ASSESSMENT OF EPISTEMIC UNCERTAINTY IN FLOOD INUNDATION MODELING

by

Zhu Liu

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Lyles School of Civil Engineering

West Lafayette, Indiana

August 2018

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Venkatesh Merwade, Chair

Department of Civil Engineering

Dr. Dennis Lyn

Department of Civil Engineering

Dr. Antoine Aubeneau

Department of Civil Engineering

Dr. Bernard Engel

Department of Agricultural and Biological Engineering

Approved by:

Dr. Dulcy Abraham

Head of the Graduate Program

To my family

ACKNOWLEDGEMENTS

I would like to express my gratitude for my advisor, Dr. Venkatesh M. Merwade, for his immeasurable guidance and encouragement throughout my graduate studies. I would also like to extend sincere thanks to my PhD committee members, Dr. Dennis Lyn, Dr. Antoine Aubeneau and Dr. Bernard Engel, who have provided insights and helpful comments for improving the quality of my research. I especially wish to thank Lan Zhao from Rosen Center for Advanced Computing in Purdue University for her help with the High Performance Computing, and Dr. Jacques W. Delleur for the travel grant. I would like to appreciate Lyles family for the Lyles Teaching Assistant support. I also like to thank all my colleagues and friends in hydro group: Keighobad Jafarzadegan, Siddharth Saksena, Sayan Dey, Jessica Eisma, Zhenglyu Lyu, Ganesh Mallya, David Canon, Nicholas Olsen, Abhinav Gupta, Abhishek Abhishek, Brace Wang, Anzy Lee and former labmates: Dr. Adnan Rajib, Liuying Du, Nikhil Sangwan, Kyungmin Sung, Dr. Becca Essig, Dr. Meenu Ramadas and Dr. Jun Choi. I will never forget sharing my graduate life with all of you. I am truly grateful to my parents for their support during my PhD studies. To my wife Jina Yin, thanks so much for being with me and share plenty of happiness and struggle during this track. I would never complete my PhD without your accompany and encouragement.

TABLE OF CONTENTS

LIST OF TABLES viii
LIST OF FIGURES ix
ABSTRACT xii
CHAPTER 1. INTRODUCTION 1
1.1 Why epistemic uncertainty study is important for flood modeling? 1
1.2 Research gaps related to epistemic uncertainty in flood modeling? 3
1.3 Research Objectives 4
1.4 Organization of this Dissertation 5
CHAPTER 2. INVESTIGTING THE ROLE OF MODEL STRUCTURE AND SURFACE ROUGHNESS IN GENERATING FLOOD INUNDATION EXTENTS USING 1D AND 2D HYDRAULIC MODELS 6
2.1 Abstract 6
2.2 Introduction 6
2.3 Study areas and data 10
2.4 Model description 13
2.5 Methodology 17
2.5.1 Assigning distributed floodplain surface roughness based on land use map 18
2.5.2 Flood modeling and inundation mapping 20
2.5.3 Comparison of flood inundation maps 22
2.6 Results 23
2.6.1 Influence of model structure on simulated inundation extent 23
2.6.2 Influence of channel surface roughness on model performance 30
2.6.3 Effect of floodplain surface roughness on model performance 32
2.6.4 Comparison of applying uniform versus distributed floodplain surface roughness 37
2.7 Conclusions 39
CHAPTER 3. ACCOUNTING FOR MODEL STRUCTURE, PARAMETER AND INPUT FORCING UNCERTAINTY IN FLOOD INUNDATION MODELING USING BAYESIAN MODEL AVERAGING 42
3.1 Abstract 42

3.2 Introduction 43
3.3 Study area and data 46
3.4 Methodology 49
3.4.1 Hydraulic modeling 49
3.4.2 Bayesian Model Averaging method 50
3.4.3 Box-Cox transformation and Expectation Maximization (EM) method 52
3.4.4 Sources of uncertainty and LISFLOOD-FP configurations 53
3.4.5 Evaluation of BMA prediction in training and validation periods 56
3.5 Results and discussion 57
3.5.1 Convergence of Expectation Maximization iterations and final model weights 57
3.5.2 Validation of BMA prediction performance 62
3.5.3 Comparison of probabilistic flood maps based on BMA_G and ensemble mean 65
3.5.4 Evaluation of uncertainty associated with BMA deterministic prediction 68
3.6 Conclusion 70
CHAPTER 4. SEPARATION AND PRIORITIZATION OF UNCERTAINTY SOURCES IN A RASTER BASED FLOOD INUNDATION MODEL USING HIERARCHICAL BAYESIAN MODEL AVERAGING 74
4.1 Abstract 74
4.2 Introduction 75
4.3 Study area and Data 78
4.4 Methodology 80
4.4.1 Hydraulic modeling 80
4.4.2 Uncertainty sources and model implementations 80
4.4.3 Hierarchical Bayesian Model Averaging (HBMA) 82
4.4.3.1 Bayesian model averaging 82
4.4.3.2 Hierarchical Bayesian model averaging 84
4.4.3.3 Conditional model probability and posterior model probability 86
4.4.3.4 Prediction mean and variance of hierarch model 86
The impact of uncertainty sources on water stage prediction and prediction variance 88
4.5 Results and discussion 89
4.5.1 BMA trees on flood peak day 89

4.5.2 The relative importance of each uncertainty sources 95
4.5.3 Model mean prediction accuracy in the hierarchical structure 97
4.5.4 Uncertainty associated with mean prediction at each level 101
4.6 Conclusion 104
CHAPTER 5. SYNTHESIS 108
5.1 Impact of model structure uncertainty on flood predictions 108
5.2 Impact of model parameter and input uncertainty on flood predictions 109
5.3 Impact of individual and combined uncertainty sources on flood predictions 110
5.4 Limitations and future work 110
REFERENCES 112
APPENDIX A 125
VITA 134

LIST OF TABLES

Table 2.1 Details of river geometry and flood event for each reach 11
Table 2.2 Manning's n values used for NLCD map 19
Table 2.3 Summary of two simulation scenarios 22
Table 2.4 Model simulated average water surface elevation (m) for each reach 26
Table 2.5 Comparison of maximum F index by using uniform (U) and distributed (D) floodplain Manning's n when the channel Manning's n is 0.03 39
Table 2.6 Comparison of maximum F index by using uniform (U) and distributed (D) floodplain Manning's n when the channel Manning's n ranges from 0.01 to 0.05 39
Table 3.1 USGS gauge IDs for the streamflow input and validation stations 48
Table 3.2 Widths of 5% (low), 50%(average), and 95% (high) thresholds for each river 55
Table 3.3 Four sources of uncertainty and their varied types 55
Table 3.5 Maximum distance between input (upstream) and validation (downstream) stations along each river 69
Table 4.1 USGS gauge IDs for the streamflow input and validation stations 79
Table 4.2 Widths of 5% (low), 50% (average), and 95% (high) thresholds for each river 82
Table 4.3 Five sources of uncertainty and their varied types 82
Table 4.4 Root mean square error (RMSE) between predicted and observed water stage at all levels in the BMA tree for four validation stations 98

LIST OF FIGURES

Figure 1.1 Schematic description of proposed research for uncertainty analyses in flood inundation modeling 5
Figure 2.1 Layout map of study reaches and observed inundation for the selected event 12
Figure 2.2 Flow and gage height data for upstream USGS stations during historical flood events at (a) White River (USGS 07074850); (b) Black River (USGS 07072500); (c) East River (USGS 03365500) and (d) Flatrock River (USGS 03363900)..... 13
Figure 2.3 The schematic diagram of methodology 18
Figure 2.4 NLCD 2011 land use condition at (a) White River; (b) Black River; (c) East River and (d) Flatrock River. 20
Figure 2.5 Comparison map when utilizing unique channel Manning's n value (n_{ch}=0.01) and distributed floodplain Manning's n value at (a) White River; (b) Black River; (c) East River and (d) Flatrock River..... 27
Figure 2.6 Comparison map when utilizing unique channel Manning's n value (nch=0.03) and distributed floodplain Manning's n value at (a) White River; (b) Black River; (c) East River and (d) Flatrock River. 28
Figure 2.7 Comparison map when utilizing unique channel Manning's n value (nch=0.05) and distributed floodplain Manning's n value at (a) White River; (b) Black River; (c) East River and (d) Flatrock River..... 29
Figure 2.8 Effect of channel Manning's n values on hydraulic model outputs at (a) White River; (b) Black River; (c) East River and (d) Flatrock River. 31
Figure 2.9 Effect of channel Manning's n values on model simulated inundation area at (a) White River; (b) Black River; (c) East River and (d) Flatrock River. 32
Figure 2.10 Comparison map when utilizing unique channel Manning's n value (n_{ch}=0.03) and unique floodplain Manning's n value (n_{pl}=0.06) at (a) White River; (b) Black River; (c) East River and (d) Flatrock River. 34
Figure 2.11 Comparison map when utilizing unique channel Manning's n value (nch=0.03) and unique floodplain Manning's n value (npl=0.12) at (a) White River; (b) Black River; (c) East River and (d) Flatrock River. 35
Figure 2.12 Effect of floodplain Manning's n values on hydraulic model outputs at (a) White River; (b) Black River; (c) East River and (d) Flatrock River. 36

Figure 2.13 Effect of floodplain Manning's n values on model simulated inundation area at (a) White River; (b) Black River; (c) East River and (d) Flatrock River. 37
Figure 3.1 Black River watershed and layout of input and validation USGS gages 48
Figure 3.2 Convergence of four indicators after EM iterations at four USGS validation stations during training period 61
Figure 3.3 BMA_S and BMA_G Weights for each model configuration at convergence based on training data 61
Figure 3.4 Predicted water stages from model ensemble, BMA_S and BMA_G, ensemble mean as well as observed water stage at four USGS stations in the validation period 64
Figure 3.5 Rank of daily average RMSEs of water stage from model ensemble, BMA_S, BMA_G, ensemble mean as well as observed water stage at four USGS stations during the validation period 65
Figure 3.6 Probabilistic flood maps based on BMA_G and ensemble mean 67
Figure 3.7 Evaluation of the performance of two types of probabilistic flood maps 68
Figure 3.8 Uncertainty associated with BMA deterministic predictions of water stage under 95% confidence interval. A1, B1, C1 and D1 are station based prediction ranges while A2, B2, C2 and D2 are global based prediction ranges at USGS 07069500, USGS 07072000, USGS 07068000 and USGS 07064000 respectively 70
Figure 4.1 Layout map of Black River watershed and input/validation USGS stations..... 79
Figure 4.2 The schematic diagram of hierarchical structure (BMA tree)..... 88
Figure 4.3 BMA tree of model weights and conditional model weights on flood peak day (April 26, 2011) at validation station USGS 07069500 92
Figure 4.4 BMA tree of mean water stage predictions on peak day (April 26, 2011) at validation station USGS 07069500..... 93
Figure 4.5 BMA tree of prediction variances on peak day (April 26, 2011) at validation station USGS 07069500 94
Figure 4.6 Relative impact of uncertainty sources on predicted water stage using model perturbation approach at four validation USGS stations (S is channel cross-sectional shape, P is parameter, T is topography resolution, F is flow magnitude and W is channel width). 96

Figure 4.7 Relative impact of uncertainty sources on water stage prediction variance considering BMA weights at four validation USGS stations (S is channel cross-sectional shape, P is parameter, T is topography resolution, F is flow magnitude and W is channel width). 97
Figure 4.8 Model (weight>=0.0001) prediction ensemble, high weight model, low weight model as well as observed water stage at four validation USGS stations. 99
Figure 4.9 Simulated and observed water stages of high weight model branch at different levels of BMA tree for USGS 07069500 99
Figure 4.10 Simulated and observed water stages of low weight model branch at different levels of BMA tree for USGS 07069500 100
Figure 4.11 R^2 of predicted water stages on different levels of BMA tree for four validation USGS stations 101
Figure 4.12 Mean water stage predictions at different levels of BMA tree and their 95% confidence interval bounds for the high weight branch at USGS 07069500..... 103
Figure 4.13 Mean water stage predictions at different levels of BMA tree and their 95% confidence interval bounds for the low weight branch at USGS 07069500..... 104

ABSTRACT

Author: Liu, Zhu. PhD
Institution: Purdue University
Degree Received: August 2018
Title: Assessment of Epistemic Uncertainty in Flood Inundation Modeling
Committee Chair: Venkatesh M. Merwade.

Flooding is one of the most devastating natural disasters in the world and it exacerbated during the past decades. In order to reduce the loss of lives and properties from repeating flooding events, reliable flood predictions are required. Currently, there exist a series of hydrodynamic models that have different model structures which solve different forms of governing equations in one- (1D), two- (2D) or three- (3D) dimensions, thus providing various possible predictions for decision makers to choose from. Even for the same model, depending on how the model is implemented for a specific parameter set, input data and channel geometry representation, the prediction is different. Therefore, investigating the reducible uncertainty (epistemic uncertainty) in flood inundation modeling and finding a proper way to generate robust predictions are very crucial for future modelers. In this dissertation, epistemic uncertainty sources from model structure, model parameter and model input are investigated and evaluated by using stream reaches ranging from reach scale to watershed scale in different geographical settings. The three objectives of this dissertation are to: (1) evaluate the impact of hydrodynamic model structure uncertainty on predicted water stages and inundation extents under different geophysical settings, and explore the influence of channel and floodplain roughness on model performance respectively, (2) investigate and apply a multi-model combing approach, Bayesian model averaging (BMA), to produce reliable predictions by considering four uncertainty sources including channel width, channel cross-sectional shape, channel roughness and flow forcing and

(3) separate and prioritize different uncertainty sources, including DEM resolution, channel width, channel cross-sectional shape, channel roughness and flow forcing, based on their relative influences using hierarchical Bayesian model averaging (HBMA).

In the first objective, the performance of four hydraulic models including HEC-RAS 1D, HEC-RAS 2D, LISFLOOD-FP diffusive and LISFLOOD-FP subgrid are evaluated at four rivers that have different geophysical settings in the United States. The results show that HEC-RAS 2D does not perform well at low channel roughness condition. However, at high channel roughness condition, the performance of HEC-RAS 2D and HEC-RAS 1D are comparable. The performance of the subgrid version of LISFLOOD-FP (LS) is more stable under different channel roughness conditions, and in general it performs better than the diffusive version (LD) in simulating floodplain inundation. Moreover, applying distributed floodplain roughness does not necessarily improve model performances.

In the second objective, LISFLOOD-FP subgrid model is applied for a relatively large catchment-Black River watershed in Missouri and Arkansas considering four uncertainty sources using BMA approach. The results indicate that although BMA deterministic prediction may not always outperform all the model members in the ensemble, this approach is able to provide a relatively robust water stage prediction. Typically, BMA deterministic prediction behaves better than most of the member predictions in the ensemble and ensemble mean prediction. BMA has better performance than ensemble mean prediction for high-chance flood regions at Black River watershed. On the other hand, there is no significant difference between two types of probabilistic flood maps for low-chance flood regions.

In the third objective, LISFLOOD-FP subgrid model is also set up in the Black River watershed to find out the relative influence of five different uncertainty sources. The results demonstrate

that channel width and topographical data resolution have largest impact on the hydrodynamic model predictions. These two sources are followed by flow forcing, which has relatively greater influence than channel cross-sectional shape and model parameter. However, when model weights are taken into account, input (topography and input forcing) and model parameter (roughness) have larger impact on prediction variance than model structure (channel shape and width).

