafter, constraining of this uncertainty is attempted by introducing spatially distributed field water stages one at a time. This stepped GLUE procedure allows (a) to assess the value of each individual field-recorded water stage to constrain SAR water stage uncertainty and (b) to determine whether all field water stages are needed to achieve maximum constraining thereby implicitly addressing the following question: 'Where and how much would we need to sample in the field to achieve successful model uncertainty constraining most efficiently?'.

## 2. Study site and available data

### 2.1. Study site

The uncertainty associated with water stages from a single SAR flood image is estimated for a 10 km reach of the Alzette River north of Luxembourg City (G.D. of Luxembourg) which experienced a high magnitude flood event in early January 2003 (Fig. 1). The Alzette catchment has an area of 1175 $km^2$, of which the study reach drainage area represents 34% (404 $km^2$). It is characterised by a relatively large and flat floodplain compared to most of the rest of the river, which cuts through steep, narrow valleys. The reach floodplain has an average width of approximately 300 m. The stream channel has an average depth of around 4 m and an average slope of 0.08%. The villages along this stretch have been subject to frequent flooding in the past two decades. The investigated flood event of January 2, 2003 has a return period of five years with a peak discharge of 75 $m^3\ s^{-1}$.

### 2.2. Available data

The event was acquired by the ASAR instrument onboard the ENVISAT satellite at the time of flood peak with C-band (5.3 GHz) in VV-VH mode and an incidence angle of 35° (Fig. 1). SAR, the wavelengths of which are reflected away from the antenna by smooth water bodies, seems especially promising for flood disaster management due to its all-weather as well as day and night capability, which is particularly useful in areas of rapid flood recession. The spatial ground resolution of the acquired ASAR image is 25-30 m, which has been downscaled to 12.5 m in pixel resolution using multi-look filtering [28].

Also, during a field campaign, several evenly distributed water stages were collected in the field (Fig. 1). Field water stages were also automatically recorded at three bridges. Moreover, a 2 m LiDAR DEM with a vertical accuracy of 15 cm was assembled for the floodplain and supplemented with ground-surveyed evenly spaced river channel cross sections. The DEM corresponds to a bare ground DEM after removal of individual buildings, structures and occasional high vegetation. The LiDAR was flown in late winter and thus deteriorating effects commonly caused by short grass or pasture (which is the main cover type of the study area) are assumed negligible.

## 3. Methodology

This section briefly outlines how water stages can be derived from a single remote sensing flood image and describes the different steps involved in the estimation of the uncertainty associated

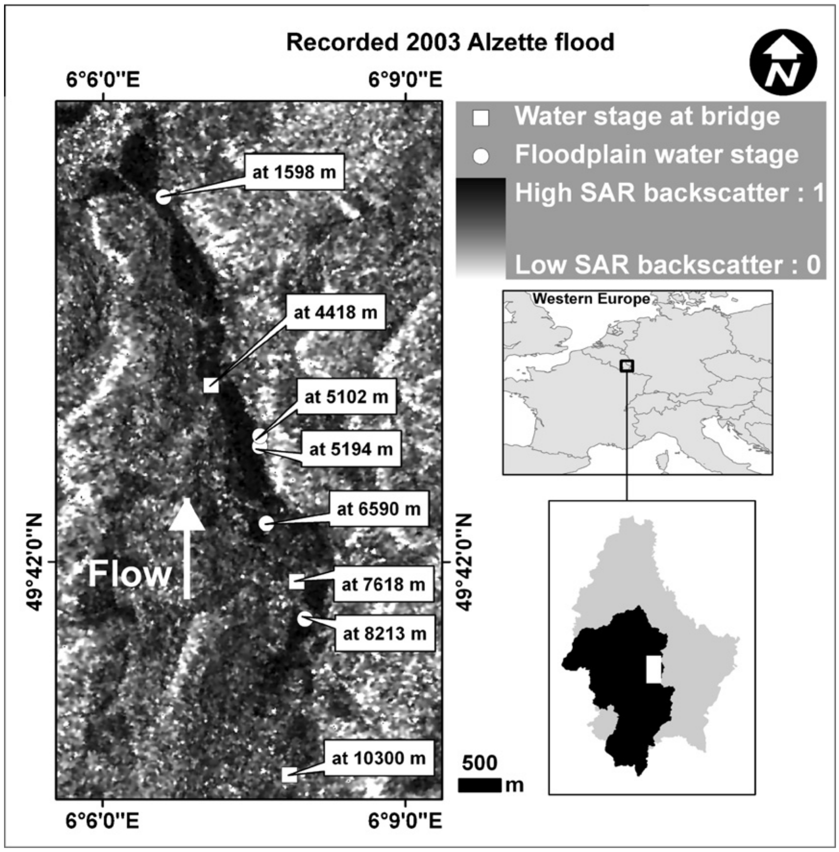


**Fig. 1.** Study site showing the ASAR VH flood image as well as the location of surveyed field water stages with associated chainage from the downstream boundary condition.

with those water stages. Notice should be given here that knowledge about parameters is defined by upper and lower bounds on 'plausible' (see Section 3.2.1 for further explanation) value ranges for the different model parameters. Parameter bounding is thought appropriate in this study given the simple structure of the remote sensing model [26]. The bounds set in this study are a function of the SAR image processing complexity to merely avoid irrealistic model outputs from an image processing perspective. After estimating water stage uncertainty within a GLUE framework, field water stages are added one at a time to constrain the uncertainty using a stepped GLUE approach.

### *3.1. Deriving water stages from single flood imagery*

A known technique to derive water stages from a single flood image is to extract the heights from a high resolution DEM at the flood extents [27]. However, as SAR is often a preferred mode of acquisition during a flood event, given the specular reflection of open water bodies resulting in dark tones on the image, but the imagery suffers from considerable distortion [34], having to cope with uncertainty in height extraction at the often blurred flood boundaries is inevitable. Accounting for most of these uncertainties, Schumann et al. [44] proposed a Regression and Elevation-based Flood Information eXtraction model (REFIX) that applies linear, piecewise linear or non-linear regression modelling to the extracted water heights. For the 2003 Alzette flood, they suggested a simple linear regression equation (1) to reliably model the water surface slope and to estimate water stages at 55 cross sections along the river, in an upstream to downstream direction and perpendicular to the direction of flow. The reported root mean squared error (RMSE) accuracy was as low as 0.18 m. Although, the REFIX model is a simple river stage model and so might only be valid for simple stream networks, it has been demonstrated that these data can be successfully used to generate flood depth maps and improved flood extent maps [44]. Moreover, the remotely sensed water stages can be employed to evaluate flood inundation models within a GLUE framework [42] or to improve the functioning of such models [43].

