Uncertainty Analysis in Flood Inundation Mapping

Younghun Jung[1] and Venkatesh Merwade[2]

Abstract

The accuracy of flood inundation maps is determined by the uncertainty propagated from all variables involved in the overall process including input data, model parameters and modeling approaches. This study investigates the uncertainty arising from key variables (discharge, topography, and Manning's n) among model variables in the East Fork White River near Seymour, Indiana. Methodology of this study involves the first order approximation (FOA) method to estimate the propagated uncertainty rates and the generalized likelihood uncertainty estimation (GLUE) to quantify the uncertainty bounds. Uncertainty bounds in the GLUE procedure are evaluated by selecting a likelihood function, which is a statistic (F-statistic) based on the area of observed and simulated flood inundation map. The results from GLUE show that the uncertainty propagated from multiple variables produce an uncertainty bound of about 15% in the inundation area compared to observed inundation.

Introduction and Objectives

Quantifying the role of uncertainty is critical for the improvement of flood prediction capabilities. Uncertainty in flood inundation mapping arises from input data as well as modeling approaches including hydraulic modeling, hydrologic modeling, and terrain analysis. Although the variables contributing to uncertainties in flood inundation mapping are well documented by several studies (Romanowicz and Beven 1998; Pappenberger et al. 2005, 2006; Merwade et al., 2008), it is impossible to completely remove these uncertainties due to constraints imposed by time, cost, technology, and knowledge. Similarly, although the uncertain variables in flood inundation mapping are known, not all of them contribute equally to the final uncertainty in the flood inundation map for a given circumstance. Therefore, deciding the priority among the elements that cause uncertainty is the first step, and reducing the sources of uncertainty for the prioritized variables is the second step in reducing the overall uncertainty in flood inundation mapping.

---

[1] Graduate Student, School of Civil Engineering, Purdue University, West Lafayette, IN 47907
[2] Assistant Professor, School of Civil Engineering, Purdue University, West Lafayette, IN 47907

The objectives of this study are to: (i) estimate the propagated uncertainty rates of key variables in flood inundation mapping by using the first order approximation (FOA) method; and (ii) quantify the uncertainty bounds arising from multiple variables using the generalized likelihood uncertainty estimate (GLUE). Monte Carlo (MC) simulations using HEC-RAS and triangle based interpolation are performed to investigate the uncertainty arising from discharge, topography, and Manning's n in East Fork White River near Seymour, Indiana as a study site.

Study Area and Data

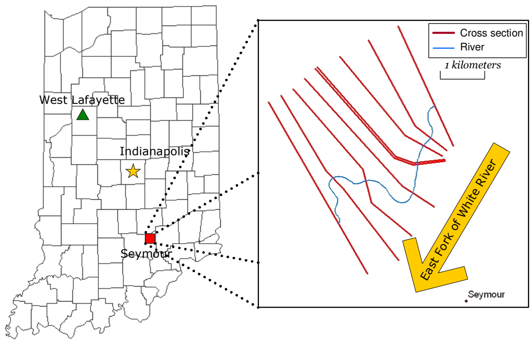


Figure 1. Study Area

This study is performed along a 5 km reach (Seymour reach) of the East Fork of the White River near Seymour Indiana. The East Fork of the White River begins in Columbus, Indiana, and joins the West Fork of the White River before draining into the Wabash River. The region around the selected Seymour reach was affected by the July 2008 flood event. The Seymour reach is characterized by a relatively wide floodplain with U shaped cross-sections. The topography data for extracting cross-sections and flood inundation mapping is obtained from the digital elevation model (DEM) from the 2005 IndianaMap Color Orthophotography Project by Indiana University. A total of nine cross-sections are extracted from the 1.5m horizontal resolution DEM. The average width of the Seymour reach cross-sections is 3.9km with an average spacing of 700m. The flow data used for hydraulic modeling of the Seymour reach include the observed discharge of 2729.7m^3/s with a reach boundary condition of downstream normal depth. The land use for the Seymour reach main channel ranges from a Manning's n value of 0.04 to 0.05. In the floodplains, the Manning's n value ranges from a value of 0.04 to 0.12.