CHAPTER 1. INTRODUCTION

1.1 Why epistemic uncertainty study is important for flood modeling?

Flooding is a global phenomenon that causes casualties and property losses on every inhabited continent. It is probably the most devastating, widespread and frequent natural disaster for human societies (Bates et al., 1997; Cook and Merwade, 2009; Horritt et al., 2001; Jiang et al., 2009; Jonkman and Kelman, 2005; Leskens et al., 2014; Merz et al., 2007; Schumann et al., 2016; Tingsanchali, 2012; Zhang et al., 2002). There is a growing concern that anthropogenic actions and climate change have increased the risk of flooding in the world for the past decades (Freer et al., 2013; Hirabayashi et al., 2013). Throughout human history, there has been a constant endeavor to understand and predict flood events and their impact. One of the keys in preventing and reducing losses from flooding is to provide reliable information to the public about the risk associated with flooding (Merwade et al., 2008). Representation of natural processes through the use of a hydraulic model will be macroscopic in comparison to reality, and therefore, model outputs will always contain uncertainties associated with the approximation of the real system (Apel et al., 2004; Aronica et al., 1998; Beven and Hall, 2014). Despite the development of many sophisticated hydraulic models, our ability to make better predictions is hindered by our lack of understanding of the channel-floodplain water propagation processes, and the spatial and temporal variability of key model inputs and parameters such as topography and surface roughness. Thus, quantifying and reducing uncertainty in flood modeling is an important research topic for improved flood modeling and prediction (Teng et al., 2017a). Although there are many sources of uncertainty in flood inundation modeling, it is important to recognize two basic types that are fundamentally different from each other: natural uncertainty and epistemic uncertainty (Merz and Thieken, 2005; Merz and Thieken, 2009; Neuhold et al.,

2009; Ross et al., 2009; Tung and Mays, 1981). Natural uncertainty (or aleatory uncertainty) stems from variability of the underlying stochastic or random processes and refers to quantities that are inherently variable over time, space, or populations such as the amount of annual rainfall in consecutive years, the pore size distribution of the soil in a field, or the evapotranspiration rate of a tree in the forest (Cullen and Frey, 1999). Such phenomena need probabilistic models for their descriptions. On the other hand, epistemic uncertainty is related to our ability to understand, measure and describe the system under study. For instance, when a hydraulic model is used for flood mapping, epistemic uncertainty exists in the numerical equations and the way natural system is described or represented in the model. Parameter uncertainty arises from using a specific value or values for the modeling process because it is impossible to measure all necessary variables at all points in time and space. Noise in the measurements due to technological limitations adds to epistemic uncertainty as well. Cullen and Frey (1999) stated that natural uncertainty is a property of the system while the epistemic uncertainty is a property of the analyst. The different states of knowledge and data resources may result in different levels of epistemic uncertainty regarding predictions from different analysts. An obvious difference in natural and epistemic uncertainty is that epistemic uncertainty could be reduced whereas natural uncertainty is not reducible (Merz and Thieken, 2005; Sun et al., 2012). Therefore, many researchers recognized that these two types of uncertainty should be treated separately (Apel et al., 2008; Apel et al., 2004; de MOEL and Aerts, 2011; Hall and Solomatine, 2008; Tung and Mays, 1981). Although both natural and epistemic uncertainty are equally important in flood modeling, investigating epistemic uncertainty is more practical since it is reducible. As a result, the proposed research is designed to study epistemic uncertainty in the context of flood modeling

to help provide more accurate and reliable flood level or inundation extent predictions in the future.

1.2 Research gaps related to epistemic uncertainty in flood modeling?

Many past studies address natural and epistemic uncertainty for flood modeling. Aronica and Bates (2002) applied generalized likelihood uncertainty estimation (GLUE) to assess parameter uncertainty in LISFLOOD-FP model prediction using observed binary pattern information derived from satellite synthetic aperture radar images. Pappenberger et al. (2005a) estimated the uncertainty of effective roughness parameters, boundary conditions and model structures in HEC-RAS 1D model using inundation and downstream level observations. Jung and Merwade (2011) applied GLUE to quantify the uncertainty of input data and modeling approach in generating flood inundation maps. Tiwari (2010) used a wavelet-bootstrap-ANN surrogate modeling framework to investigate the numerical and sampling uncertainty to provide reliable hourly flood forecast. Pappenberger et al. (2013) applied expert preference technique to probabilistic flood forecast to study the hydro-meteorological uncertainty. Yu et al. (2015) developed a GLUE incorporating moving least squares with entropy method for stochastic sampling to perform the uncertainty analysis in flood inundation modeling Despite all these studies to assess uncertainty in flood modeling, many questions still exist related to epistemic uncertainty. For instance, a few studies raised the issue of model structure to guide data collection and model development for enhanced decision making and hydraulic risk mitigation. Additionally, many of these studies assessed uncertainty using data from only a single or two reaches using a single flood event (Aronica et al., 2002; Horritt and Bates, 2002a; Jung and Merwade, 2011). According to Merwade et al. (2008), individual uncertainty sources and how they affect the overall uncertainty in the final flood inundation map are not well understood. The

objective of this study is to address some of these gaps related to uncertainty in flood inundation modeling.

1.3 Research Objectives

In order to investigate some of the remaining issues related to the epidemic uncertainty associated with flood modeling, three objectives will be studied in the proposed research. The first objective is to evaluate the epistemic uncertainty by comparing the performance of four commonly used one- dimensional (1D) or two- dimensional (2D) hydraulic models (HEC-RAS 1D, HEC-RAS 2D, LISFLOOD-FP diffusive and LISFLOOD-FP subgrid) for simulating flood inundation extents at four reaches and explore the role of surface roughness (channel roughness and floodplain roughness) on simulation outcomes. This objective addresses the influence of model structure (governing equations and ways for numerical discretization) and parameterization on flood inundation modeling. The second objective considers a wider range of epistemic uncertainty sources by using the Bayesian model averaging (BMA) statistical approach to combine the ensemble flood predictions of different LISFLOOD-FP subgrid model configurations. These model configurations are created by incorporating four epistemic uncertainty sources, including channel width, channel cross section shape, channel roughness and input flow accuracy. BMA approach is used to provide robust and reliable flood prediction at watershed scale. Although BMA is a rigorous multi-model combining approach, this method does not provide the relative contribution or influence from individual epistemic uncertainty source to the overall uncertainty. To address this limitation, the third objective uses the Hierarchical Bayesian Model Averaging approach (HBMA) to segregate and prioritize sources of uncertainty in a hierarchical structure (BMA tree) and evaluate the relative importance of different epistemic uncertainty sources (Chitsazan and Tsai, 2014). Fig. 1.1 shows the flowchart

of how the three objectives are related within the context of uncertainty analyses for flood inundation modeling. Details of the multi-model comparison and averaging methods including Bayesian model averaging and hierarchical Bayesian model averaging approaches used for first, second and third objectives will be discussed in Chapter 2, Chapter 3 and Chapter 4, respectively.

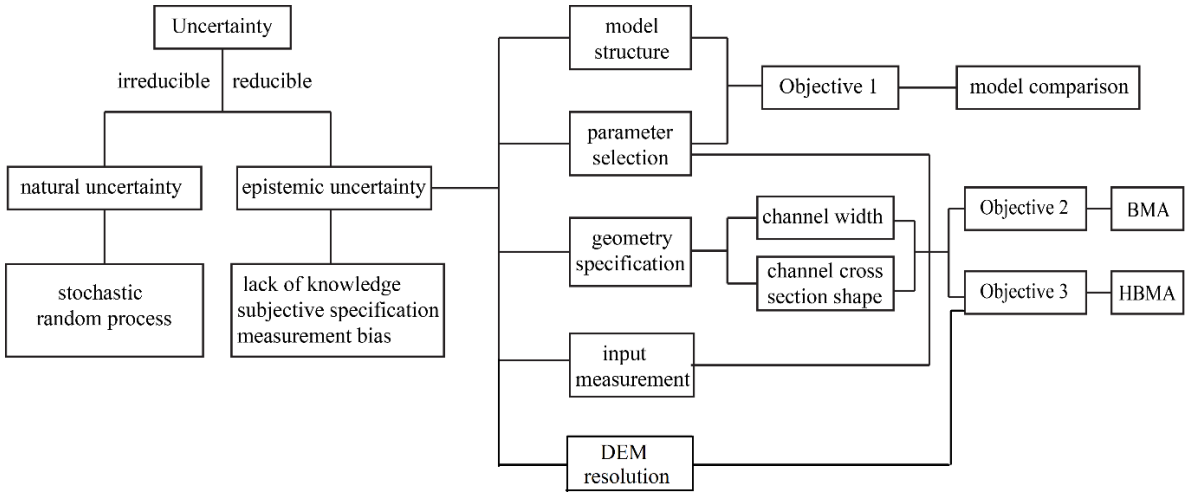

Figure 1.1 Schematic description of proposed research for uncertainty analyses in flood inundation modeling

1.4 Organization of this Dissertation

This dissertation consists of five chapters. Chapters 2-4 describe the three major topics of this dissertation. These chapters come from separate studies and they are presented in a self-contained manner, i.e., each chapter has an abstract, introduction, description of study area and data, methods, results, and summary sections. However, all of these chapters are linked under the topic of assessment of epistemic uncertainty in flood inundation modeling. The overall concluding remarks from this research are synthesized in chapter 5.

CHAPTER 2. INVESTIGTING THE ROLE OF MODEL STRUCTURE AND SURFACE ROUGHNESS IN GENERATING FLOOD INUNDATION EXTENTS USING 1D AND 2D HYDRAULIC MODELS

2.1 Abstract

Hydraulic models play an essential role in determining the flood inundation areas. Considering a wide array of one – (1D) and two-dimensional (2D) hydraulic models, selecting an appropriate model and its calibration are crucial in accurate prediction of flood inundation. This study compares the performance of four commonly used 1D and 2D hydraulic models including HEC-RAS 1D, HEC-RAS 2D, LISFLOOD-FP diffusive and LISFLOOD-FP subgrid with respect to their model structure and their sensitivity to surface roughness characterization. Application of these models to four study reaches with different river geometry and roughness characterization shows that for a given set of roughness condition, the geometry, including the sinuosity, reach length and floodplain width, did not affect the performance of a 1D or 2D model. Overall, the performance of 1D model was comparable to the 2D models used in the study, with the 2D models showing slightly better results. The performance of 2D models is affected by low channel roughness, and it improves with increasing channel roughness that enables more water enter into the floodplain. On the contrary, the performance of 1D model is affected positively with increasing floodplain roughness. When the models are evaluated for uniform versus distributed roughness characterization in the floodplain, the uniform surface characterization provided the best results compared to the distributed roughness characterization.

2.2 Introduction

Floods are expected to become more frequent and intense due to climate change, thus exacerbating their negative impacts (Arnell and Gosling, 2016; Das et al., 2013; Gobeyn et al.,

2017; Hirabayashi et al., 2013; Kay et al., 2009; Næss et al., 2005; Parinussa et al., 2016; Rosser et al., 2017; Spence et al., 2011; Wilby et al., 2008). Considering the increasing threat of frequently occurring high magnitude floods, there is a growing interest to create reach scale and regional scale flood inundation extents corresponding to near future storm forecasts (Bates, 2012; Bermúdez et al., 2017; Chang et al., 2010; Nguyen et al., 2015; Pappenberger et al., 2011). Flooding in natural channels is a three dimensional hydrodynamic process that is typically simulated by using models that are simplified for idealized environmental system based on certain assumptions (Marsooli et al., 2016; Zhang et al., 2016a). The structure of a hydraulic model can be described by its governing equations for river channel and floodplain as well as how these equations are solved in one, two or three dimensions. Accordingly, this structure dictates the physical description of the river geometry in the model. For example, a river is described by a set of discrete cross-sections in a 1D model; whereas it is described by a continuous mesh in a 2D model. The 2D mesh can be defined by either a raster grid, a mesh of equal size square cells, or a network of triangular or irregular cells. A 1D model assumes that the water moves only longitudinally in the direction of river, a 2D model assumes that water moves both longitudinally and laterally, and a 3D model assumes vertical movement in addition to the 2D flow. Some models couple 1D with 2D depending on how the water is flowing in the main channel and in the floodplains. Thus depending on the model structure and its parameterization, a flood simulation and the subsequent inundation can be very different based on the river morphologic characteristics and hydrodynamics (Beck, 2016; Blöschl et al., 2008; Cook and Merwade, 2009; Pappenberger et al., 2006b; Prestininzi et al., 2011). The result from a single hydraulic model provides just one possible realization of the flood extent without capturing the

uncertainty associated with the physical representation of the river in the model, its governing equation or parameters.

Selection of model structure plays a significant role in accurate simulation and prediction of floods. Although previous studies (Alho and Aaltonen, 2008; Horritt and Bates, 2001b; Leandro et al., 2009; Vojinovic and Tutulic, 2009) have compared the performance of 1D and 2D models, evaluation of recently released models such as HEC-RAS 2D and LISFLOOD-FP (subgrid) needs to be performed. Additionally, past studies focused on a single reach without considering the influence of variations in channel geometry, such as channel slope, cross section shape, and floodplain width, on model performance. Most hydrodynamic models, irrespective of their structure, are greatly influenced by topography and surface roughness characterization (Aronica et al., 1998; Elshorbagy et al., 2017; Fu et al., 2016; Horritt and Bates, 2001a; Pappenberger et al., 2005a; Pourali et al., 2016) because both topography and roughness coefficients affect flow area and velocity. The role of topography on flood modeling has been discussed in many past studies (Bates et al., 2003; Bren and Gibbs, 1986; Cook and Merwade, 2009; Saksena and Merwade, 2015; Zhou et al., 2004), but the role of surface roughness has received relatively less attention, especially within the context of model comparison. Specifically, the sensitivity of models to surface roughness characterization in the channel and floodplain needs to be investigated. For example, in a reach scale study many calibrated hydraulic models have a uniform value of surface roughness for the entire floodplain and another value for the entire channel (Horritt and Bates, 2001b; Horritt and Bates, 2002a; Tayefi et al., 2007). Therefore, the effect of applying uniform roughness for an entire floodplain compared to using distributed roughness based on landuse needs to be studied.

As noted above, most hydraulic modeling studies have been conducted on a single river segment or reach, but there is a growing interest now to create flood inundation maps for an entire stream network at regional to continental scales to facilitate flood preparedness and emergency response services. One of the models that has been used for large scale applications is LISFLOOD-FP, which has been applied for the European continent(Alfieri et al., 2014), lower Zambezi River (Schumann et al., 2013) and the Amazonian region (Wilson et al., 2007). AutoRAPID is another model that incorporates a continental scale river routing model known as Routing Application for Parallel Computation of Discharge (RAPID) with a regional scale flood delineation model called AutoRoute(Follum et al., 2017). It has been used in generating flood inundation extents for Midwestern United States (Tavakoly, 2015). When a hydraulic model is used over a larger region, its calibration and validation is limited to only a few reaches where data are available (Schumann et al., 2013; Wilson et al., 2007). As a result, a model that performs satisfactorily for one reach with one set of parameters may not perform equally on another reach, thus requiring a newer set of parameters. Under such circumstances, it becomes critical to know the limitations of a given model in simulating river hydraulics under different physical settings within a large basin. Considering the above discussion on the role of hydraulic model structure and parameterization in simulating floods under different settings, this study aims to compare the performances of multiple hydraulic models in generating flood inundation extents under different geophysical settings. Specifically HEC-RAS 1D, HEC-RAS 2D, LISFLOOD-FP diffusive and LISFLOOD-FP subgrid are used to answer the following questions: (1) What is the impact of models structure in simulating flood extents for identical input data and boundary conditions?(2) What is the influence of channel surface roughness and floodplain surface roughness in simulating flood

extents? and (3) What is the relative influence of distributed versus uniform floodplain surface roughness characterization on the performance of a hydraulic model?

2.3 Study areas and data

This study aims to explore the effect of model structure and roughness characterization on hydraulic modeling. Four study reaches are selected in this study, including the White River and Black River in Arkansas, East Fork White River (East River hereafter) and Flatrock River in Indiana (Fig. 2.1 and Table 2.1). These study areas provide distinct topographical, geomorphic and climate settings, varying reach lengths, sinuosity and floodplain width, thus offering good test beds for addressing the study objectives. With regards to model structure, the flow in a relatively straight channel with a valley shaped floodplain topography can be considered one-dimensional, and a 2D model may not provide significant improvement in flow hydrodynamics (Horritt and Bates, 2002a). The Flatrock reach, the shortest and relatively straight among the study reaches, fits this criteria. Assuming that 2D effects will dominate the flow hydrodynamics with increasing sinuosity of a stream reach (Finaud-Guyot et al., 2011), the study reaches provide a range of sinuosity to test the validity of this assumption. Similarly, 2D effects will dominate in a relatively flat and wider floodplain, such as that of the White River, compared to a narrow valley shaped floodplain (Mark et al., 2004). Finally, the distinct land use and floodplain topography of the study reaches will be useful in testing the effect of roughness characterization using 1D and 2D models. Besides meeting the criteria needed to accomplish the objectives of this study, the selected reaches have a history of flooding, along with the availability of inundation data for some of these events.