The REFIX regression model for the Alzette reach as proposed by Schumann et al. [44] is of the form

$$H_{\mathrm{R}} = a \cdot d + b \tag{1}$$

where $H_{\mathrm{R}}$ is the REFIX modelled water stage in m above sea level, $d$ is the downstream distance along the stream centreline in m, $a$ is the regression slope (m/m) and represents the average gradient of the water surface and $b$ is the regression intercept (m), i.e., the upstream boundary condition in SAR water stage. It is worth noting that the regression coefficients are not set by field data but are conditioned on the SAR water stages extracted at the flood boundaries of river cross sections drawn perpendicular to the direction of flow.

In this study, the REFIX model is applied to estimate water stages along the Alzette reach for the 2003 flood with the aim to estimate the associated uncertainty.

### *3.2. Spatial patterns of remotely sensed water stage uncertainty*

It is now well established that uncertainty may be quantified using a Monte Carlo technique [37] such as GLUE. This involves comparison with field data and the definition of some performance measure threshold above which model outputs are acceptable [12,13,11]. Setting a threshold (or behavioural definition) allows the different REFIX model outputs to be qualified as behavioural, $B$, if the REFIX performance does not exceed that threshold, and non-behavioural, $\overline{B}$, otherwise. As the REFIX output is distributed in space, its associated uncertainty can be estimated and visualised spatially.

*3.2.1. Performing multiple model simulations*

Five different input factors (or REFIX processing parameters) are considered as contributing most to the uncertainty of the water stage output:

- geolocation accuracy, $X_G$, i.e., the positional accuracy of the image on the ground;
- image filter window size, $X_F$, i.e., the size of a moving window, given by the number of pixels, used to remove random noise (speckle) in a SAR image (see 'salt and pepper' effect on the dark water areas in Fig. 1);
- image thresholding value, $X_T$, the value on the image greyscale used to classify areas as flooded;
- vertical precision of the DEM (assessed by adding random errors of a given variance to the DEM values), $X_E$;
- number of river cross sections at which water stages are derived, $X_S$, i.e., number of data points for regression modelling.

It is worth noting that the uncertainty level of the regression model in terms of its parameters (slope and intercept) is not examined in this study. Uncertainty of the regression parameters may be expressed by confidence limits which can be easily and more directly calculated using the sum of the squares of the residuals from the regression and Monte Carlo analysis is thus not needed. Moreover, an autocorrelation test as well as a test on the order of residuals from the regression model used by Schumann et al. [44] suggest that a linear model is indeed the most adequate equation for the Alzette reach. Furthermore, it is expected that the SAR water stage uncertainty associated with the above listed factors will encompass the regression parameter uncertainty.

For each input factor, an appropriate range of values is chosen from which a large number of REFIX simulations are performed within a Monte Carlo environment (Table 1). It is worth noting that the resulting uncertainty values are modelled using Monte Carlo simulation and a particular sampling design (see Table 1). The one-factor-at-a-time (OFAT) sampling design used in this study is in itself an 'intuitive uncertainty reduction', in the sense that only plausible REFIX simulations are considered and highly irrealistic ones (e.g., a geolocation inaccuracy of more than eight image pixels is hardly likely or an irrealistic flood masking value close to white on the greyscale sampling range) are omitted. Table 2 gives the values of the factors that remained unchanged when varying only one factor at a time during the OFAT analysis. As only one factor is varied at a time, the order of the factor variation is negligible. Hence, factor interactions cannot be detected and would therefore require additional investigation. The values listed are those used by Schumann et al. [44]. More details on these values can be found in their paper.

Another point worth noting is that although the Monte Carlo analysis setup for most factors is self-explanatory, the geolocation error (or flood boundary position uncertainty) is modelled using an automated Monte Carlo-based algorithm proposed by Schumann et al. [40] that performs a high number of shifts of remotely sensed flood boundaries in each geographical direction. In this study, an ASAR image subset constrained to the flooded area extracted using the same image processing values as in Schumann et al. [44] is shifted 5000 times in each of the four geographical directions by a random amount up to 100 m. The starting position of the multiple image shifts is given by the left and right intersection marks between each of the 55 river cross sections (drawn perpendicular to the stream centreline [44]) and the LiDAR DEM. In other words, the $X$ coordinate together with the corresponding $Y$ coordinate of every mark along the flood extents are moved to the East (+n m) or West (−n m) and to the North (+n m) or South (−n m), respectively. So, each image shift is composed of a pair of $X$- and $Y$-coordinate shifts.

**Table 1**
Range of values for uncertain input factors

| Uncertain factor | Variable (unit) | Value range | Value increment | Number of REFIX simulations |
|---|---|---|---|---|
| $X_G$ | Distance (m) | [−100, 100] | Random | 5000 |
| $X_F$ | Window size (pixels) | [3, 35] | 2 | 17 |
| $X_T$ | Greyscale value (DN) | [80, 160] | 10 | 9 |
| $X_E$ | Variance (m) | [0.01, 0.3] | 0.005 | 59 |
| $X_S$ | Number of river cross sections | [2, 55] | 1 | 54 |

**Table 2**
Values set for factors kept unchanged during the OFAT analysis

| Uncertain factor | Value |
|---|---|
| $X_G$ | 1–2 Image pixels |
| $X_F$ | 5 × 5 Frost filter window |
| $X_T$ | 130 on the greyscale |
| $X_E$ | A vertical precision of ±0.15 m |
| $X_S$ | 55 River cross sections |

*3.2.2. Setting an acceptability criterion*

Defining an objective and appropriate model acceptability threshold within GLUE is not an easy task [11]. There are many different ways this can be done. The easiest way is probably to (subjectively) define an adequate model performance value above or below which models are retained or rejected. Justifying the adequacy of this choice is for obvious reasons hardly possible. Another way to decide on acceptable models is to set observational uncertainty limits inside which model simulations are required to fall (this can be done using fuzzy measures [23,30] or minimum and maximum bounds [42] within, e.g., the extended GLUE approach suggested by Beven [11]). Although the extended GLUE approach has the advantage of making the 'choice' of the limits of acceptability more objective, it is rather difficult to set appropriate uncertainty limits for observations, and again these are likely to be prone to a great deal of subjectivity. It is worth noting that subjectivity in choosing thresholds may be impossible to avoid but may be justified and adequate, as it is most often based on experience and knowledge.