Methods

First order approximation (FOA) method

First-order approximation (FOA) is a relatively simple technique for estimating the amount of uncertainty, or scatter, of prediction by a deterministic model transferring from multiple variables in a functional relationship. The moment analysis of a function associated with independent random variables is the basis of FOA. The FOA approach to hydrologic problems was suggested by Benjamin and Cornell (1970), and the method has been applied to flood risk analysis (Johnson and Rinaldi 1998; Liu et al. 2001). The uncertainty (σ_y) of model output (Y) is computed by knowing the uncertainty (σ_x) of independent variables (X) and the associated propagated uncertainty rate (dy/dx) as given in Eq. 1.

$$\therefore \sigma_y^2 = \left(\left. \frac{dy}{dx} \right|_{x=\bar{x}} \right)^2 \sigma_x^2 \tag{1}$$

Generalized Likelihood Uncertainty Estimation (GLUE)

The GLUE method involves forward MC simulations using different parameter values sampled from a feasible range. The objective of the GLUE method is to identify a set of 'behavioral' or acceptable models within the possible model/parameter combination (Beven and Binley, 1992). Outputs from all the simulations that are created by using the feasible parameter sets are weighted by a likelihood measure, which is a function that describes how well the simulated model matches the observed data. Generally, likelihood measures based on Bayes equation (Eq. 2) can be estimated by several likelihood functions, such as inverse of sum of squared error, inverse of sum of absolute error, and Nash-Sutcliffe efficiency.

$$P(Z < z) = \sum \left(L[M(\Theta, I)] \mid Z < z \right) \tag{2}$$

where, P is posterior likelihood values and Z is the value of z simulated by model. $L[M(\Theta, I)]$ is likelihood measure by model prediction, M, for given parameter, Θ, and set of input data, I. Thus, a higher likelihood measure indicates better fit between the model output and the observed data, and vice versa. A cutoff threshold for likelihood measure then classifies the simulated outputs as behavioral (acceptable) or non-behavioral. The likelihood measures of the behavioral models are then rescaled to obtain the cumulative density function (CDF) of the output prediction. The median of the rescaled CDF is generally taken as the deterministic model prediction (Blasone et al. 2008), and the uncertainty bound corresponding to this prediction is quantified by the 90% confidence interval selected at 5% and 95% confidence levels.

Methodology

The methodology involves: (i) creating probability distribution for each variable (discharge, Manning's n and topography); (ii) running Monte Carlo simulations using the HEC-RAS hydraulic model; and (iii) uncertainty analysis using GLUE and FOA. A brief description about each step is provided below.

Probability distributions for discharge, Manning's n and topography

A uniform distribution is assumed for Manning's n, and the values for Manning's n are assigned based on four types of land use including cultivated land, tree, urban area and water. The range to define the uniform distribution for Manning's n for each land use type is extracted from Chow (1959). For discharge data at the Seymour reach, a stage-discharge rating equation based on historic peak flows is developed through regression. By assuming a t-distribution for the stag-discharge raging curve, discharge values within the 95% confidence bounds of the observed flow of the 2008 flood event (2729.7 m^3/s) in the regression equation are used to define the range of flow rate values. The DEM used in this study has a vertical accuracy of ± 69 cm, and therefore a uniform distribution is assumed for topography to generate random digital elevation models to extract cross-sections for HEC-RAS, and to map flood inundation. The range of values used for each random variable is presented in Table 1. In the case of Manning's n, a random number actually represents a percentage, and this percentage is applied to each Manning's n within a cross-section. For example, if a cross-section has three Manning's n of 0.03 (left bank), 0.02 (main channel), and 0.04 (right bank), a random number of -10 % will reduce these Manning's n to 0.027, 0.018 and 0.036 to represent a change of -10%.

Table 1. Random variables (RV) in Monte Carlo Simulations

Initial (variables)	Modeling Variables estimated by RV	Min	Max	Probability Type	No. of Chosen RV
N_i Manning's n	$N = N_i\,(1+RV)$	-0.375	0.375	Uniform	1
F_i Discharge	$F = RV$ [m^3/s]	2257	3301	T-distribution	1
E_i Topography	$E = E_i + RV$ [cm]	- 69	69	Uniform	1

Monte Carlo (MC) Simulations

After defining probability distribution for each uncertain variable (Manning's, discharge and topography), random values are picked from these distributions to run HEC-RAS in MC simulations. A total of 1000 HEC-RAS simulations are conducted for each individual variable, and 5000 HEC-RAS simulations are conducted by using a combination of all variables. All HEC-RAS simulations are conducted with steady state assumption.

Estimation of the propagated uncertainty rate using the FOA method

FOA method requires a mathematical equation that relates random variables with model output to define the rate of propagation uncertainty in Eq. 1. In this study, a regression equation is developed to define a mathematical relationship between each target random variable (Manning's n, discharge and topography) and flood inundation area. The propagated uncertainty rate is estimated through 1000 MC simulations for each variable. The propagated uncertainty rate is computed by taking the ratio of flood inundation area to the change rate (%) of each target random variable. This ratio defines the dy/dx term in Eq.1.