Hydraulic modeling of each study reach is conducted by using flow data of a historical flood event from an upstream United States Geological Survey (USGS) gauging station (Table 2.1).

The flow hydrographs for the study reaches are presented in Fig. 2.2, but to reduce the simulation time, the models are run for only days when the water stage is above the flood stage as described by the National Oceanic and Atmospheric Administration's Advanced Hydrologic Prediction Service (NOAA, 2015). Hydraulic models are validated using reference flood inundation maps that are derived using either the Landsat images (Dimitriadis et al., 2016; Mueller et al., 2016) or ground survey based high water marks. The Landsat images of flood inundation for White River and Black River can be found at USGS Earth Explorer website (USGS, 2015), which are classified into water and non-water areas with a supervised classification technique through ERDAS Imagine software. The ground survey based high water marks for June 7-9, 2008 floods at East River and Flatrock River are available from USGS (USGS, 2008). All these reference maps (Fig. 2.1) are treated as truth for this study to verify the model performance while the uncertainty associated with the data acquisition and analysis are neglected. The topography and land use dataset for all study areas are obtained in the form of National Elevation Dataset 30m DEMs (USGS, 1998) and NLCD 2011 land use (MRLC, 2011) respectively.

Table 2.1 Details of river geometry and flood event for each reach

Study Area	White River	Black River	East River	Flatrock River
Reach length (km)	56.7	25.2	11.2	3.9
Average channel width (m)	175	115	60	35
Channel slope (%)	0.0053	0.052	0.027	0.077
Floodplain width (km)	13	3.5	2.5	2
Bankfull depth (m)	4.5	2.5	2	1
Peak flow (cms)	1500	1600	2700	1700
USGS station ID	07074850	07072500	03365500	03363900
Flood stage (m)	7.9	4.3	3.7	3.4
Start of flood	22Dec2013	4Mar2015	5Jun2008	4Jun2008
End of flood	6Jan2014	27Apr2015	17Jun2008	9Jun2008
Image acquisition date	1Jan2014	25Mar2015	9Jun2008	9Jun2008

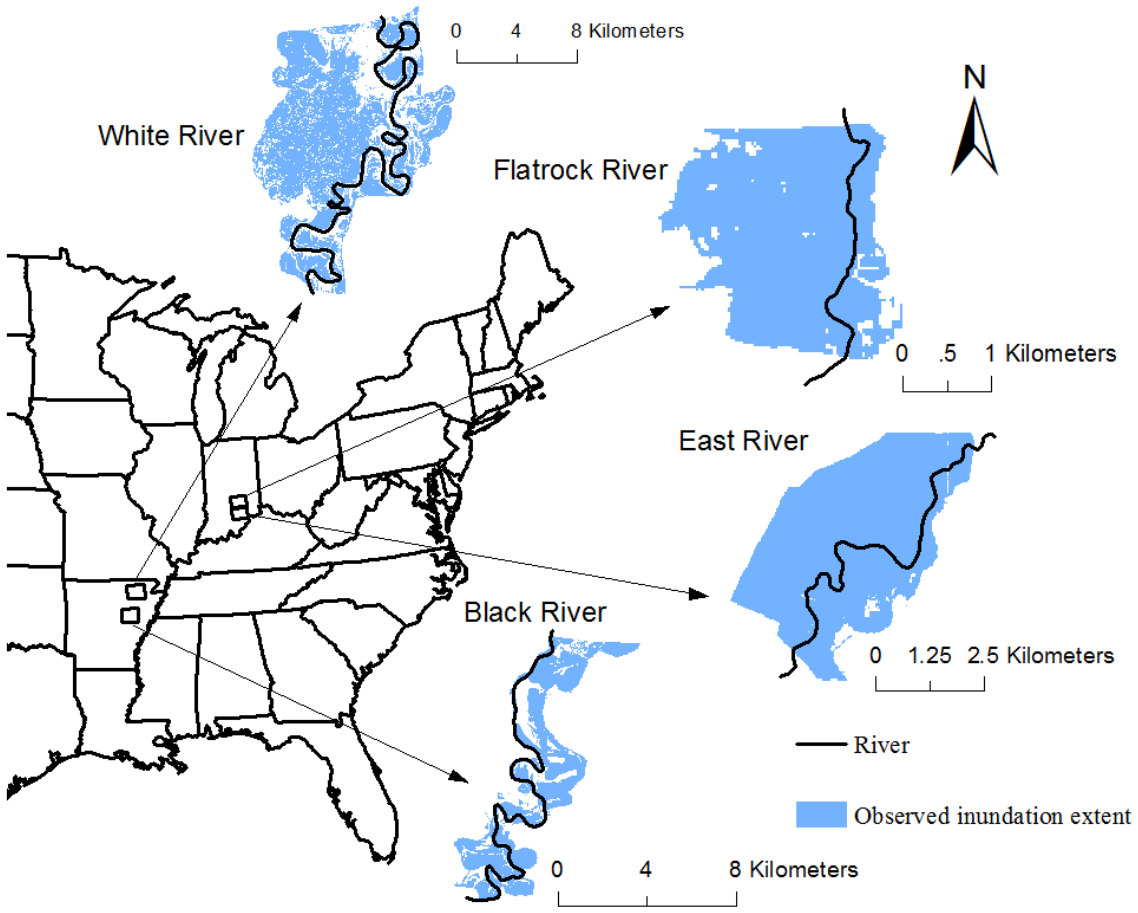

Figure 2.1 Layout map of study reaches and observed inundation for the selected event

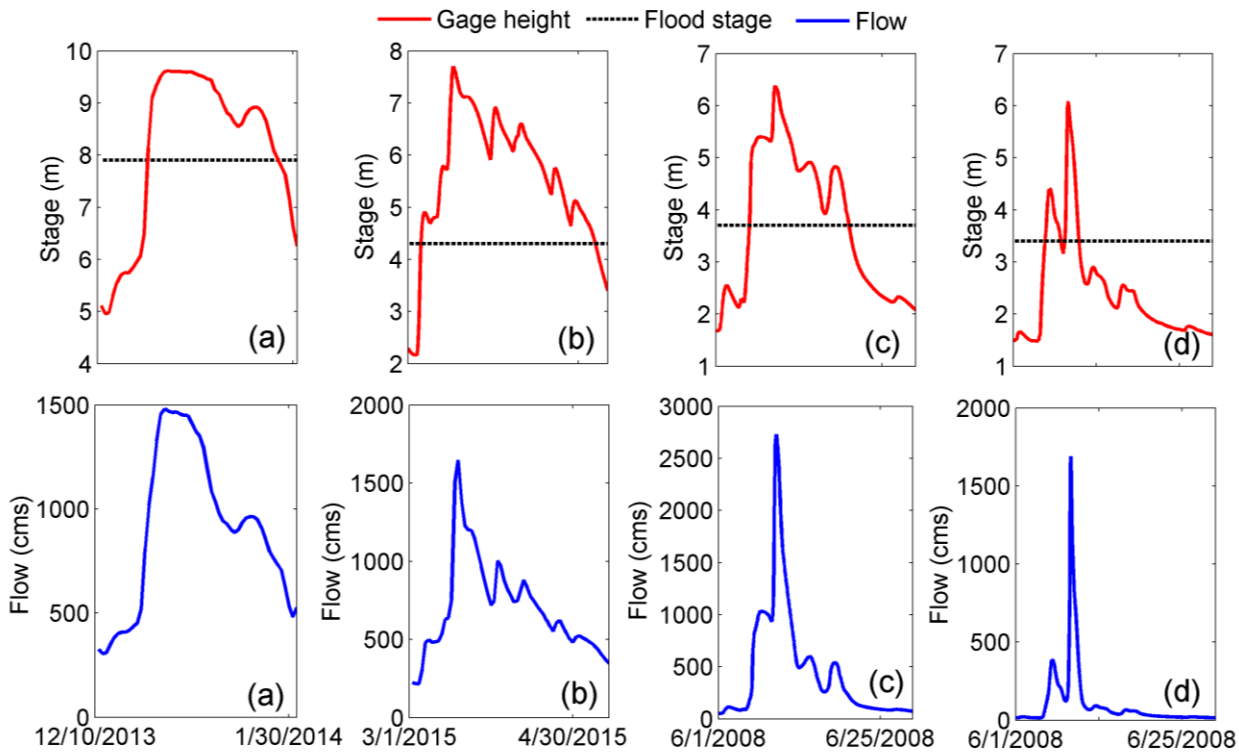

Figure 2.2 Flow and gage height data for upstream USGS stations during historical flood events at (a) White River (USGS 07074850); (b) Black River (USGS 07072500); (c) East River (USGS 03365500) and (d) Flatrock River (USGS 03363900).

2.4 Model description

Four hydraulic models are used in this study, including HEC-RAS 1D, HEC-RAS 2D, LISFLOOD-FP diffusive version and LISFLOOD-FP subgrid version. The first two models are developed by the United States Army Corps of Engineers Hydrologic Engineering Centre (Brunner, 2002; USACE, 2015) and the last two models are developed at Bristol University in the United Kingdom (Bates and De Roo, 2000; Neal et al., 2012a). HEC-RAS 1D (H1 hereafter) is a 1D hydraulic model that computes cross-sectional average water surface elevation (WSE) and velocity at discrete cross-sections by solving a full version of 1D Saint-Venant equations [2.1-2.4] (USACE, 2015)using implicit finite difference method. The calculated 1D WSE is then transferred to 2D inundation extent by: (i) linearly interpolating the WSE to create a water

surface grid; and (ii) subtracting the DEM from the water surface over the floodplain. Grid cells with positive results are designated as flooded.

$$\frac{\partial A}{\partial t} + \frac{\partial \phi Q}{\partial x_c} + \frac{\partial (1-\phi) Q}{\partial x_f} = 0 \qquad (2.1)$$

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x_c}\left(\frac{\partial \phi^2 Q^2}{A_c}\right) + \frac{\partial}{\partial x_f}\left(\frac{(1-\phi)^2 Q^2}{A_f}\right) + gA_c\left(\frac{\partial z}{\partial x_c} + S_c\right) + gA_f\left(\frac{\partial z}{\partial x_f} + S_f\right) = 0 \qquad (2.2)$$

$$\phi = \frac{K_c}{K_c + K_f}, \text{ where } K = \frac{A^{5/3}}{nP^{2/3}} \qquad (2.3)$$

$$S_c = \frac{\phi^2 Q^2 n_c^2}{R_c^{4/3} A_c^2}, S_f = \frac{(1-\phi)^2 Q^2 n_f^2}{R_f^{4/3} A_f^2} \qquad (2.4)$$

Where Q is the total flow down the reach, x_c and x_f are distances along the channel and floodplain, A (A_c, A_f) is the cross sectional area in channel and floodplain respectively, P is the wetted parameter, R_c and R_f are the hydraulic radius (A/P) for channel and floodplain, respectively, n is the Manning's roughness value and S is the friction slope. According to the conveyances K_c and K_f, ϕ determines how the flow is partitioned between the channel and floodplain. All HEC-RAS simulations performed in this study assume unsteady state flow conditions.

HEC-RAS 2D 5.0 version (H2 hereafter) is a recently developed hydrodynamic model to include purely two-dimensional flow routing within the unsteady flow analysis portion in the HEC-RAS model. The floodplain is represented in the model by one or several storage areas that are divided by mesh grids for calculation. The program solves either the full version of 2D Saint-Venant equations (Eq. 2.1-2.4) or the 2D diffusive wave equations by ignoring inertial terms as per users' input over the floodplain and 1D Saint-Venant equation inside the channel. The 1D channel and

2D floodplain needs to be linked using a lateral structure such as a levee. This study uses the 2D Saint-Venant equations that are solved with implicit finite volume algorithm for all floodplain simulations.

LISFLOOD-FP diffusive model (LD hereafter) is a raster based 1D and 2D coupled hydraulic model. Channel flow is simulated in 1D to capture the downstream propagation of a flood wave and the response of flow to free surface slope by using continuity and momentum equations [2.5-2.6](Bates et al., 2013).

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = q \tag{2.5}$$

$$S_0 - \frac{n^2 P^{4/3} Q^2}{A^{10/3}} - \frac{\partial h}{\partial x} = 0 \tag{2.6}$$

Where Q is the volumetric flow rate in the channel, q is the flow into the channel from tributary and floodplain, A is the cross sectional area of the flow, S_0 is the bed slope, n is Manning's coefficient of friction, P is wetted perimeter, and h is the flow depth.

Floodplain flow is described in terms of continuity and momentum equation and discretized as a grid of square cells to represent 2D flow on the floodplain. The 2D flow over the floodplain is actually quasi-2D in LISFLOOD-FP diffusive and subgrid solvers, but to keep the major distinction between 1D and 2D consistent, LIDFLOOD-FP is referred as 2D in this study. Flow between adjacent cells is calculated based on the free surface slope using Equations 2.7-2.8 (Bates et al., 2013).

$$\frac{dh^{i,j}}{dt} = \frac{Q_x^{i-1,j} - Q_x^{i,j} + Q_y^{i,j-1} - Q_y^{i,j}}{\Delta x \Delta y} \tag{2.7}$$

$$Q_x^{i,j} = \frac{h_{flow}^{5/3}}{n} \left(\frac{h^{i-1,j} - h^{i,j}}{\Delta x} \right)^{1/2} \Delta y \tag{2.8}$$

Where $h_{i,j}$ is free water surface height at the node (i,j), n is the Manning's friction coefficient for the floodplain, Δx and Δy are the cell dimensions, Q_x and Q_y represents the volumetric flow rate along x and y direction respectively, and h_{flow} describes the depth available for flow.

The numerical scheme of LISFLOOD-FP subgrid model (LS hereafter) is based on an extension of the LISFLOOD-FP base model (Bates et al., 2010) to incorporate a subgrid scale representation of channelized flows. The subgrid version enables a river channel with any width below that of the DEM grid resolution to be simulated in the model (Neal et al., 2012a; Schumann et al., 2013). In addition, channels are not necessarily assumed to be wide and rectangular in LS where hydraulic radius is assumed to be equal to water depth. Thus, LS can be used to simulate relatively small, narrow, deep channels. By considering the flow from channel and floodplain respectively, the momentum equation for channel, momentum equation for floodplain and continuity equations for LISFLOOD subgrid solver are described by Equations 2.9-2.11(Bates et al., 2013), respectively.

$$Q_c^{t+\Delta t} = \frac{Q_c^t - gA_{c,flow}^t \Delta t S_c^t}{\left\{1 + g\Delta t n_c^2 |Q_c^t| / \left[\left(R_{c,flow}^t\right)^{4/3/} A_{c,flow}^t\right]\right\}} \tag{2.9}$$

$$Q_{f,i+1/2}^{t+\Delta t} = \frac{q_{f,i+1/2}^t - gh_{f,flow}^t \Delta t S_{f,i+1/2}^t}{\left[1 + g\Delta t n_f^2 \left|q_f^t\right| / \left(h_{f,flow}^t\right)^{7/3}\right]} \left(\Delta x - w_{c,flow}\right) \tag{2.10}$$

$$Q_{i+1/2}^{t+\Delta t} = Q_{c,i+1/2}^{t+\Delta t} + Q_{f,i+1/2}^{t+\Delta t} \tag{2.11}$$

Where Q is the total volumetric flow rate, Q_c and Q_f are volumetric flow rate in channel and floodplain, respectively, A_c is the cross sectional area for the channel, S_c and S_f are the friction slope in the main channel and floodplain, respectively, n_c and n_f are the Manning's n friction coefficients for the channel and floodplain, respectively, i and j are cell spatial indices in x and y

directions, respectively, q_f is the unit width flow rate for floodplain, Δx is the cell width and w_c is the channel width.

Some studies (Bladé et al., 2012; Dimitriadis et al., 2016; Horritt and Bates, 2002a) have shown that H1 model is efficient and could reach similar levels of performance compared with 2D models, However, when the 1D model cross-sections spacing is greater than the element size of 2D models in the floodplain, H1 can be deficient in simulating complex floodplain processes (Bladé et al., 2012; Huang et al., 2007). The H2 model has shown to produce good results by solving the full Saint-Venant equation in 2D on the floodplain (Morales-Hernández et al., 2013; Quiroga et al., 2016; Şensoy et al., 2016), but it comes at a higher computational cost compared to H1. Even though the LD model neglects the channel-floodplain momentum transfer and the effects of advection and secondary circulation on mass transfer (Guidolin et al., 2016), its simple way of coupling 1D channel with 2D floodplain provides better representation of the real system compared to LS (Aronica et al., 2002; Horritt and Bates, 2002a). LS considers both mass and momentum transfer terms between channel and floodplain and it includes subgrid channels representations on the floodplain to increase the model simulation accuracy (Wood et al., 2016a), but a major shortcoming of LS is the use of global channel parameterization (Neal et al., 2012a).