As an extension to the extended GLUE, a benchmark model could be defined as an improved threshold against which to compare model outputs. In rainfall–runoff modelling this kind of comparison is implicitly applied when using the Nash–Sutcliffe efficiency criterion, where model outputs are tested against the mean of the observations over a certain time period to be acceptable (where acceptability may be defined as >0 [18] with 1 being a perfect fit). In this context, Seibert [45] calls for alternative benchmarks in hydrological modelling and proposes to simply substitute the mean of observations in the Nash criterion by some other kind of benchmark (for runoff modelling, Seibert [45] suggests e.g., a long-term seasonal variation instead of one constant value, or the observed runoff shifted by several time steps). In this study, this acceptability approach is adopted and the Nash efficiency is reformulated in Eq. (2) for the space domain rather than the time domain by comparing the spatially distributed water stages from the REFIX to those generated by a benchmark model for the 2003 Alzette flood:

$$E_{\text{bench}} = 1 - \frac{\sum_{i=1}^{n}(H_{O_i} - H_{R_i})^2}{\sum_{i=1}^{n}(H_{O_i} - H_{B_i})^2} \tag{2}$$

where $E_{\text{bench}}$ is the proposed efficiency criterion, $H_O$ is the water stage observed in the field in m above sea level at river cross section

$i$, $H_R$ is the REFIX modelled water stage and $H_B$ is that given by the benchmark model in Eq. (3). It is worth noting that the measure proposed by Seibert [45] is for time series, whereas in this study $E_{bench}$ is used for comparing data over space.

A suitable benchmark model for the studied reach is defined as a linear regression model that uses only the water stage measured at the upper boundary and the surveyed average river bed gradient (−0.08%). A very similar simple model has been successfully applied by Wang et al. [48] using gauged upstream and downstream water stages to infer the flow gradient. Reliable trends of a river gradient may even be reliably derived from coarse resolution DEMs (e.g., SRTM DEM) as demonstrated on the same Alzette river reach by Schumann et al. [41]. It is worth noting that $H_B$ in Eq. (2) can be obtained from any other type of appropriate benchmark just as easily as $H_R$ may be modelled with another type of flood model (e.g., 1D or 2D hydrodynamic model). The benchmark model equation used in this study for the 2003 Alzette flood is of the form

$$H_B = -0.0008 \cdot d + 226.31 \tag{3}$$

where $H_B$ is the water stage in m above sea level simulated by the benchmark model, the slope of the regression model is the surveyed average river bed gradient (−0.08%) in m/m and the regression intercept is the upstream boundary condition and is represented by the water stage automatically recorded at the bridge at 10,300 m (Fig. 1). This simple model has a root mean squared error (RMSE) of 48 cm when compared to all field water stages. To recall, the benchmark model used for evaluation with Eq. (2) is purely field based whereas the SAR model in Eq. (1) is purely remote sensing based.

Although Schumann et al. [44] have shown that the REFIX model performs much better than the simple benchmark model (18 cm compared to 48 cm), weighting a SAR-based flood model against this simple field-based benchmark allows to highlight the following:

- A single SAR or other remote sensing flood image combined with a high resolution DEM provides useful and accurate information about hydraulic variables.
- Water stages derived from a single remote sensing flood image provide additional information to water stages simulated using gauged water stages at the boundary condition in combination with the surveyed average river bed gradient.
- The distributed nature of remote sensing may be regarded as an additional source of hydraulic information to be used in conjunction with other models when no spatially distributed field data are available.

*3.2.3. Estimating SAR water stage uncertainty*

During a first stage of the analysis, the uncertainty associated with SAR water stages is estimated within a GLUE framework using a weighting of 1 for all REFIX model simulations from the plausible parameter ranges. Then, during a second stage, after definition of a behavioural threshold, each simulation from the feasible parameter space can be classified into behavioural and non-behavioural outputs using Eq. (2). According to $E_{bench}$ a REFIX model simulation is accepted if it is a better predictor of water stages than the benchmark model (Fig. 2), compared to the field water stages in Fig. 1. In other words, during the stepped GLUE analysis all REFIX simulations with a performance value $\leqslant 0$ are rejected. A stepped version of GLUE is adopted to assess the uncertainty constraining power of each water stage recorded in the field.

All acceptable outputs are weighted according to their performance and a cumulative distribution function (CDF) is then generated. This technique is also used within the GLUE framework to assess the uncertainty of environmental models and allows the

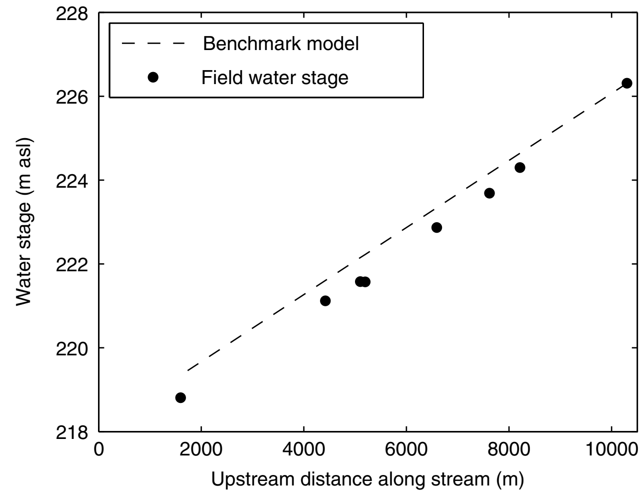


**Fig. 2.** Visualising the benchmark model performance compared to field water stages.

derivation of uncertainty quantiles from the generated CDF. In hydraulic modelling, for example, 2D or even 3D maps may be produced from these quantiles to visualise the uncertainty associated with flood extent (2D) and/or depth (3D).

This procedure is applied to the REFIX model in this study to allow for the quantification as well as visualisation of the uncertainty associated with remote sensing-derived water stages and flood extents and depths. Water stage uncertainty will be represented graphically as a range delimited by the 5th and 95th quantiles of the water surface lines whereas flood extent and depth uncertainty will be visualised by generating 3D maps of the 5th and 95th quantile. The maps are produced in a GIS environment following the REFIX mapping procedure [44], whereby water stages at each river cross section (in this study, these will be the water stage uncertainty values for each cross section) will be used to generate a triangular irregular network (TIN) mesh of the water surface of the flooded reach. As a final processing step, this TIN is subtracted from the DEM to derive flood extent and depth.

## 4. Results of the uncertainty analysis

This section reports the results of the quantitative assessment of the uncertainty associated with the REFIX model output. During a first analysis stage, plausible parameter bounds for uncertain model factors are determined and used to generate multiple REFIX simulations that are all equally weighted in a GLUE analysis. During a second stage, field data are introduced to the analysis one by one using the proposed $E_{bench}$ efficiency criterion (Eq. (2)). Thereafter, all acceptable simulations ($E_{bench} > 0$) are performance weighted to estimate and visualise the uncertainty of remotely sensed water stages.