Quantification of uncertainty using GLUE

After MC simulations, all outputs are evaluated by a likelihood measure to reflect how well the simulated model compares with the observed or baseline output. The selection of a likelihood measure is a subjective process, and the uncertainty bound obtained using GLUE is affected by the choice of the likelihood measure. In this study, F-statistic (Eq. 3) that includes the spatial aspects of a flood inundation map is used to estimate the likelihood measure for the uncertainty bound.

$$\text{F statistic of i}^{\text{th}}\text{ iteration}, \quad F_i = \left(\frac{A_{op,i}}{A_o + A_{p,i} - A_{op,i}} \right) \times 100 \qquad (3)$$

where A_o indicates the observed inundation area, A_p refers to the predicted flood inundation area, and A_{op} represents the intersection of both observed and predicted inundation areas. Uncertainty bound using GLUE is estimated based on the output of MC simulations.

Table 2. MC simulation results for Seymour Reach (Area in Km^2)

	Combination	Manning's n	Topography	Discharge
Min	6.706	9.935	9.204	10.441
Max	11.085	10.814	10.957	10.675
Deviation	4.379	0.879	1.753	0.234

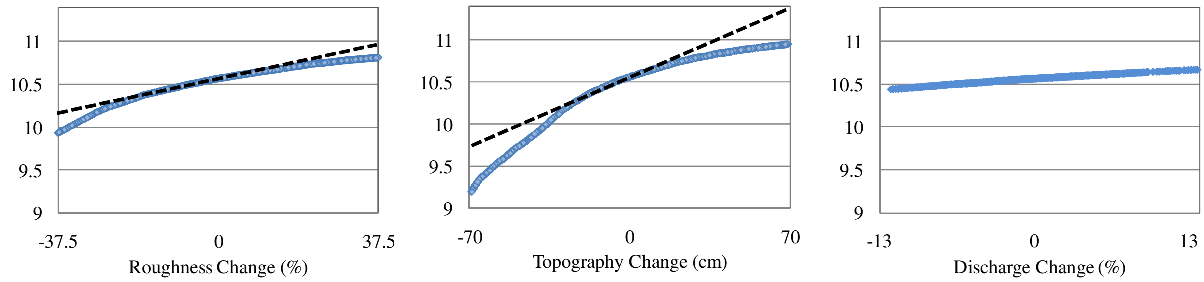


Figure 2. FOA method for Seymour Reach. X axis shows the change of variables and Y axis indicates inundation area (km2). Solid line shows the plotted inundation area and the dotted line is a linear line by FOA method.

Results

Results from MC simulations for each variable including Manning's n, topography and flow, and a combination of all variables are presented in Table 2. The simulated inundation area is in the range of 6.70 km^2 to 11.09 km^2 for the combined parameters, and is 9.20 km^2 to 10.96 km^2 for each variable. The results of the FOA analysis show how much uncertainty from each variable is transferred to the flood inundation area (Fig. 2). Quantitatively, a 1% change in uncertainty of Manning's n produces a corresponding change of 0.011 km^2 in the flood inundation area. A 1% change in discharge produces a corresponding change of 0.009 km^2 in the flood inundation area, and a 1cm change in topography produces a 0.012 km^2 change in the flood inundation area.

Results from GLUE analysis show the uncertainty bound in the flood inundation area from individual random variable as well as from the combination of all variables including Manning's n, discharge and topography (Table 3 and Fig. 3). The uncertainty bound for flood inundation area is in the range of 0.15 to 1.27 km^2 for individual variable, and is 1.61 km^2 for