2.5 Methodology

This study involves the following steps for each reach (Fig. 2.3): (i) setting up each model by assigning distributed Manning's n values in floodplain based on land use characterization; (ii) flood modeling and inundation mapping using each model with different combinations of channel and floodplain surface roughness; and (iii) evaluating each model's performance by comparing model simulated flood inundation extent with reference flood map using commonly used performance measures, including the fit (F) index, correctness (C) index, and inundation

area (Alfieri et al., 2014; Sangwan and Merwade, 2015). Because there is lack of observed flood map in time and it needs fair comparison among the four models, including their sensitivity to channel roughness description, all simulations are conducted without calibration.

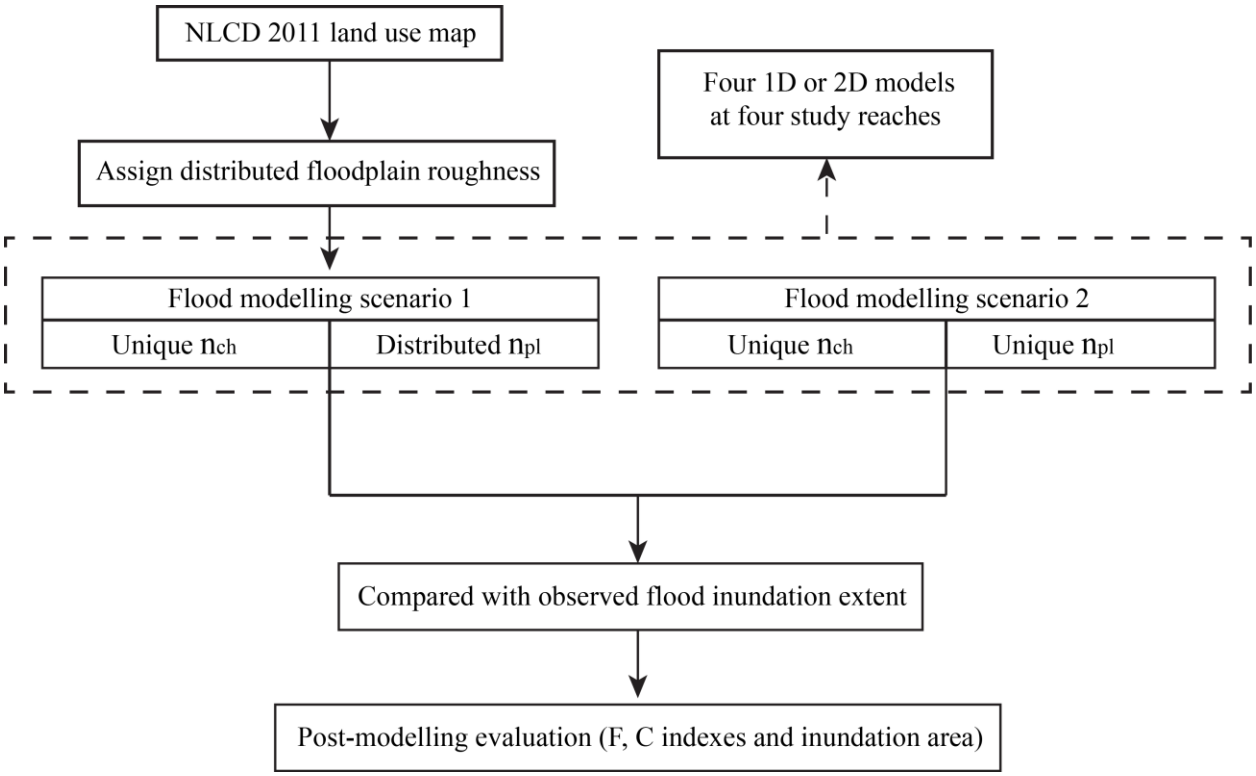

Figure 2.3 The schematic diagram of methodology

2.5.1 Assigning distributed floodplain surface roughness based on land use map

The flood events included in this study occurred between 2008 and 2015, so the NLCD 2011 land use data are used to assign Manning's n for all areas in the floodplain (Fig. 2.4). This roughness characterization is referred as "distributed roughness" in this study. The White River floodplain is characterized by the woody wetlands on the downstream side of the reach and cropland on the upstream. Black River has a narrow cropland floodplain; East River also has a cropland type floodplain but it is relatively wide. Flatrock River has cropland on one side and a combination of deciduous forest and cropland on the other side. Different land use categories

allow computation of flow resistance coefficients using measured or predicted floodplain and riparian vegetation properties (Shields Jr et al., 2017). The Manning's n value for each land use characterization (Table. 2.2) are therefore assigned based on the values from previous studies (Kalyanapu et al., 2010; McCuen, 1989).

Table 2.2 Manning's n values used for NLCD map

Land Cover	Description	Manning's n
11	Open water	0.001
21	Developed, open space	0.0404
22	Developed, low intensity	0.0678
23	Developed, medium intensity	0.0678
24	Developed, high intensity	0.0404
31	Barren land	0.0113
41	Deciduous forest	0.36
42	Evergreen forest	0.32
43	Mixed forest	0.40
52	Shrub/scrub	0.40
71	Grassland/herbaceous	0.368
81	Pasture/Hay	0.325
82	Cultivated crops	0.037
90	Woody wetlands	0.086
95	Emergent herbaceous wetlands	0.1825

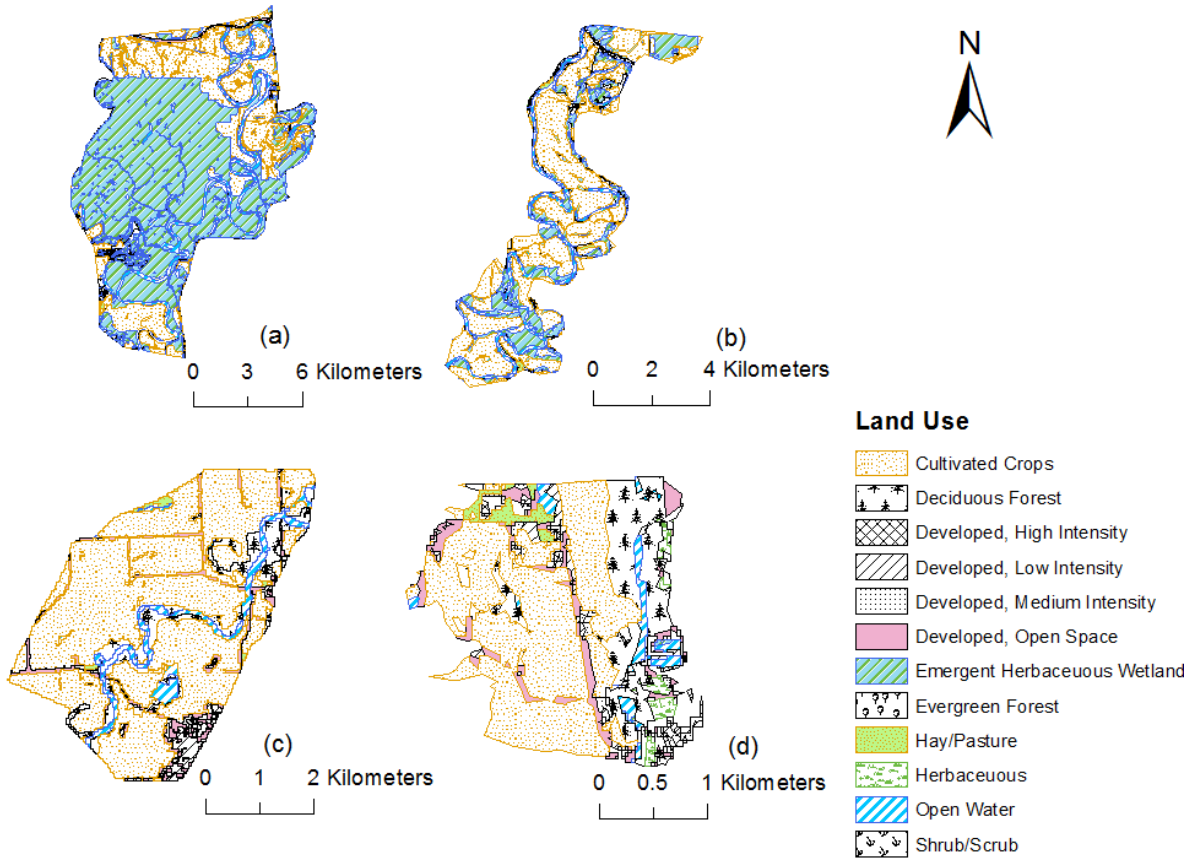

Figure 2.4 NLCD 2011 land use condition at (a) White River; (b) Black River; (c) East River and (d) Flatrock River.

2.5.2 Flood modeling and inundation mapping

Setting up H1 or H2 model using HEC-GeoRAS (Ackerman, 2005; Dysarz et al., 2014) requires the description of channel geometry, which is mainly provided in the form of cross-sections at multiple locations along the channel, bank lines to delineate the flow domains (channel and floodplain) and flow path lines to compute distances along the flow. HEC-GeoRAS is also used for post-processing the H1 model output and generate flood inundation extents. Besides the elevation raster, input data for LD and LS model include the description of stream centreline along with its ground profile and related attributes such as the average river width, slope, and surface roughness.

Hydraulic models are generally calibrated by changing the Manning's n for the channel and the floodplain when simulating flood inundation extents. To investigate the effect of Manning's n on 1D and 2D simulations, two modeling scenarios are created for H1, H2, LD, and LS models based on the roughness characterization in the channel and on the floodplain (Table 2.3). The first modeling scenario is designed to evaluate the effect of channel surface roughness on model performance by changing the channel Manning's n for each simulation, and keeping the distributed Manning's n for the floodplain unchanged. Nine simulations are performed with each model by changing the channel Manning's n from 0.01 to 0.05, with an increment of 0.005. This results in 36 simulations for each study site and a total of 144 simulations for this scenario. When conducting hydraulic simulations for all four models, the flow input and boundary conditions remain unchanged.

In the second scenario, the effect of floodplain surface roughness on model performance is investigated by keeping the channel Manning's n value at 0.03 and using different uniformly assigned floodplain Manning's n for each simulation. While the distributed floodplain roughness is available for all the study reaches from the landuse, the objective of creating this scenario is to test whether and to what extent a single roughness value for the entire floodplain affects the hydrodynamic simulation compared to distributed roughness characterization. Accordingly, five floodplain Manning's n values ranging from 0.03 to 0.15 with an increment of 0.03 are adopted based on its common range as found in previous studies (Horritt and Bates, 2002a). A Manning's n of 0.03 is representative of a channel that is clean, straight, full stage, no rifts or deep pools (Te Chow, 1959). In addition, 0.03 is also the mean of the channel surface roughness values, ranging from 0.01 to 0.05, used in this study, and also the range suggested for LISFLOOD-FP model user

manual (Bates et al., 2013). In this scenario, five simulations are created for each model at one study reach, thus generating a total of 80 simulations.

Table 2.3 Summary of two simulation scenarios

Modeling Scenarios	Scenario 1	Scenario 2
Channel Manning's n	0.01-0.05	0.03
Floodplain Manning's n	Distributed value	0.03-0.15
Total Simulations	144	80

2.5.3 Comparison of flood inundation maps

Results from all simulations are compared with inundation extents derived for each site from Landsat images on a single day during the flooding. To enable the comparison, inundation results from each model is extracted for the same day when the images were collected, and are examined using comparison map and three performance measures: F index (equation 2.12), C index (equation 2.13) and flood inundation area. The comparison map divides the simulated inundation extents into three categories as fit, overestimation and underestimation.

$$F = 100 * \left(\frac{A_{om}}{A_o + A_m - A_{om}}\right) \qquad (2.12)$$

$$C = 100 * \frac{Aom}{Ao} \qquad (2.13)$$

Where A_o refers the observed area of inundation, A_m is the model simulated flood inundation area, A_{om} refers to area that is both observed and simulated as inundated.

Both F and C indexes range from 0 to 100. A value of F = 100 means a perfect match between observed and predicted areas of inundation, and a lower F indicates discrepancy between the two.

The C index describes the percentage of the reference flood extent that is correctly predicted by the model output. Similarly, a value of C = 100 means all of the model extent falls within the reference extent, and a value of zero means the model prediction misses all of the reference flood map extent.

2.6 Results

2.6.1 Influence of model structure on simulated inundation extent

Three sets of comparison maps for Scenario 1 are presented in Figs. 2.5-2.7 to show the performance of each model for the study reaches by using three distinct uniform Manning's n values for the channel (n_{ch}=0.01, n_{ch}=0.03 and n_{ch}=0.05) and same distributed floodplain roughness values. Figs. 2.5-2.7 indicate that H1 model performs better for White River, Black River and Flatrock River than H2 model when the channel Manning's n is low, but as the channel Manning's n increases, H2 performs better than H1. When the channel surface roughness is really small, the flow depth is low to completely inundate the floodplain. When the channel surface roughness increases, flow velocity decreases and the flow depth increases to enter into the floodplain. Once water propagates into the floodplain, H2 model performs better than H1 model by capturing the floodplain dynamics in 2D (Dimitriadis et al., 2016; Manfreda et al., 2015). Moreover, when channel roughness (n_{ch}) is small, the partition variable ϕ becomes larger and more water is routed inside the channel (Equation 2.3).

Figure 2.5c shows that East River simulations are not affected by low channel Manning's n. The flow depth in an open channel is directly related to flow volume and inversely related to cross-sectional area and slope. A high flow discharge of 2700 m^3/s passing through a relatively small cross-sectional area of 120 m^2 at East River creates a large flow depth to enter into the floodplain even when channel surface roughness is small. The peak flow at both White River and Black

River are relatively small (1500 and 1600 m^3/s) and the cross sectional areas are large (788 m^2 and 288 m^2 respectively). Although the cross sectional area of Flatrock River is small (35 m^2), the low flow peak (1700 m^3/s) and relatively high slope (0.077%) at the Flatrock River cause the flow depth to be lower than the banks at low channel roughness.

Both LD and LS model do not use the full version of Saint-Venant equation in their model structure. The comparison maps (Figs. 2.5-2.7) show that in general the LD model overestimates the flood inundation extent for the study reaches. Comparatively, the LS model produces better results for all study reaches and channel Manning's n values. The only exception is the northwest corner of the Flatrock River floodplain, which remains dry for all Manning's n values. This portion of Flatrock River floodplain has a steep natural rise in topography, and it is not simulated well by any of the models in the study. The elevation data used in this study corresponds to the latest DEM from the USGS which was not available in 2008. It is possible that the difference in the observed and simulated inundation in this area is related to the discrepancy in the topography data.

The difference in the behaviour of LD and LS models can be related to several factors, including how the channel is represented in the model, the topography of the study reaches, and model setup resolution. The LD model assumes the channel cross section to be rectangular, which is not true for any of the study reach, but LS is able to better represent relatively complex cross section shape based on hydraulic geometry theory. The LD model only considers mass transfer between channel and floodplain, but neglects channel-floodplain momentum transfer, and the effects of advection and secondary circulation on mass transfer. Equations 2.6-2.8 indicate that local acceleration and convective acceleration terms are neglected from 1D Saint Venant momentum equation (equation 2.2) for LD model, but the LS model considers both mass and momentum

transfer (local acceleration term in momentum equation 2.9-2.10) between river channel and floodplain. As a result, LS model provides more detailed information about whether a specific cell is flooded or not around the channel while LD model gets almost all the cells near the channel flooded especially when surface roughness is high.

The LS model does not behave well at any channel roughness value for northwest corner of the Flatrock River due to a rise of 1m in topography. Performance of H2 model for Flatrock River improves as the channel roughness increases beyond 0.03 when the flow enters the floodplain. One possible reason is that the full Saint-Venant equations (2.1-2.4) which include the advection term produces higher WSE for H2 model (about 0.8 m higher than the 1m rise in topography) to inundate the northwest corner. The LS model does not include advection term in its governing equation and the simulated WSE is 1.0 m less than H2 in the northwest corner of Flatrock River. Additionally, the ways these two models conduct 2D floodplain calculation based either on mesh grids using finite volume method (H2 model) or raster cells using finite difference method (LS model) may also contribute to their difference.