It is important to note that a very specific water stage sampling plan for the field has been assumed for the stepped GLUE analysis and reflects the most likely approach in an 'intuitive' sampling campaign. Within the reach there are two bridges other than the one positioned at the upstream boundary at 10,300 m (Fig. 1). The latter, however, cannot be added as a singleton to the analysis, as the benchmark model is conditioned on that gauged water stage (and so, according to Eq. (2), all REFIX simulations will inevitably be rejected when using this measured water stage). The automatically-gauged water stages at the two remaining bridges would of course be used first before attempting any additional sampling in the field. It was also assumed that the bridge furthest downstream would be added to the analysis first, so as to keep water stages more or less evenly distributed over space when constraining REFIX model uncertainty. After the two bridges have been added, sampling in the field would probably start at the lower boundary condition of the reach before moving gradually upstream the river gradient whilst ensuring sampled water stages are at equidistance.

Before assessing the uncertainty of the REFIX model, an appropriately large number of simulations were performed for each uncertain input factor using the defined parameter value ranges (Table 1). Uncertainty was estimated within a GLUE framework using a uniform performance value weighting of 1 for each simulation. The results are shown in Fig. 3a which plots the 5th and 95th uncertainty quantiles. As can be seen on the graph, uncertainty in SAR-derived water stages assuming no prior knowledge about SAR image processing is large, with 2.8 m on average. This uncertainty was considerably reduced to 0.6 m on average by introducing in a step-wise manner field water stages to the analysis and by weighting the simulations each time to the benchmark model conditioned on field observations (Eq. (3)).

Actual errors of the SAR-based model, which are potentially interesting, can be given by an RMS accuracy between the best estimate (median value) and the field water stages. A target value range may be defined by the uncertainty in field stages, set at [10 cm, 30 cm] by Schumann et al. [43]. Comparing the constrained median (Fig. 3d) to the remaining four field stages not used in the constraining gives an RMS accuracy of 13 cm, which is situated close to the lower bound of the target range. Comparing the unconstrained uncertainty median (Fig. 3a) to all water stages in the field gives still a reasonable value of 47 cm, although outside the target range. This illustrates clearly the utility of field data (to constrain uncertain data sets) but at the same time also highlights the potential of remote sensing in case of no field data.

All field water stages have been added according to the design sampling plan described earlier and uncertainty constraining results for the REFIX model following the stepped GLUE procedure are shown in Fig. 3, graphs b–h. As a final step of the GLUE procedure, the 5th and 95th percentiles in Fig. 3 are taken to create uncertainty maps in terms of flood extent following the REFIX flood extent and depth mapping procedure [44]. Fig. 4 shows flood extent uncertainty maps before and after constraining REFIX model uncertainty with field water stages.

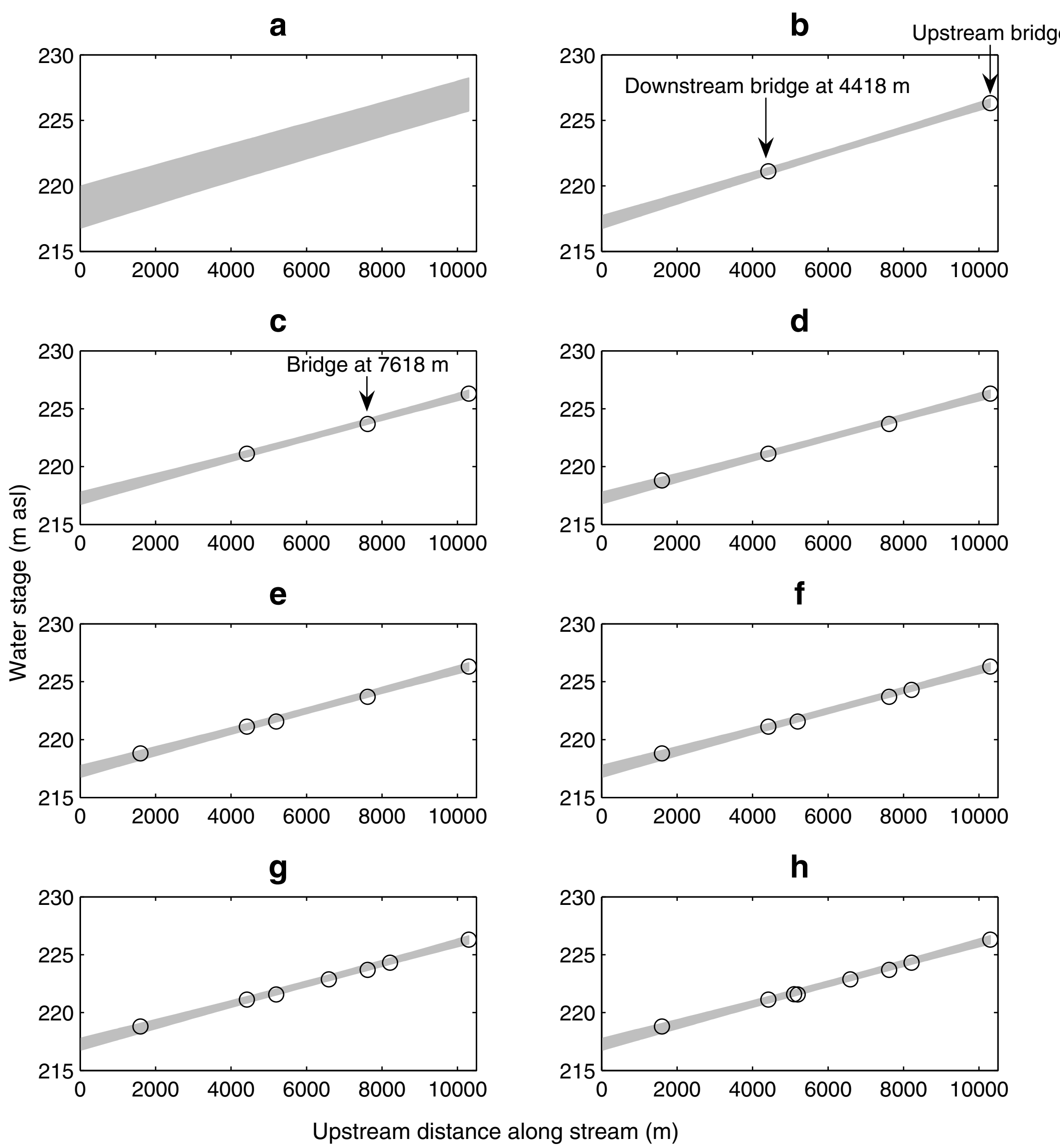


**Fig. 3.** The 5th and 95th uncertainty quantile ranges as well as the location and number of additional field water stages used to constrain remote sensing-based flood modelling. Graphs show uncertainty with no field data constraining, i.e., assuming no or very little prior knowledge about SAR image processing (a) and uncertainty constraining with two (b), three (c), four (d), five (e), six (f), seven (g), and all eight (h) field water stage locations. The exact location of each field water stage is shown in Fig. 1.