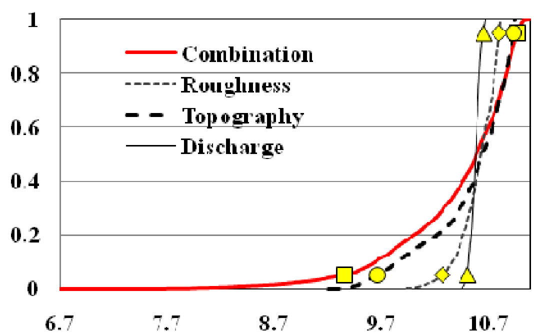
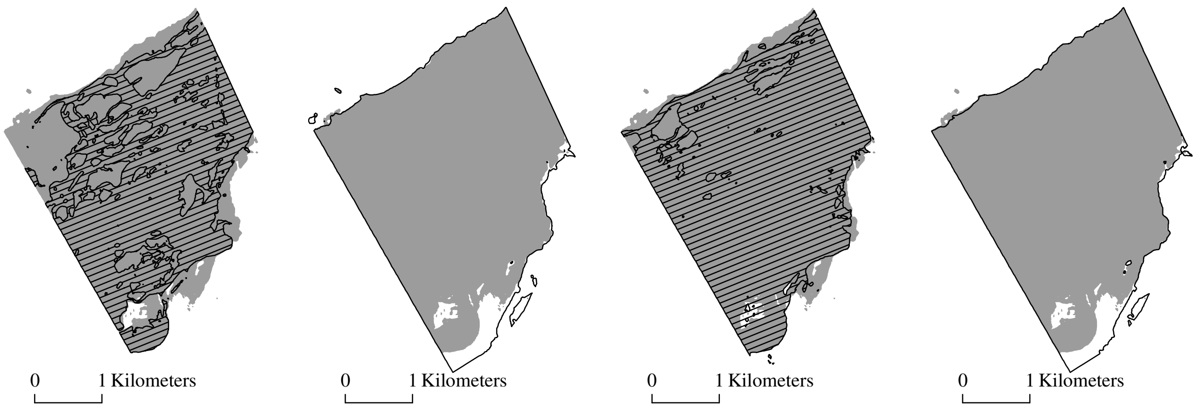


Figure 3. GLUE for Seymour Reach. X axis shows the inundation area and Y axis indicates CDF.

Table 3. Uncertainty Bounds from GLUE results

	Combination	Manning's n	Topography	Discharge
Lower 5%	9.356	10.273	9.662	10.501
Upper 95%	10.969	10.801	10.936	10.654
90% Bound	1.613	0.528	1.274	0.153



a) min (hatched) and obs (shaded)
b) max and obs (shaded)
c) 5% and obs (shaded)
d) 95 % and obs (shaded)

Figure 4. Flood inundation maps for Seymour Reach.

combined variables. Similar to MC simulations, combination of all variables produce the widest uncertainty bound (1.61 km^2) followed by topography, Manning's n and discharge. Considering the observed inundation area of 10.57 km^2 for the Seymour reach, the uncertainty bound for inundation area ranges from 1.4 % to 15.3 % of the base inundation area. Flood inundation maps for the Seymour reach are shown in Fig. 4.

Conclusions

The following conclusions are drawn from this study:

- This study presents an approach for quantifying the uncertainty and the propagation of uncertainty in flood inundation mapping using FOA and GLUE methods. FOA analysis using the 2008 flood data on the Seymour reach shows that the propagation of uncertainty is highest for topography followed by Manning's n and discharge.

- The GLUE analysis also showed that topography emerged as the top uncertain variable for the Seymour Reach. This finding can be attributed to the accuracy of topography data in flood inundation modeling and mapping. This conclusion is consistent with past studies that have found the accuracy of topography data to play a major role in flood inundation mapping.

- The uncertainty bound from each variable does not add up to produce the combined uncertainty bound, thus demonstrating the non-linear nature of uncertainty propagation in the overall flood inundation mapping process.

- The findings of this study are based on one single reach in Indiana. More studies using different topographic and flow conditions are needed to generalize the role of uncertainty and uncertainty propagation in flood inundation mapping.

Reference

Benjamin, J.R. and Cornell, C.A. (1970) Probability, Statistics and Decision Making for Civil Engineers. McGraw-Hill, New York, N.Y.

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

Blasone, R.-S., J. A. Vrugt, H. Madsen, D. Rosbjerg, B. A. Robinson, and G. A. Zyvoloski (2008) Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov chain Monte Carlo sampling, Adv. Water Resour., 31, 630–648.

Chow, V.T., (1959) Open-channel hydraulics, New York, McGraw-Hill

Johnson, P. A., and Rinaldi, M. (1998). Uncertainty in stream channel restoration. Uncertainty modeling and analysis in civil engineering, B. M. Ayyub, ed., CRC Press, Boca Raton, Fla., 425–437.

Liu, J., Tian, F., and Huang, Q. (2001) A risk analyzing method for reservoir flood control. *Hydrology*, 21(3): 1-3.

Merwade, V.M., Olivera, F., Arabi, M., and Edleman, S. (2008) Uncertainty in flood inundation mapping – current issues and future directions. ASCE Journal of Hydrologic Engineering, 13 (7), 608–620

Pappenberger, F., K. Beven, et al. (2005) Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations. Journal of Hydrology 302(1-4): 46-69

Pappenberger, F., P. Matgen, et al. (2006) Influence of uncertain boundary conditions and model structure on flood inundation predictions. Advances in Water Resources 29(10): 1430-1449

Romanowicz, R. and Beven, K.J. (1998) Dynamic real-time prediction of flood inundation probabilities. Hydrol. Sci. J. 43: 181–196