Interestingly, only 1D model (H1) generates acceptable results even when the channel roughness is small. Table 2.4 indicates H1 model produces higher average water surface elevation at small channel roughness to inundate the northwest corner of Flatrock River floodplain. For similar reason, H1 model performs well at low channel roughness for other study reaches as well. According to Chow (1959), the roughness coefficient of main channel in the natural stream that is winding and has some weeds, stones, pools and ineffective slopes and sections could reach up to 0.055. It is possible that the actual Manning's n for the study channels may be higher than 0.03 at which the 2D models with better physics provide better results than H1. On the other hand, H1 which may treat Manning's n as a purely calibration parameter may give better results

when the high WSE inundates the entire floodplain at low n_{ch} values, thus providing better match with the reference flood map.

Table 2.4 Model simulated average water surface elevation (m) for each reach

Model	n_{ch}	White River	Black River	East River	Flatrock River
H1	0.01	59.6	74.6	173.1	189.0
	0.03	59.9	74.9	173.6	189.2
	0.05	60.1	75.1	173.8	189.6
LD	0.01	58.3	75.2	174.1	188.7
	0.03	60.3	75.4	174.9	189.9
	0.05	60.5	75.5	175.1	190.3
LS	0.01	59.8	74.3	173.1	189.3
	0.03	59.9	74.5	173.3	189.4
	0.05	60.2	74.6	173.4	189.5
H2	0.01	58.7	74.2	174.8	188.9
	0.03	59.5	74.3	175.0	189.8
	0.05	59.8	75.0	175.2	190.2

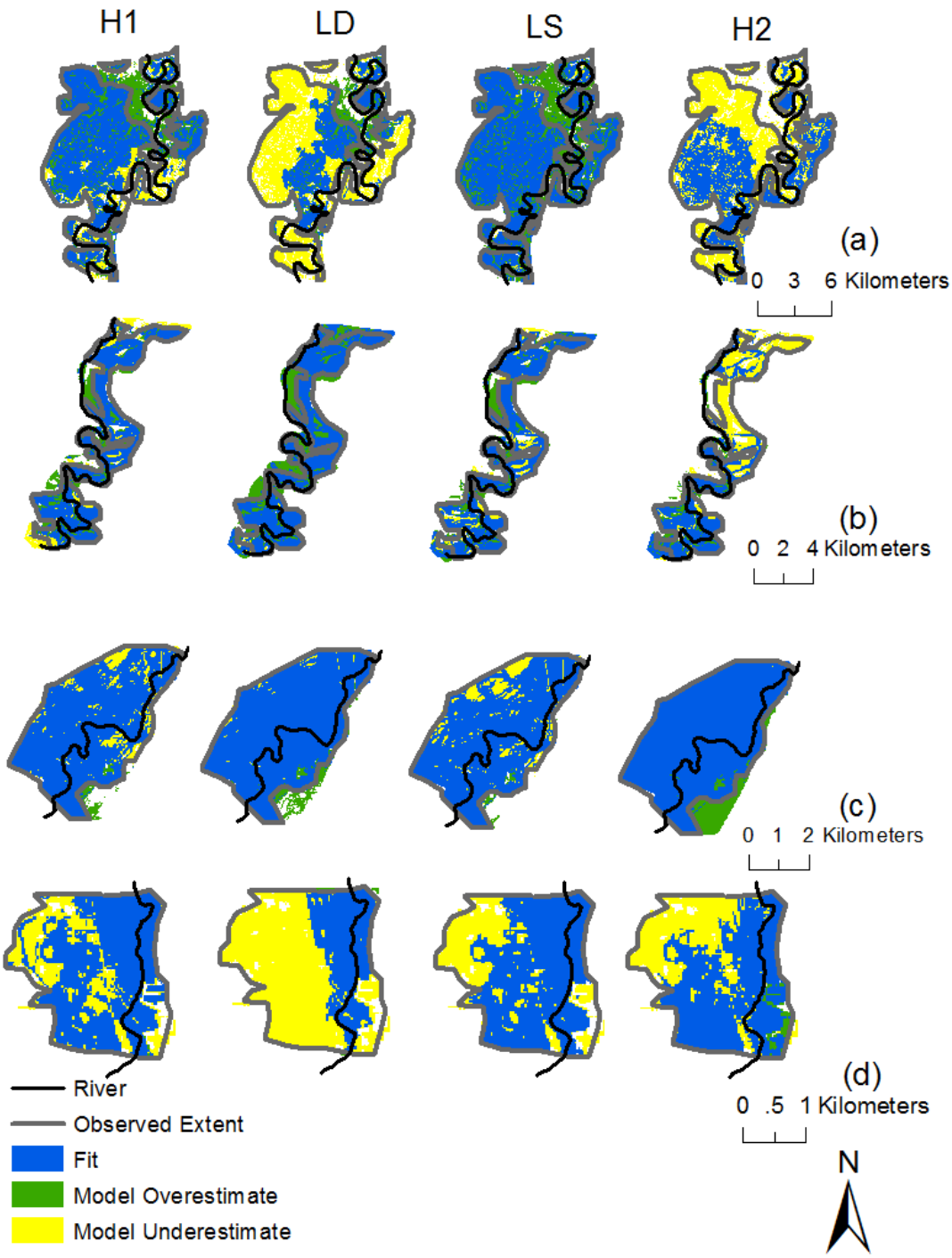

Figure 2.5 Comparison map when utilizing unique channel Manning's n value (n_{ch}=0.01) and distributed floodplain Manning's n value at (a) White River; (b) Black River; (c) East River and (d) Flatrock River.

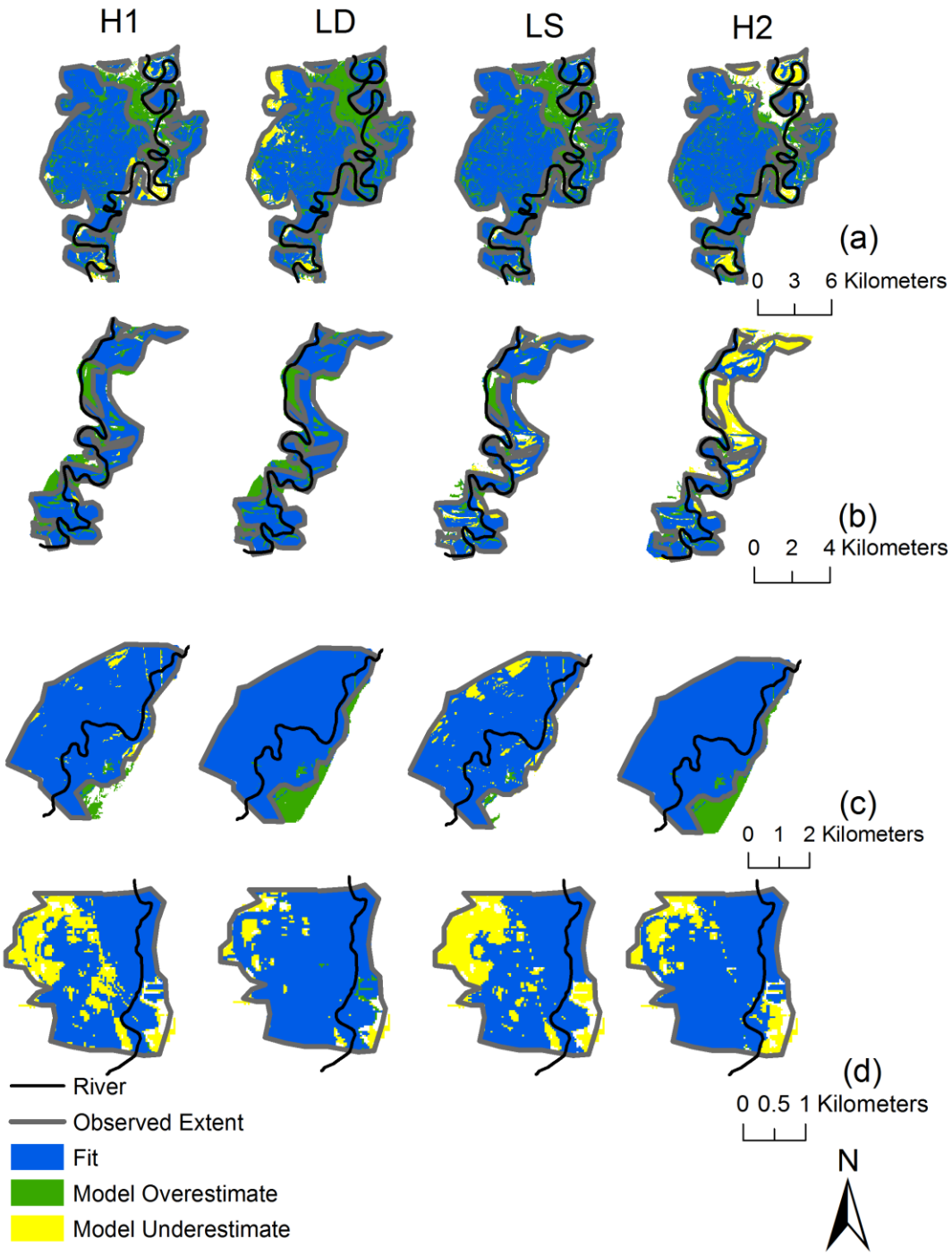

Figure 2.6 Comparison map when utilizing unique channel Manning's n value (nch=0.03) and distributed floodplain Manning's n value at (a) White River; (b) Black River; (c) East River and (d) Flatrock River.

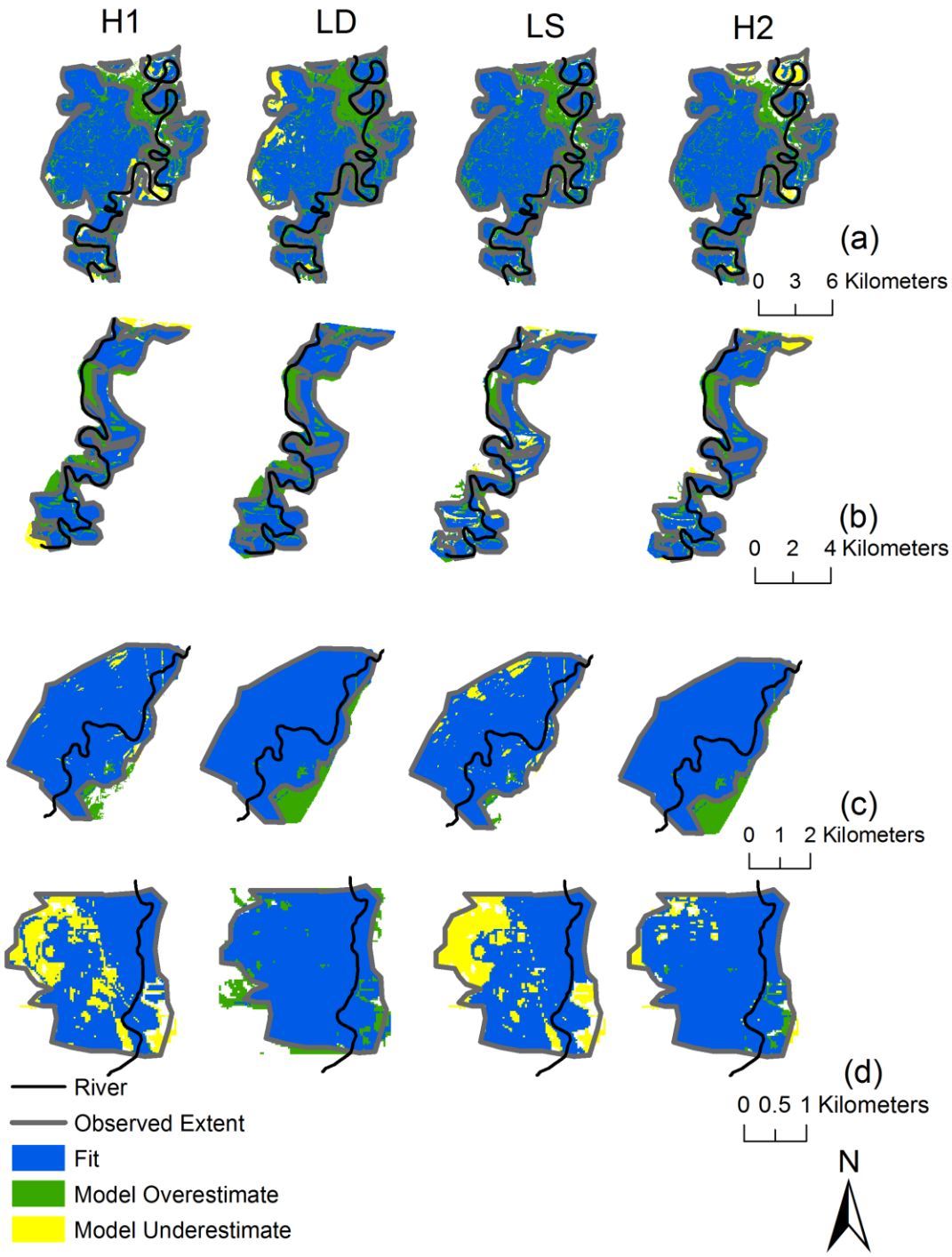

Figure 2.7 Comparison map when utilizing unique channel Manning's n value (nch=0.05) and distributed floodplain Manning's n value at (a) White River; (b) Black River; (c) East River and (d) Flatrock River.

2.6.2 Influence of channel surface roughness on model performance

Surface roughness plays an important role in simulating accurate flow hydrodynamics in both channel and the floodplain. The effect of channel Manning's n values on each model is evaluated by finding out the change in three performance measures: F and C indices (Fig. 2.8) and inundation area (Fig. 2.9). Overall, H1, LD and H2 models seem to be more sensitive to channel surface roughness at low n values compared with LS model for the four study reaches as both F and C indices change with Manning's n for these three models. Interestingly, LS shows sensitivity to channel surface roughness only for the East River because the floodplain is relatively more heterogeneous in terms of topography. Thus, as the channel roughness increases more water flows overbank to inundate the floodplain and the inundation area increases with channel roughness. In the case of more homogeneous topography, much of the floodplain gets inundated with overbank flow thus showing insensitivity to the channel roughness.

In addition, the F and C indices show sudden spike when the channel surface roughness increases beyond a certain value except for the East River. This is caused by the combined effect of the flow within the channel, channel cross-sectional area, channel slope and the definition of lateral structure in the H2 model. The large cross-sectional area keeps the flow in the channel up to a certain roughness value, thus underestimating the flood extent. Once the water enters the floodplain, the F and C indices show greater agreement with observed and simulated inundation extents. In the case of East River, which has a relatively high flow and small cross-sectional area, the water enters the floodplain even for low roughness values, thus eliminating the certain values for other reaches. This also explains why all models perform well for East River using any channel roughness value. As expected, Fig. 2.9 shows increase in flood inundation area with higher channel surface roughness for all models (reference line shows the observed inundation area) as increased roughness creates higher flow depth and inundation. The relative smaller

increase in inundation area for LS model than other models also indicates the relative insensitive of this model to channel surface roughness.