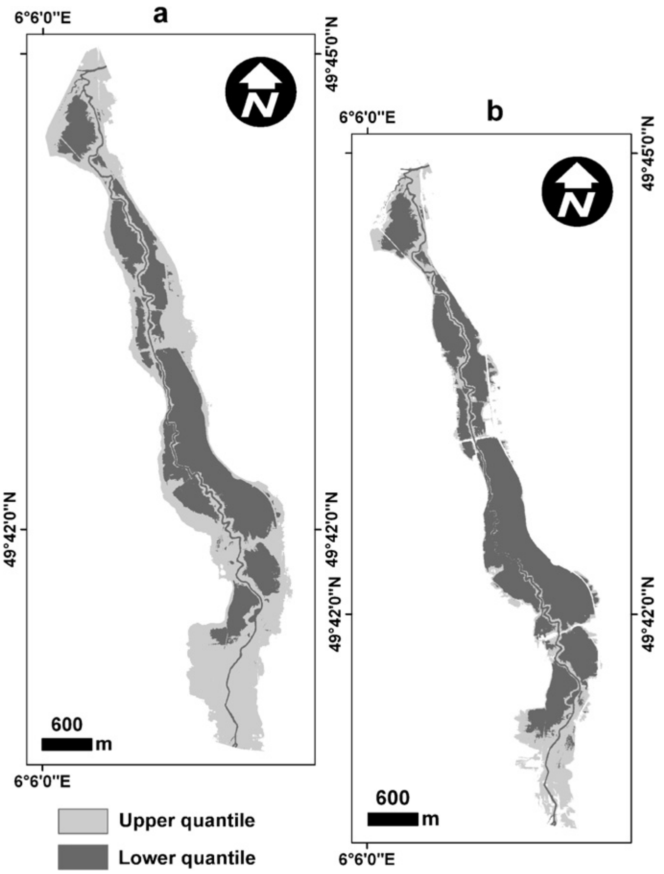


**Fig. 4.** Flood extent uncertainty maps of the 95th and 5th quantiles for SAR water stage uncertainty assuming no prior knowledge about remote sensing processing (no field data constraining) (a) and maximum constraining achieved with only three out of seven additional field stages (b).

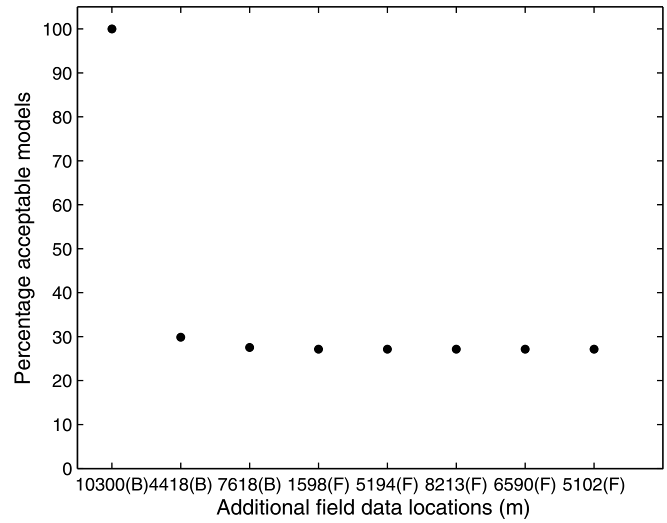


**Fig. 5.** Constraining SAR water stage uncertainty by adding field data. (*N.B.* Adding more than three field water stages does not have any additional constraining effect.)

From the graphs b–h in Fig. 3 it is difficult to see how many field water stages are necessary to achieve maximum constraining and where these water stages were recorded in the field. This information is provided in Fig. 5, which plots the constraining power of each additional water stage in terms of percentage of models accepted. From the plot in Fig. 5 it is obvious that adding only the water stage gauged at the bridge most downstream, uncertainty of the REFIX model is drastically reduced (>70%). Adding the remaining bridge does nearly lead to maximum constraining. Only the location most downstream has additional very minor constraining power but reducing the number of acceptable models by only additional 0.39% may be negligible.

The fact that only water stages gauged at the bridges (one upstream at 10,300 m, one at 7618 m and one downstream at 4418 m) are necessary to achieve maximum constraining of the SAR model uncertainty is primarily due to a combination of the linearity of the reach in hydraulic terms (Fig. 2) and the structure of the REFIX model. The linear regression of the REFIX model only requires one water stage at approximately the upper boundary and one at the lower boundary of the reach to be well conditioned, whilst the third bridge does merely fine-tune the regression model. Moreover, it was shown that the REFIX easily outperforms the benchmark model and so additional water stages are unnecessary to achieve a (moderate) performance of only >0, which is the acceptability threshold in this study. It is worth noting though that when searching for higher model performances, additional water stages are expected to be necessary. However, this could eventually lead to all models being rejected (see [42]).

Constraining uncertainty allowed not only to assess the value of each additional water stage sampled in the field but also enabled identification of appropriate values for each uncertain REFIX processing factor that provides adequate model performance. Fig. 6 plots all model simulations for each uncertain model input factor before and after maximum uncertainty constraining.

In order to increase its performance, the REFIX model reacts strongly to the geolocation accuracy ($X_G$), the size of the edge-preserving (Frost) filter window ($X_F$) and the image thresholding value ($X_T$). This is not surprising, as these factors are considered crucial parameters in a SAR image processing chain. In flood mapping, they determine the positional accuracy of flood boundaries which is crucial when overlaying these with a DEM to model water stages. Different geolocations as well as different greyscale flood masking values may change the position of the flood boundaries dramatically with adverse effects on water stage estimations. Also, an inappropriate filter window size, particularly larger ones, even with an edge-preserving filter [22] has deteriorating blurring effects on flood boundaries and as a result positional accuracy is decreased. The number of cross sections ($X_S$) used to model water stages with remote sensing seems far less important. The quasi-linear behaviour of the investigated reach may be reliably approximated by a linear regression model conditioned on only very few data points (Fig. 2). The vertical precision of the LiDAR DEM ($X_E$) seems also less important. Relatively small errors in this data set are smoothed out by the regression and so a moderate variation in LiDAR precision (up to 30 cm maximum) does not alter the regression line considerably. For a sensitivity study on DEMs with coarser resolutions, both in the vertical and horizontal dimensions, see the study by Schumann et al. [41].

Identified appropriate values for the three most sensitive REFIX processing factors for the 2003 Alzette flood are (preferably):

- $X_G$: an accuracy of <50 m (i.e., 2 ground resolutions). It is worth noting here that changes in the *Y* coordinate (in flow direction in this study) are much less impacting on the results of a SAR-based 1D flood model than changes in *X*, due to the S–N orientation of the floodplain (minor changes in *X* may very rapidly result in considerable changes in elevation values extracted at the flood boundaries with possible adverse effects on model output, especially if these positional changes are near suddenly sloping terrain [24]).

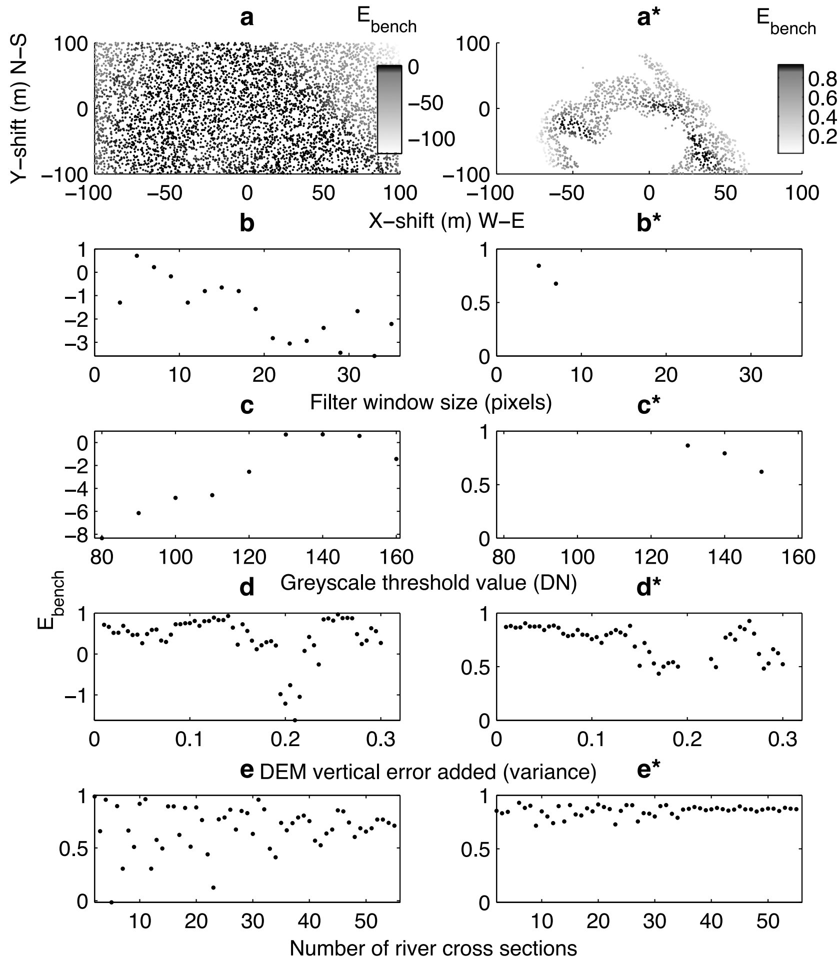


**Fig. 6.** Scatterplots on the left show all the REFIX simulations with the corresponding $E_{bench}$ values for each uncertain input factor: $X_G$ (a); $X_F$ (b); $X_T$ (c); $X_E$ (d); $X_S$ (e). Scatterplots on the right show all REFIX simulations that are retained after maximum constraining with field water stages has been achieved.

- $X_F$: an edge-preserving filter window size of either 5 × 5 or 7 × 7 image pixels.
- $X_T$: a greyscale threshold value between 130 and 150.

It is noteworthy that although a relatively large number of REFIX model simulations were accepted for $X_G$, $X_E$ and $X_S$ even after maximum constraining with field data, constraining the value ranges of these three parameters further may become very important when higher performances (i.e., $E_{bench}$ approaching 1) of the model are required. So, for example for the $X_G$ factor, Schumann et al. [40] recommend a geolocation accuracy of <1 ground resolution to achieve a REFIX RMSE of <20 cm when evaluated with field data. In addition, it is clear that for the $X_E$ factor, a DEM with very low errors in the vertical should be used to get to that kind of REFIX accuracy [44] and also the number of river cross sections used in the regression model, i.e., factor $X_S$, becomes more important to achieve a better REFIX performance.

Although the results presented here are specific for the ASAR-VH image acquired over the Alzette, the methodology proposed to estimate the uncertainties associated with remote sensing-derived water stages can easily be applied to another data set elsewhere.

## 5. Discussion

This section will discuss two major points: (i) the value of additional field data locations to constrain model uncertainty and (ii) using the uncertainty bounds to some meaningful end.

### 5.1. The value of additional field data locations

Constraining the uncertainty associated with water stages modelled from a single SAR flood image by introducing field data using a stepped GLUE approach has enabled to assess the value of each additional water stage location in the field. It has been argued that a combination of the hydraulic complexity or simplicity of a reach and the structural properties of the REFIX or model topology determine the value of additional water stage locations. Thus, the value of each water stage location is expected to change with another model structure and also with a different performance (or likelihood) measure. Although this is particularly important for events that experience considerable backwater effects or for rivers where hydraulically 2D water stages are more important during floods (i.e., for events or river reaches with an increased hydraulic complexity), it is worth noting that this is expected to a much lesser extent for the Alzette flood event investigated due to the linear behaviour of the field water stages during this particular event (Fig. 2) that may be reliably approximated by a simple linear regression model. However, for events of higher magnitude with significant backwater effects, the additional water stage locations in this study that had no constraining power for this particular event may prove important.

The findings of the stepped GLUE procedure implicitly suggest the need of a strategic sampling plan for field water stage locations. A possible sampling guidance could be that water stage locations need to be tied to the structural properties of the model and the evaluation criterion used and the nature of the reach under investigation (a similar suggestion is made by Schumann et al. [42]). This means that it is crucial to recognise that a 'point of diminishing returns' [10] is associated with field data or any other type of data used for model uncertainty constraining. Given simplification and thus imperfection of models, there is always a threshold beyond which gathering more data for evaluating a given model will prove unnecessary. From this, it may be concluded that the first step of any modelling exercise would be to select an appropriate model type and performance measure or objective function that meet the needs of the user [35]. Thereafter, the model requires adequate evaluation data, the amount and spatial allocation of which is defined by the model topology. Furthermore, it could be that there is simply not more information content when sampling beyond a certain number of locations due to the nature of the region or event. Thus, the sampling design has to be optimum not only for the model but also for the area and expected event. In this context, the entire model selection and evaluation exercise is a continuous process between modeller and user [35].

In the case of the Alzette 2003 flood event, it has been shown that a very simple linear model may already be sufficiently appropriate to the user and so only limited field data (two marks at the model boundaries and one additional for fine-tuning) are required to evaluate the model and to achieve sufficient uncertainty constraining using a fit-for-purpose model acceptability threshold. It is clear that the REFIX model may be inadequate for representing dynamic flood processes in the floodplain but is certainly highly valuable to describe flood characteristics at a certain point in time (i.e., at the time of image acquisition) and in areas where no other model or data source is available [41]. Associating an adequate level of uncertainty with the REFIX model output is necessary to gain additional knowledge on model topology, value, credibility and on the information it conveys. Also, for evaluating uncertain flood inundation predictions from better performing hydrodynamic models, constrained REFIX uncertainty quantiles are believed to be useful [42]. Important to note is that a similar uncertainty constraining study could be performed with any other, more dynamic model that would be more suitable for river reaches or events of higher hydraulic complexity.