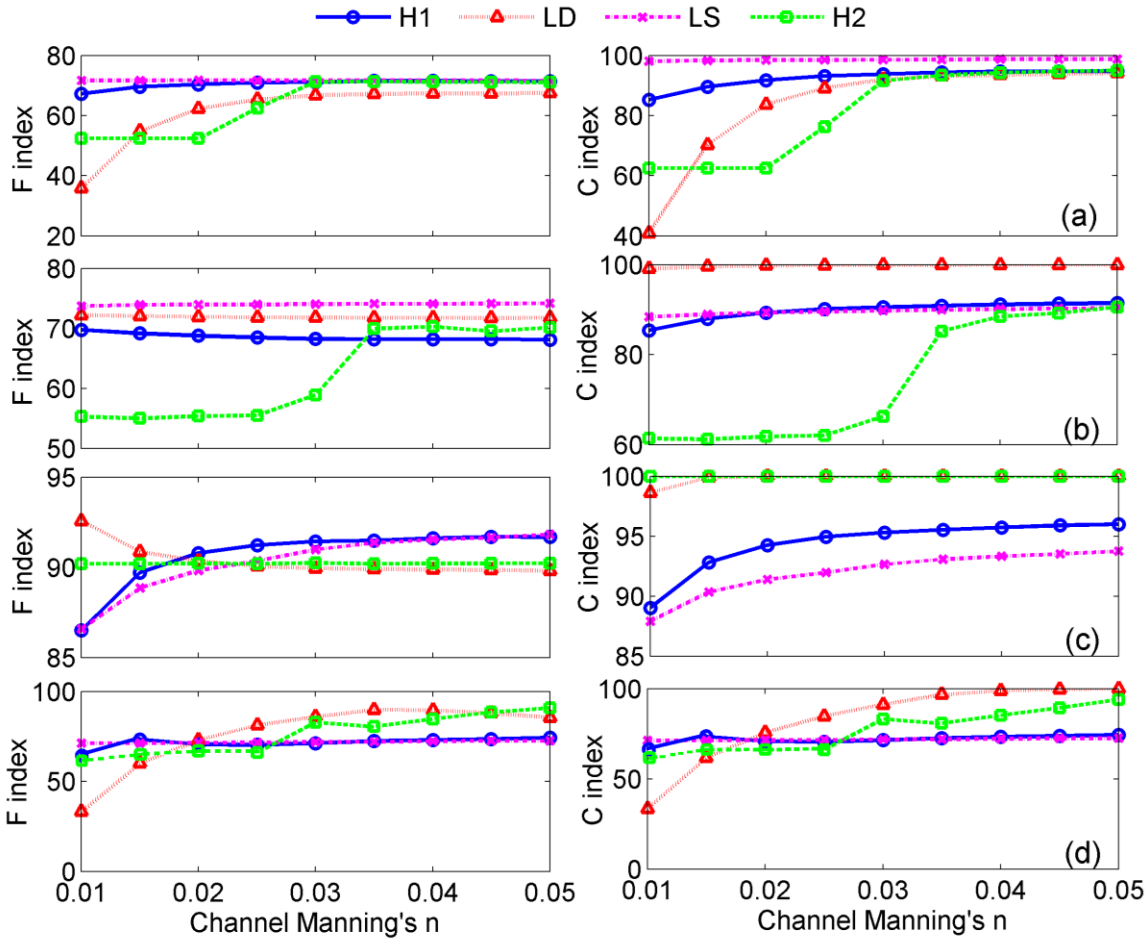

Figure 2.8 Effect of channel Manning's n values on hydraulic model outputs at (a) White River; (b) Black River; (c) East River and (d) Flatrock River.

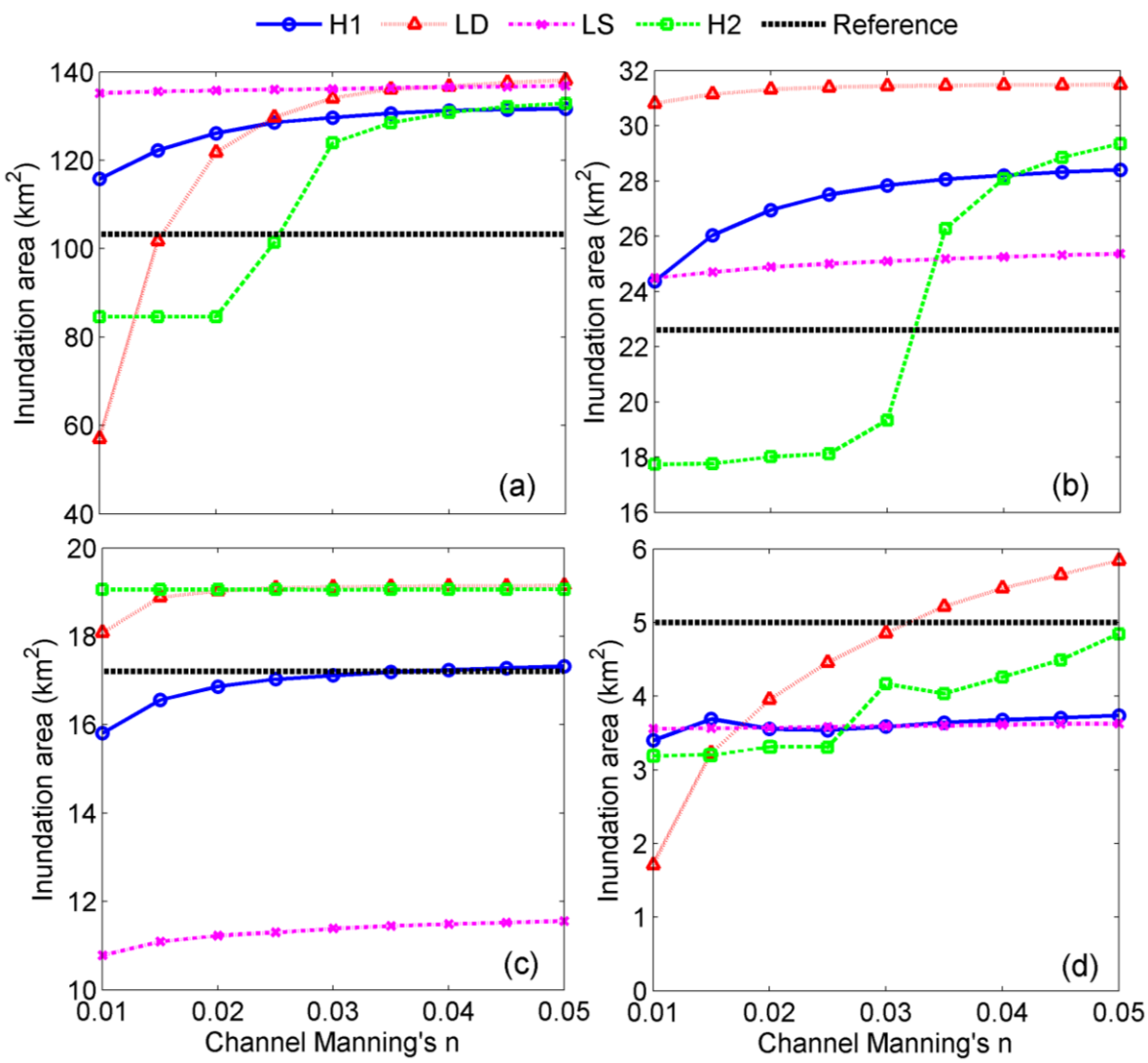

Figure 2.9 Effect of channel Manning's n values on model simulated inundation area at (a) White River; (b) Black River; (c) East River and (d) Flatrock River.

2.6.3 Effect of floodplain surface roughness on model performance

The simulation outputs from scenario 2 describe the model sensitivity to floodplain surface roughness when a single channel Manning's n (0.03 in this case) and single floodplain Manning's n value are used for each simulation. The comparison maps from Figs. 2.10-2.11 together with the F and C indexes from Fig. 2.12 show that H1 and LS models are relatively more sensitive to floodplain Manning's n value compared with the other two models. However, when results from

Fig. 2.12 are compared with results from Fig. 2.8, it is clear that floodplain roughness has relatively little effect on model simulation for both White and Black River. In the case of these two rivers, the LD and H2 models' performance improves significantly when the channel roughness value is higher (greater than 0.03) as seen in Fig. 2.8ab. At higher channel roughness values, the increase in floodplain roughness does not affect the model simulation as seen in Fig. 2.12. The LD model's insensitivity to the floodplain surface roughness shows that floodplain processes (due to momentum effects) are unimportant in the model, which has also been found in other studies (Horritt and Bates, 2002a) since the local acceleration and advection acceleration terms in the momentum equation (Equations 2.6-2.8) are neglected in this model. H1 model shows more sensitivity at low floodplain roughness value, but as the n value increases, the results are less affected. As seen from Figs. 2.10-2.11, the floodplain inundation from LS model increases for all reaches when the floodplain surface roughness change from 0.06 to 0.12. Additionally, the F and C indices increase dramatically for the Flatrock River (about 30). Fig. 2.13 shows that inundation area increases for both H1 and LS models as the floodplain surface roughness increases, but the change in flood inundation area is not significant for LD model and H2 model. Similar to Fig. 2.12, Fig. 2.13 indicates that floodplain roughness plays a trivial role for water propagation on floodplain process of LD and H2 model though it has relatively significant effect for H1 and LS model.

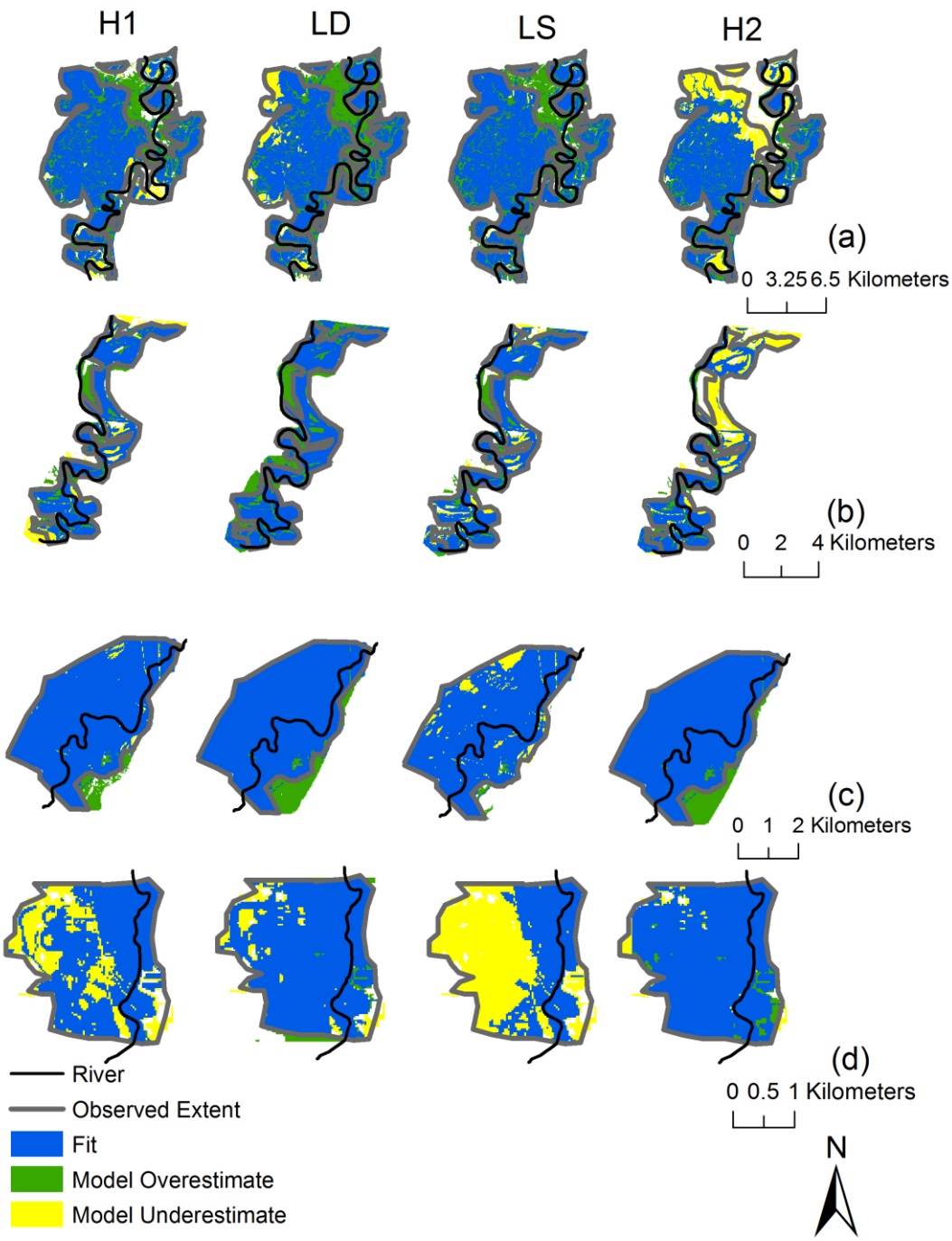

Figure 2.10 Comparison map when utilizing unique channel Manning's n value (n_{ch}=0.03) and unique floodplain Manning's n value (n_{pl}=0.06) at (a) White River; (b) Black River; (c) East River and (d) Flatrock River.

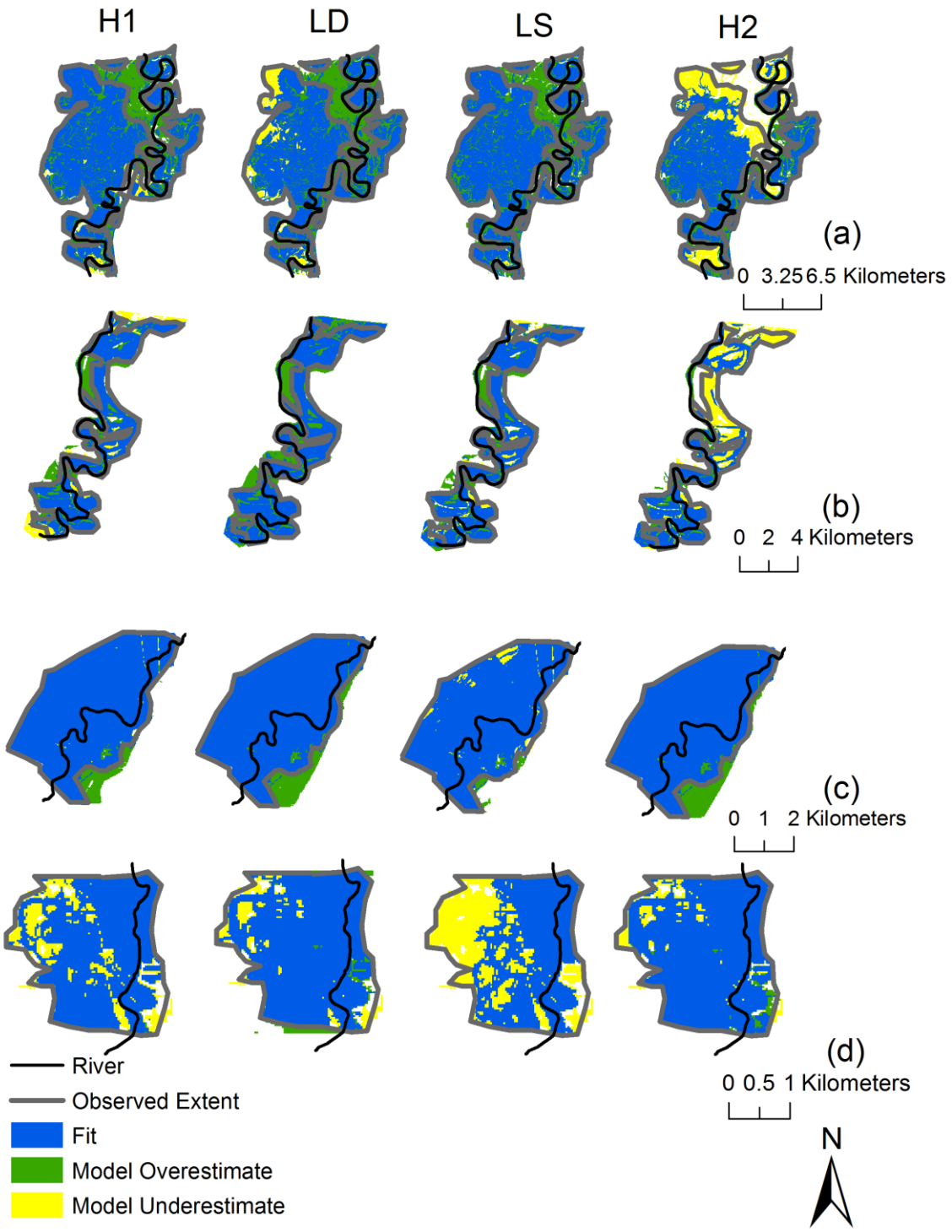

Figure 2.11 Comparison map when utilizing unique channel Manning's n value (nch=0.03) and unique floodplain Manning's n value (npl=0.12) at (a) White River; (b) Black River; (c) East River and (d) Flatrock River.