### 5.2. Making use of the SAR water stage uncertainty bounds

There is very little doubt that remote sensing will increasingly be used to monitor, map or model environmental processes from space [19]. Consequently, there will be a growing demand for appropriate evaluation of such data and estimation of the associated uncertainties. The literature on integrating remote sensing observations with other types of models is expanding rapidly. As a contribution, this study has attempted to examine the 'quality' of water stages derived from a remote sensing flood image.

On the one hand, the estimated SAR water stage uncertainty helps both the modeller and the user to establish some additional knowledge on model topology and also impacts upon the model credibility. On the other hand, the derived uncertainty intervals can be used to evaluate uncertain flood inundation predictions from hydrodynamic models [42]. A possible further extension to the extended GLUE procedure for evaluating flood inundation models may be to generate and convert flood depth uncertainty maps (as those produced in this study) into a flood depth probability map expressing levels of 'certainty' in flood depth estimation using a linear function (where 1 equals total agreement between the 5th and the 95th uncertainty percentiles and 0 where they disagree largest). This enables a weighting scheme for a fuzzy membership function [23,31,11] to evaluate hydrodynamic models. This could be further augmented by defining a split fuzzy membership function that combines this weighting scheme with one for distance to the channel. This allows the modeller to concentrate specifically on model performance in urban areas that are (often) located along the edges of a flood. However, this requires the development of a spatial performance measure that is sensitive to and targeted at subtle changes in flood depth/extent mapping. The feasibility of such an evaluation scheme is currently being investigated.

Another use of the SAR water stage uncertainty intervals could be that of assimilation in hydrodynamic models using sequential re-initialisation or 'hard' updating as proposed by Matgen et al. [25] or Kalman filtering techniques as applied by Andreadis et al. [3]. An ensemble Kalman filter could be conditioned on the remote sensing observed water stage uncertainty at each location along the stream.

## 6. Conclusion

It is clear that assessing the magnitude of uncertainty of some estimate is needed to (a) associate some level of accuracy with that variable with some degree of certainty, (b) better understand the way such data are derived, and (c) use that quantification to some meaningful ends. In this study, the uncertainty associated with water stages derived from a single SAR image for the Alzette (G.D. of Luxembourg) 2003 flood was assessed using a stepped GLUE procedure.

Main uncertain input factors to the SAR processing chain for estimating water stages include geolocation accuracy, spatial filter window size and image thresholding value. Initial OFAT results showed that even with plausible parameter values uncertainty in water stages over the entire river reach is in the order of 2 m. Adding spatially distributed field water stages to the GLUE analysis following a one-at-a-time approach helped considerably reduce SAR water stage uncertainty down to more or less half a metre (the median RMSE is as low as 13 cm), thereby identifying appropriate value ranges for each uncertain SAR water stage processing factor. For the GLUE analysis a Nash-like efficiency criterion adapted to spatial data was proposed whereby acceptable SAR model simulations are required to outperform a simpler hydraulic model based on the field-surveyed average river bed gradient. Weighted CDFs for all factors based on the proposed efficiency criterion allows the generation of reliable uncertainty quantile ranges and 2D maps that show the uncertainty associated with SAR-derived water stages.

The stepped GLUE procedure demonstrated that not all field data collected are necessary to achieve maximum constraining. A possible efficient way to decide on relevant locations at which to sample in the field would be to tie field data locations to the model's structural properties and the model acceptability criterion chosen, which in turn depends on the model performance required by the modeller or user. It is suggested that the resulting uncertainty ranges and flood extent or depth maps may be used to evaluate 1D or 2D flood inundation models in terms of water stages, depths or extents. For this, the extended GLUE approach, which copes with the presence of uncertainty in the observed data, may be adopted. Also, the uncertainty associated with SAR water stages could be used to condition an assimilation scheme for remote sensing-observed water stages.

## Acknowledgements

This study is supported by the 'Ministère Luxembourgeois de la Culture, de l'Enseignement Supérieur et de la Recherche'. Additional financial support has been provided by the PREVIEW project, a European Commission FP6 programme, and the HYDRASENS project, a study funded by the Belgian Science Policy Office (BelSPO) under the framework of the STEREO II programme. The LiDAR DEM was made available by the 'Ministère Luxembourgeois de l'Intérieur et de l'Aménagement du Territoire'.

## References


[1] Alsdorf D, Bates P, Melack J, Wilson M, Dunne T. Spatial and temporal complexity of the Amazon flood measured from space. Geophys Res Lett 34, doi:10.1029/2007GL029447.

[2] Alsdorf DE, Rodrguez E, Lettenmaier DP. Measuring surface water from space. Rev Geophys 45, doi:10.1029/2006RG000197.

[3] Andreadis KM, Clark EA, Lettenmaier DP, Alsdorf DE. Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. Geophys Res Lett 34, doi:10.1029/2007GL029721.

[4] Aronica G, Bates PD, Horritt MS. Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. Hydrol Process 2002;16:2001–16.

[5] Atkinson P, Foody GM. Uncertainty in remote sensing and GIS: fundamentals. In: Foody G, Atkinson P, editors. Uncertainty in remote sensing and GIS. Chichester: John Wiley and Sons; 2002. p. 1–18.

[6] Atkinson PM. Geographical information science: geostatistics and uncertainty. Progr Phys Geogr 1999;23:134–42.

[7] Bates PD. Invited commentary: Remote sensing and flood inundation modelling. Hydrol Process 2004;18:2593–7.

[8] Bates PD, Horritt MS, Aronica G, Beven K. Bayesian updating of flood inundation likelihoods conditioned on flood extent data. Hydrol Process 2004;18(17):3347–70.

[9] Bates PD, Horritt MS, Smith CN, Mason DC. Integrating remote sensing observations of flood hydrology and hydraulic modelling. Hydrol Process 1997;11:1777–95.

[10] Bergstrom S. Principles and confidence in hydrological modelling. Nord Hydrol 1991;22:123–36.

[11] Beven K. A manifesto for the equifinality thesis. J Hydrol 2006;320:18–36.

[12] Beven K, Binley A. The future of distributed models: model calibration and uncertainty prediction. Hydrol Process 1992;6:279–98.

[13] Beven K, Freer J. Equifinality data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J Hydrol 2001;249:11–29.

[14] Beven K, Smith P, Freer J. Comment on 'hydrological forecasting uncertainty assessment: Incoherence of the glue methodology' by Pietro Mantovan and Ezio Todini. J Hydrol 2007;338:315–8.