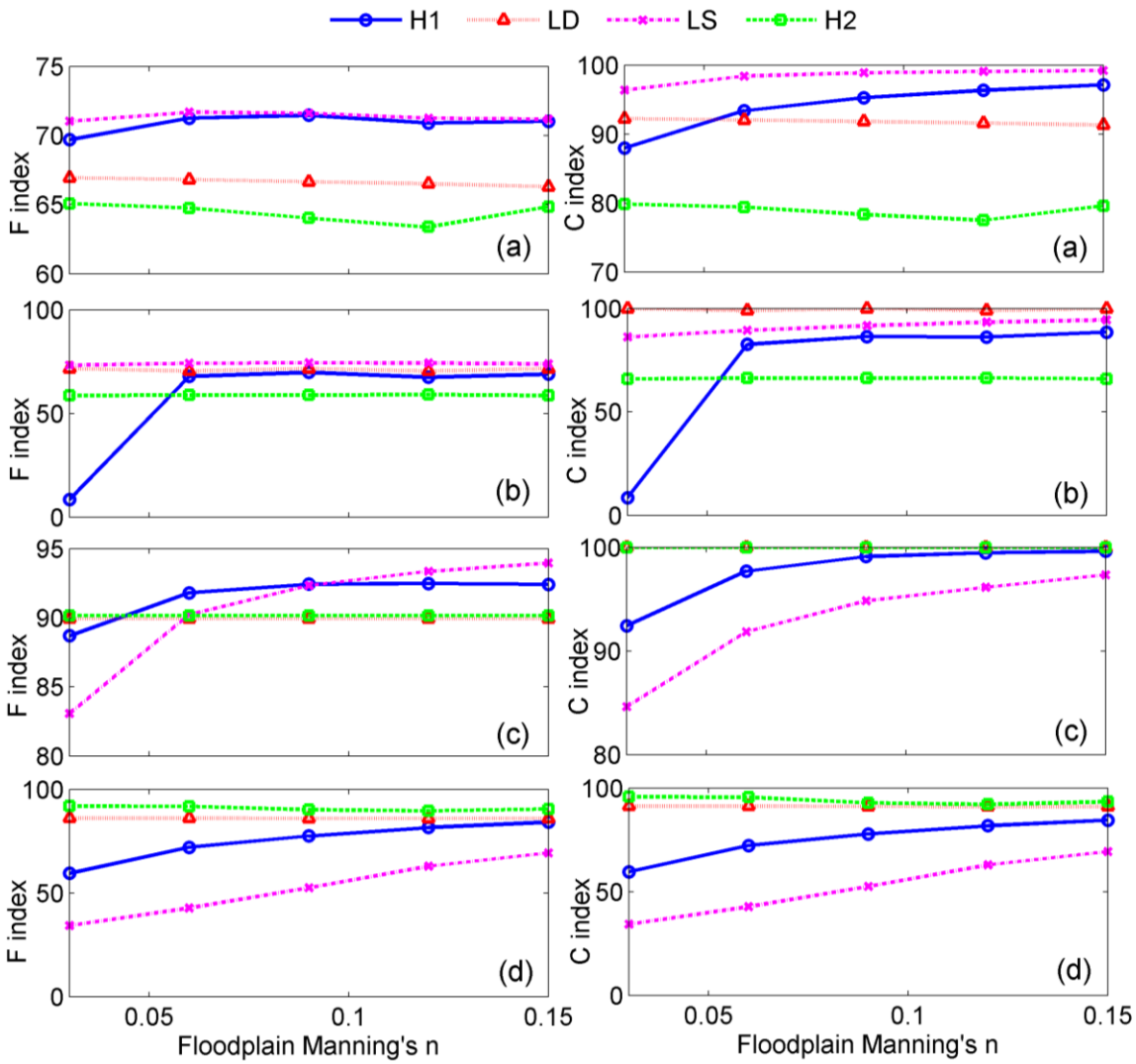

Figure 2.12 Effect of floodplain Manning's n values on hydraulic model outputs at (a) White River; (b) Black River; (c) East River and (d) Flatrock River.

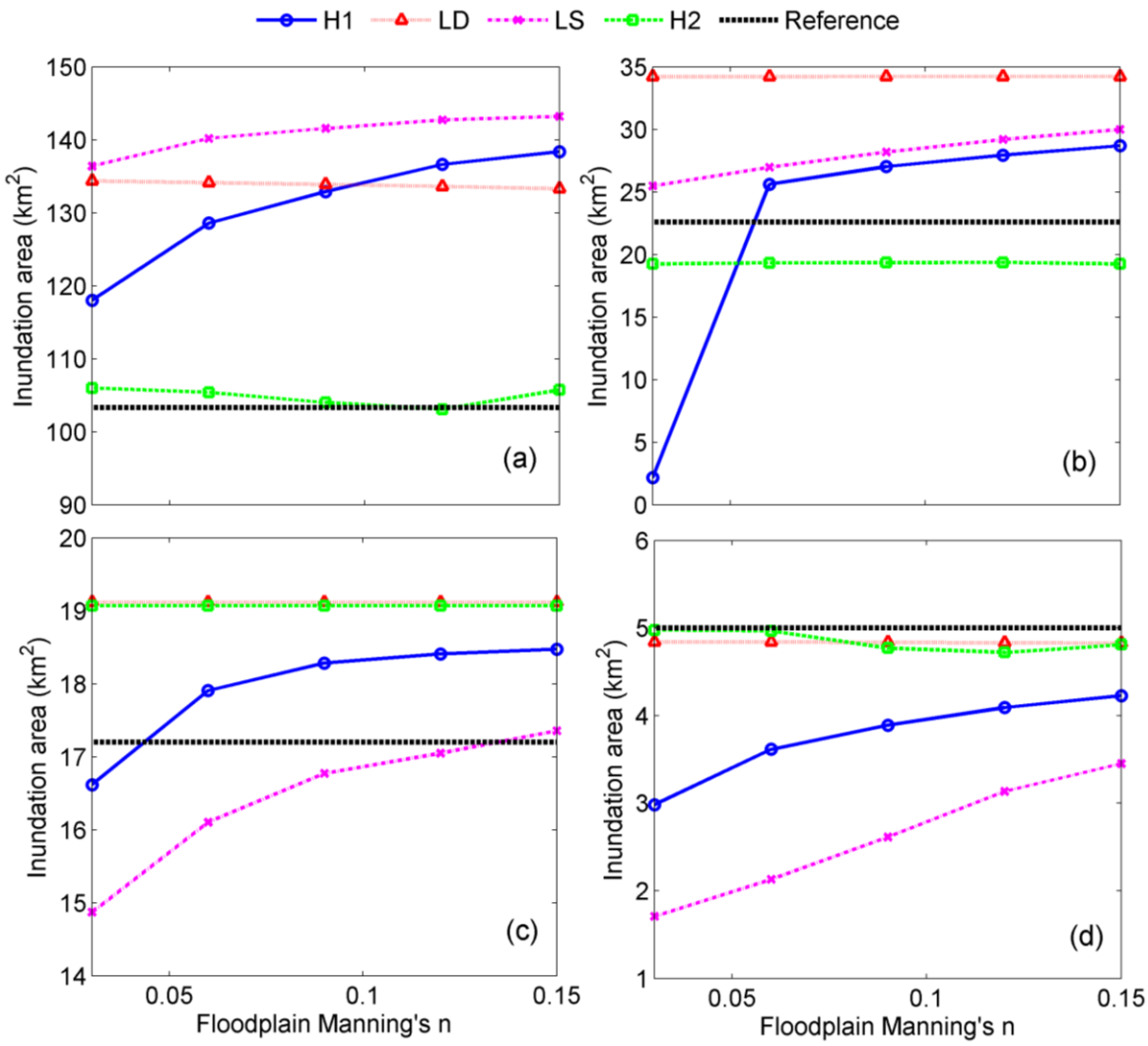

Figure 2.13 Effect of floodplain Manning's n values on model simulated inundation area at (a) White River; (b) Black River; (c) East River and (d) Flatrock River.

2.6.4 Comparison of applying uniform versus distributed floodplain surface roughness

Scenario 1 included distributed Manning's n for the floodplain whereas scenario 2 included single Manning's n value for the floodplain in each simulation. As it shown in Fig.2.8 and Fig. 2.12, no matter the single or distributed floodplain roughness values are used, the optimal solution that yielding largest F and C indexes is typically not from a small channel or floodplain roughness value. This indicates the lower roughness values in the parameter space usually underestimate the friction and could not reflect the true roughness condition in reality for our

study regions. In previous reach scale studies, combinations of single channel Manning's n and single floodplain Manning's n values are tested during the model calibration process to find out the combination that yields best model performance. The question of whether the model performance will improve when the distributed floodplain surface roughness is applied and the model is only calibrated by changing the channel Manning's n value is investigated. From the simulation results in scenario 1 and scenario 2, the maximum F index value derived from utilizing single floodplain Manning's n for each model is compared with the F index resulting from using distributed floodplain Manning's n values when using a fixed channel Manning's n of 0.03 (Table 2.5). Results indicate that maximum F index corresponding to applying single floodplain Manning's n is comparable or even better compared to applying distributed floodplain surface roughness in most situations. In order to further reinforce this finding, scenario 2 simulations are repeated for other channel Manning's n values (0.01, 0.02, 0.04, and 0.05), yielding another 320 simulations. The largest F resulting from these simulations is then compared with the largest F resulting from simulations with distributed Manning's n in the floodplain. The results (Table 2.6) show that, in general, using distributed floodplain surface roughness gives relatively poorer performance compared to using single Manning's n for the floodplain. This is because single floodplain Manning's n value results in more model parameter combinations and thus larger variance for the simulated inundation extents compared with distributed floodplain roughness values. Therefore, chance is higher for some of the generated inundation extents by combining single channel and floodplain to fit the observed inundation extent.

Table 2.5 Comparison of maximum F index by using uniform (U) and distributed (D) floodplain Manning's n when the channel Manning's n is 0.03

Models	HEC-RAS 1D		LISFLOOD-FP Diffuisve		LISFLOOD-FP Subgrid		HEC-RAS 2D	
Reach	U	D	U	D	U	D	U	D
White River	71.3	71.1	67.2	67.0	71.2	71.3	66.5	71.1
Black River	70.2	68.8	71.7	71.8	73.7	73.5	59.8	59.7
East River	92.3	91.3	89.9	89.9	93.2	91.1	90.1	90.1
Flatrock River	84.1	71.3	86.6	86.5	69.1	71.8	92.7	83.2

Table 2.6 Comparison of maximum F index by using uniform (U) and distributed (D) floodplain Manning's n when the channel Manning's n ranges from 0.01 to 0.05

Models	HEC-RAS 1D		LISFLOOD-FP Diffuisve		LISFLOOD-FP Subgrid		HEC-RAS 2D	
Reach	U	D	U	D	U	D	U	D
White River	71.8	71.5	67.7	67.5	71.8	71.8	71.5	71.3
Black River	71.2	69.7	72.0	72.2	74.8	74.2	72.0	70.3
East River	92.8	91.7	92.6	92.6	94.2	91.8	90.2	90.2
Flatrock River	97.0	93.3	89.7	89.9	69.3	72.6	93.0	90.8

2.7 Conclusions

Flood risk assessment and management rely on the accuracy of flood extent simulated by a hydrodynamic model. Faster and more accurate flood inundation predictions call for the end

users to select an appropriate a hydraulic model and its parameters that could balance his/her needs against model complexity and data requirements (Teng et al., 2017a). The differences in model structures in representing river geometry, governing equations and numerical discretization as well as model parameterization lead to different outputs. Four commonly used hydraulic models including HEC-RAS 1D model (H1), LISFLOOD-FP diffusive model (LD), LISFLOOD-FP subgrid model (LS) and HEC-RAS 2D model (H2) are used with same input data and boundary condition to perform unsteady flow analysis for four historical flood events at four reaches in the U.S.

The following conclusions are drawn in relation to the objectives of this study:

Impact of Model Structure: Overall for a given set of roughness condition, the geometry, including the sinuosity, reach length and floodplain width, did not seem to play any major role in dictating the performance of a 1D or 2D model. With regards to HEC-RAS 1D (H1) and HEC-RAS 2D (H2), H1 performed better for Flatrock River at low channel roughness conditions. However, as more water entered the floodplain at high channel roughness values, the performance of H2 and H1 were comparable, with H2 providing slightly better performance. Even for the most sinuous White river reach with the widest floodplain, H1's performance is comparable to the other 2D models. For all study sites, the performance of the subgrid version of LISFLOOD-FP (LS) was more stable under different channel roughness conditions, and it performed better than the diffusive version (LD) in simulating floodplain inundation, except for the Flatrock River.

Influence of channel surface roughness and floodplain surface roughness in simulating flood extents: Channel roughness affected the performance of both H2 and LD more compared to other two models. Both H2 and LD performed better when the channel roughness was high to let more

water flow into the floodplain. Interestingly, H1 and LS were found to be more sensitive to floodplain roughness, and their performance improved as the floodplain roughness increased. Floodplain roughness did not affect the overall floodplain hydrodynamics in both LD and H2 models for all study reaches. It is found that all models, that did not perform satisfactory for a given value of channel and floodplain roughness, improved with increasing roughness value. Relative influence of distributed versus uniform floodplain surface roughness characterization: Applying distributed floodplain roughness did not necessarily improved model performances for any reach included in the study. The conventional approach of using unique channel Manning's n value together with unique floodplain Manning's n value yielded better or similar results compared to using distributed floodplain roughness. It is possible that the water depth simulated in the floodplain may have some role to play in this finding, but this needs to be explored further.

Many past studies have shown 2D models to perform better than 1D model with respect to hydrodynamics(Morales-Hernández et al., 2013; Quiroga et al., 2016; Şensoy et al., 2016), but in some cases, such as in steep topographies with no sinuous geometries, 1D models can give equally acceptable results. Application of 1D and 2D models for reaches with different topography and sinuous geometries shows that 1D and 2D models can give equal results with the right combination of channel and floodplain roughness characterization.

CHAPTER 3. ACCOUNTING FOR MODEL STRUCTURE, PARAMETER AND INPUT FORCING UNCERTAINTY IN FLOOD INUNDATION MODELING USING BAYESIAN MODEL AVERAGING

3.1 Abstract

Reliability of flood stage and inundation extent predictions are affected by the performance of a hydraulic model. However, uncertainties at all times exist in the model setup process. Therefore, prediction from a single hydraulic model implementation may be subject to huge uncertainty. Bayesian model averaging (BMA) is applied in this study to combine ensemble predictions from different hydraulic model implementations and to develop a robust deterministic water stage prediction as well as the prediction distribution. The BMA approach is tested over the Black River watershed in Missouri and Arkansas based on water stage predictions from 81 LISFLOOD-FP model configurations that integrate four sources of uncertainty including channel shape, channel width, channel roughness and flow input. Model ensemble simulation outputs are trained with observed water stage data during one flood event to obtain the weight and variance for each model member, and BMA prediction ability is then validated for another flood event. The results indicate that the BMA approach is able to provide consistently good and reliable deterministic flood stage prediction across the basin, though it does not always outperform the best model in the ensemble. The BMA water stage prediction has better performance than the ensemble mean prediction. Additionally, high-chance flood inundation extent derived from a BMA probabilistic flood map is more accurate than the probabilistic flood inundation extent based on the equal model weights in the Black River watershed.

3.2 Introduction

With the increasing threat of frequently occurring intense storms, hydrodynamic models are expected to play a bigger role in understanding and predicting the floods and their corresponding extents. There are several ongoing efforts to simulate floods at multiple spatial scales ranging from single reach to continental scale stream networks (Cook and Merwade, 2009; de Paiva et al., 2013; Horritt and Bates, 2002a; Knebl et al., 2005; Schumann et al., 2013). Additionally, all these efforts use different approaches ranging from simplistic digital elevation model (DEM) based models such as HAND (Nobre et al., 2011) to more sophisticated 1D or 2D hydraulic models such as HEC-RAS and LISFLOOD-FP (Pappenberger et al., 2005a; Wood et al., 2016a). All approaches require two primary inputs including the topography to construct the river geometry and flow magnitude to simulate the hydraulics. Simulation of hydraulics require adjustment of model parameters, which is primarily the channel roughness specified in the form of Manning's n. Depending on the flood modeling approach, the final result is affected by several sources of uncertainty including the model structure, flow magnitude, topography and model parameters, among others (Bermúdez et al., 2017; Cook and Merwade, 2009; Dottori et al., 2013; Mukolwe et al., 2016; Teng et al., 2017a).