[15] Foody G. Uncertainty, knowledge discovery and data mining in GIS. Progr Phys Geogr 2003;27:113–21.

[16] Grayson RB, Bloeschl G, Western AW, McMahon TA. Advances in the use of observed spatial patterns of catchment hydrological response. Adv Water Resour 2002;25:1313–34.

[17] Horritt MS, Bates PD. Evaluation of 1D and 2D numerical models for predicting river flood inundation. J Hydrol 2002;268:87–99.

[18] Krause P, Boyle DP, Baese F. Comparison of different efficiency criteria for hydrological model assessment. Adv Geosci 2005;5:89–97.

[19] Lincoln T. Flood of data. Nature 2007;447:393.

[20] Mantovan P, Todini E. Hydrological forecasting uncertainty assessment: incoherence of the glue methodology. J Hydrol 2006;330:368–81.

[21] Mantovan P, Todini E, Martina MLV. Reply to comment by Keith Beven, Paul Smith and Jim Freer on 'hydrological forecasting uncertainty assessment: incoherence of the glue methodology'. J Hydrol 2007;338:319–24.

[22] Marques RCP, Carvalho EA, Costa RCS, Medeiros FNS. Filtering effects on SAR images segmentation. In: Souza JD, Dini P, Lorenz P, editors. 11th International conference on telecommunications – ICT 2004. Lecture notes in computer science. ICT No. 3124/2004. Berlin: Springer; 2004.

[23] Matgen P, Henry JB, Pappenberger F, De Fraipont P, Hoffmann L, Pfister L. Uncertainty in calibrating flood propagation models with flood boundaries observed from synthetic aperture radar imagery. In: 20th ISPRS congress. Istanbul: ISPRS; 2004 [cd-ROM].

[24] Matgen P, Schumann G, Henry JB, Hoffmann L, Pfister L. Integration of SAR-derived inundation areas high precision topographic data, and a river flow model toward real-time flood management. Int J Appl Earth Obs Geoinf 2007;9(3):247–63.

[25] Matgen P, Schumann G, Pappenberger F, Pfister L. Sequential assimilation of remotely sensed water stages in flood inundation models. In: Symposium HS3007 at IUGG2007: Remote sensing for environmental monitoring and change detection, vol. 316. Perugia, Italy: IAHS; 2007.

[26] Norton JP. Roles for deterministic bounding in environmental modelling. Ecol Model 1996;86:157–61.

[27] Oberstadler R, Hoensch H, Huth D. Assessment of the mapping capabilities of ERS-1 SAR data for flood mapping: a case study in Germany. Hydrol Process 1997;10:1415–25.

[28] Oliver C, Quegan S. Understanding synthetic aperture radar images. Raleigh (NC): SciTech Publishing, Inc.; 2004.

[29] Pappenberger F, Beven K, Frodsham K, Romanowicz R, Matgen P. Grasping the unavoidable subjectivity in calibration of flood inundation models: a vulnerability weighted approach. J Hydrol 2007;333:275–87.

[30] Pappenberger F, Frodsham K, Beven K, Romanowicz R, Matgen P. Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations. Hydrol Earth Syst Sci 2007;11:739–52.

[31] Pappenberger F, Matgen P, Beven K, Henry JB, Pfister L, De Fraipont P. Influence of uncertain boundary conditions and model structure on flood inundation predictions. Adv Water Resour 2006;29:1430–49.

[32] Pietroniro A, Leconte R. A review of Canadian remote sensing applications in hydrology, 1995–1999. Hydrol Process 2000;14:1641–66.

[33] Pietroniro A, Leconte R. A review of Canadian remote sensing and hydrology, 1999–2003. Hydrol Process 2005;19:285–301.

[34] Rees W. Physical principles of remote sensing. Cambridge: Cambridge University Press; 2001.

[35] Refsgaard JC, Henriksen HJ. Modelling guidelines – terminology and guiding principles. Adv Water Resour 2004;27:71–82.

[36] Rotmans J, van Asselt MBA. Uncertainty management in integrated assessment modeling: towards a pluralistic approach. Environ Monitor Assess 2001;69:101–30.

[37] Saltelli A, Tarantola A, Campolongo F, Ratto M. Sensitivity analysis in practice – a guide to assessing scientific models. Chichester: John Wiley and Sons; 2004.

[38] Schmugge TJ, Kustas W, Ritchie JC, Jackson TJ, Rango A. Remote sensing in hydrology. Adv Water Resour 2002;25:1367–85.

[39] Schultz G, Engman E. Remote sensing in hydrology and water management. Berlin: Springer; 2000.

[40] Schumann G, Black A, Cutler M, Henry JB, Hoffmann L, Matgen P, et al. Hydraulic and event knowledge to reduce the positional uncertainty in SAR flood images for improved flood model calibration and development. In: Caetano M, Painho M, editors. Accuracy 2006: 7th international symposium on spatial accuracy assessment in natural resources and environmental sciences, IGP. Lisbon, Portugal: IGP; 2006.

[41] Schumann G, Matgen P, Cutler MEJ, Black A, Hoffmann L, Pfister L. Comparison of remotely sensed water stages from LiDAR, topographic contours and SRTM. J Photogram Remote Sens 2008;63:283–96.

[42] Schumann G, Matgen P, Pappenberger F, Black A, Cutler M, Hoffmann L. Evaluating uncertain flood inundation predictions with uncertain remotely sensed water stages. Int J River Basin Manage; in press.

[43] Schumann G, Matgen P, Pappenberger F, Hostache R, Pfister L. Deriving distributed roughness values from satellite radar data for flood inundation modelling. J Hydrol 2007;344:96–111.

[44] Schumann G, Matgen P, Pappenberger F, Hostache R, Puech C, Hoffmann L, et al. High-resolution 3D flood information from radar for effective flood hazard management. IEEE Trans Geosci Remote Sens 2007;45:1715–25.

[45] Seibert J. On the need for benchmarks in hydrological modelling. Hydrol Process 2001;15:1063–4.

[46] Smith LC. Satellite remote sensing of river inundation area, stage, and discharge: a review. Hydrol Process 1997;11:1427–39.

[47] Stewart JB, Engman ET, Feddes RA, Kerr YH. Scaling up in hydrology using remote sensing: summary of a workshop. Int J Remote Sens 1998;19:181–94.

[48] Wang Y, Liao M, Sun G, Gong J. Analysis of the water volume, length, total area and inundated area of the three gorges reservoir, China using the SRTM DEM data. Int J Remote Sens 2005;26:4001–12.

[49] Woodcock CE. Uncertainty in remote sensing. In: Foody G, Atkinson P, editors. Uncertainty in remote sensing and GIS. Chichester: John Wiley and Sons; 2002. p. 19–24.