The uncertainty in flood inundation modeling can be categorized into three major types: model structure, model parameter and input forcing. Model structure broadly includes the type of the model, one-dimensional or two-dimensional, type and form of numerical equations, and the assumptions in the model. For a specific model, structural uncertainty could also include how the river geometry, including the channel cross-sectional shape and planform, is extracted and represented in the model (Liu et al.; Pappenberger et al., 2006a; Teng et al., 2017b). For instance, commonly used DEMs do not have information on channel bed, and thus assuming a shape for the channel bed could add substantial uncertainty to the model output. Most hydraulic models are

calibrated for different flows using channel roughness parameter, and thus when these models are used to simulate flows that are outside the range used during calibration, the calibrated channel roughness value can add uncertainty to the model results. This is typically the case when hydraulic models are used for simulating 100-year or higher return period design flows as observed data for such flows may not exist. Finally, observed or simulated input forcing data such as streamflows and water stage used in hydraulic models also add significant uncertainty to the simulation results (Beven and Hall, 2014; Demeritt et al., 2007; Merwade et al., 2008; Pappenberger et al., 2005c). Therefore, reliance on a single hydraulic model implementation for flood prediction typically increases the statistical bias of the forecast.

One way to handle model uncertainties is by using a multi-model combination approach, in which results are extracted from a group of existing model implementations to provide a robust prediction based on the model prediction ensemble. Multi-model combining for ensemble predictions is widely used in hydrology and climate forecast using a variety of methods, including SMA (simple model average), WMA (weighted model average), MMSE (multi-model super-ensemble) and M3SE (a variant of MMSE), among others (Ajami et al., 2006; Chowdhury and Sharma, 2009; Hamill, 2001; Liu et al., 2014; Najafi and Moradkhani, 2015a; Najafi and Moradkhani, 2015b; Shamseldin et al., 1997; Xiong et al., 2001). Additionally, the performance of probabilistic ensemble merging techniques have been evaluated and compared with the deterministic model predictions for climate model and streamflow predictions (Najafi and Moradkhani, 2015a; Najafi and Moradkhani, 2015b). In recent years, the Bayesian model averaging (BMA) technique (Merlise, 1999; Raftery et al., 2005) has been used widely in surface water hydrology (Ajami et al., 2007; Duan et al., 2007; Jiang et al., 2017; Rings et al., 2012; Zhang et al., 2009), groundwater hydrology (Neuman, 2003), climatology (Zhang et al., 2016b),

biology (Yeung et al., 2005), ecology (Wintle et al., 2003), public health (Morales et al., 2006), and economics (Fernandez et al., 2001). The rationale behind the BMA method lies in the fact that some models are superior to other models and each model should not be treated exactly the same. The BMA approach evaluates model implementations and assigns each of them a weight and variance based on the model performance in the training period. The advantage of this approach over other model combing methods is that BMA not only provides a deterministic model weighted average prediction of the interested variable, but also produces the forecast distribution which reflects the uncertainty associated with the deterministic prediction (Raftery et al., 1997; Rings et al., 2012).

Considering the wide applicability of BMA in other areas of hydrology, its application in flood inundation modeling can help address the issue of accounting and presenting model structure, parameter, and input forcing uncertainty. Although the BMA approach is usually being used to account only for model structure uncertainty, there do exist previous studies in other fields to incorporate forcing/boundary condition and parameter uncertainty using the BMA method (Chitsazan and Tsai, 2015; Yen, 2012). Accordingly, the objectives of this study are to: (1) determine whether BMA can provide accurate and reliable deterministic flood predictions for a stream network by considering various uncertainty sources; (2) compare the performance of BMA prediction with predictions from model members and ensemble mean; and (3) quantify uncertainty associated with the BMA deterministic prediction. The above objectives are accomplished by applying a large scale hydraulic model for the Black River watershed that is located in Arkansas and Missouri in the U.S.

We acknowledge that uncertainty quantification is not a new topic in flood inundation mapping because several past studies have addressed uncertainties related to parameter, input data and

boundary conditions (Aronica et al., 2002; Jung and Merwade, 2011; Pappenberger et al., 2005a; Pappenberger et al., 2013; Tiwari and Chatterjee, 2010; Yu et al., 2015). However, the results from these studies are somewhat limited due to the use of isolated flood events on a single reach in the analysis. With the growing need to simulate the river hydrodynamics over an entire stream network at basin to continental scales (Huang and Hattermann, 2018; Jafarzadegan and Merwade, 2017; Merwade et al., 2018; Schumann et al., 2013; Wilson et al., 2007), it is expected that multiple sources of uncertainties will play different roles in different streams to contribute to the overall uncertainty in the final result. Hydrodynamic modeling in a stream network will involve a mix of large and small reaches. Many large reaches could be well described in the model in terms of their cross-sectional shape, roughness characterization and channel width, but the same may not apply to many low order contributing streams. As a result, many low order streams may be affected by structural, parameter and input forcing uncertainties, but the large streams may only be affected by input forcing uncertainties. Thus, there is a need to understand the cumulative effect of different uncertainty sources in flood inundation modeling over a larger stream network. This study attempts to address this need by using the BMA methodology. Once the cumulative effect of all uncertainties is accounted, the hierarchical BMA can then be used to understand the relative impact of individual uncertainties.

3.3 Study area and data

The 20,000 km^2 Black River watershed located across Arkansas and Missouri states in the U.S. is selected for this study. Historical records indicate that this region has experienced numerous flood events in the past including the recent one that occurred in May 2017. The recorded water level data for the 2017 event are available from the United States Geological Survey (USGS) gauges (Figure 3.1 and Table 3.1), and are used in validating the results from this study.

Additionally, the Black River watershed has four major rivers including Black river, Current River, Eleven Point River and Spring River, which drain towards the watershed outlet in Arkansas. These four major rivers provide distinct topographical, geomorphic settings as well as varying reach lengths and sinuosity, thus making Black River watershed a good test bed for this study. The daily streamflow data (input to the hydraulic model) and stage data (for validation) used in this study are unaffected by any major hydraulic structures, and are thus considered natural. The topography and land use dataset are obtained in the form of national elevation dataset 90m DEM (http://ned.usgs.gov) and NLCD 2011 land use data (http://www.mrlc.gov/nlcd2011.php), respectively. A 90m DEM instead of 30m or finer resolution DEM is selected for this study to strike a balance between predication accuracy and the computational demand for a large number of simulations needed for uncertainty analysis. Although the prediction accuracy might be affected by using the 90m DEM, many large scale models use 90m or coarser DEM for flood modeling as well (Neal et al., 2012a; Schumann et al., 2013). Inundation extents for selected storms, which will be used for validation of results, are derived by classifying Landsat images from the USGS earth explorer website (http://earthexplorer.usgs.gov/). The Landsat images are classified into water and non-water area with a supervised classification technique using the ArcGIS classification tools. The classification is formed using the following three steps: (i) train the tool by delineating water and non-water areas; (2) use the maximum likelihood classification approach to classify the entire Landsat image based on information obtained from the training areas; and (iii) extract the "water" area from the classified image and treat it as observed inundation extent around the streams.

Table 3.1 USGS gauge IDs for the streamflow input and validation stations

Streamflow Input Station	USGS Station ID
S1	07069305
S2	07071500
S3	07065495
S4	07064533
S5	07068510
S6	07061900
S7	07061500
Validation Station	
A	07069500
B	07072000
C	07068000
D	07064000

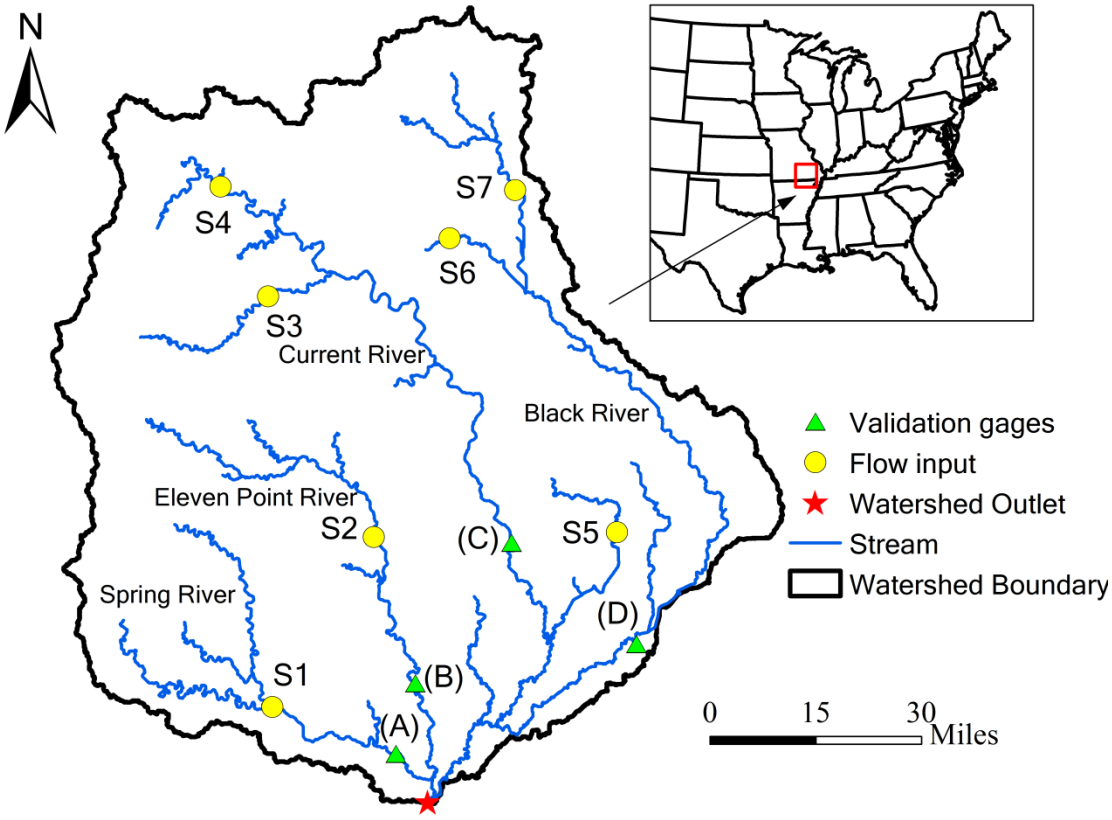

Figure 3.1 Black River watershed and layout of input and validation USGS gages

3.4 Methodology

3.4.1 Hydraulic modeling

Many one-, two- or three-dimensional models exist for conducting flood simulations. One-dimensional (1D) models use discrete cross-sections to describe the rivers and assumes that water moves only longitudinally along the direction of river; whereas two-dimensional models (2D) use continuous mesh or raster grid to define the channels and assumes water moves both longitudinally and laterally; and a three-dimensional (3D) model adds the vertical movement to the 2D flow. One of the most commonly used models in the U.S. is the Hydrological Engineering Center's River Analysis System (HEC-RAS) (USACE, 2015) which includes the classic 1D and the recently developed (2D) versions. HEC-RAS 1D model requires manual digitization of river features such as the center line, banks, and cross-sections using GIS. Application of HEC-RAS 1D model, which is commonly used for simulating few reaches within a network, can become tedious for a larger stream network containing hundreds of stream segments. The 2D HEC-RAS model overcomes the significant amount of digitization work. However, considering this version still exists at its early age and more validation work need to be done, it is not used in this study. A recently developed model, AutoRAPID (Tavakoly, 2015), incorporates a continental scale river routing model known as Routing Application for Parallel Computation of Discharge (RAPID) with a regional scale flood delineation model called AutoRoute to generate flood inundation. AutoRAPID is a computationally efficient model, which is not meant for detailed hydrodynamic modeling, and hence its results are less accurate compared to HEC-RAS and other models. LISFLOOD-FP model is a raster based 1D/2D coupled hydraulic model that has been widely used in the flood inundation modeling community in many parts of the world including Europe, North America, Latin America, Africa and Asia (Altenau et al., 2017; Komi et al., 2017; Wilson et al., 2007; Wood et al., 2016b; Wu et al., 2017).

The latest subgrid version of LIFLOOD-FP is easy to implement for large scale flood simulations and yields accurate predictions (Bates and De Roo, 2000; Neal et al., 2012a). Considering the amount of efforts in model set-up, computational demand and the results from past studies, LISFLOOD-FP is selected in this study. LISFLOOD-FP solves continuity and momentum equations to obtain the information of streamflow and water depth in each raster cell. The input data required by the LISFLOOD-FP model includes the DEM, flow values, stream centerline, and channel geometry information including average river width, slope, surface roughness and channel cross sectional shape. The floodplain roughness values used in this study are obtained from the NLCD land use dataset (Homer et al., 2015). NLCD is raster dataset specifying the land use for each pixel, which can then be related to a specific Manning's n value as described in (Kalyanapu et al., 2010). Two six-month daily time-step simulations are used for training (12/2007 to 05/2008) and validation (01/2011 to 06/2011), respectively. These two periods are selected based on the flood events and availability of input and validation data. A total number of 81 LISFLOOD-FP model configurations, described in the next section, are created and simulated using High Performance Computing (HPC) resources at Purdue University. A two month warm-up period is used in the training and validation simulations to remove unrealistic predictions, and thus the outputs in the warm-up periods are excluded from the BMA analysis.

3.4.2 Bayesian Model Averaging method

Dynamic model simulation outputs are subject to uncertainty when compared with observations. For any simulation ensemble, some models have relatively superior performances compared to others**.** Therefore, instead of relying on the results of one specific model, extracting information from the model simulation ensemble is more reliable (Ma et al., 2018; Raftery et al., 2005).

Bayesian model averaging (BMA) is a statistical approach to combine estimations from individual models and produce reliable prediction. The BMA predictive PDF of a variable, such as water depth in this study, is the weighted average of the PDFs associated with each model member forecast in the ensemble. The weight could reflect the prediction skill of a member in the training period. This method has several desirable properties, one of which is that the BMA approach not only provides the deterministic model weighted average prediction, but also provides an associated prediction distribution. In addition, instead of trusting one "best" model, the BMA approach considers the performance of all model predictions in the training period. To predict flood stage d on the basis of training data d^T using K models, the law of total probability indicates that the forecast PDF, $p(d)$, is given by Equation 3.1.

$$\mathrm{p(d)} = \sum_{\mathrm{k=1}}^{\mathrm{K}} \mathrm{p(d|M_k)} \times \mathrm{p(M_k|d^T)} \qquad (3.1)$$

Where $p(d|M_k)$ is the forecast PDF based on model M_k alone, and $p(M_k|d^T)$ is the posterior probability of model M_k given the observation data in the training period. In the BMA theory, the posterior model probability $p(M_k|d^T)$ represents the weight for each model member in the ensemble and all the weights add up to one (Equation 3.2). The model prediction is associated with a conditional PDF $g_k(d|M_k)$, which can be illustrated as the conditional PDF of d conditioned on M_k, given that M_k is the best forecast. Then the BMA predictive PDF is obtained using Equation 3.3. Typically, it is assumed that the PDF of predicted quantity follows a normal distribution centered at M_k, with standard deviation σ_k (Equation 3.4). A deterministic water stage forecast can also be obtained through the BMA method. The BMA deterministic prediction is the expectation of the quantity given the model predictions, which can be calculated using Equation 3.5.

$$\sum_{\mathrm{k=1}}^{\mathrm{K}} \mathrm{p(M_k|d^T)} = 1 \qquad (3.2)$$

$$p(d|M_1, \dots, M_K) = \sum_{k=1}^{K} w_k \times g_k(d|M_k) \qquad (3.3)$$

$$d|M_k \sim N\left(M_k, \sigma_k^2\right) \qquad (3.4)$$

$$E(d|M_1, \dots, M_k) = \sum_{k=1}^{K} w_k M_k \qquad (3.5)$$

3.4.3 Box-Cox transformation and Expectation Maximization (EM) method

For non-Gaussian data, a power transformation is needed to map the variables from their original space to the Gaussian space. The observed and simulated water stage data are normalized using the Box-Cox power transformation (Box and Cox, 1964; Osborne, 2010). This method is used to transform the skewed data into normal distribution using Equation 3.6. When the power λ is zero, the natural logarithm is applied to the data. Different powers are tested and the best power that yields the closest normal distribution of transformed water stage data is found and applied. The transformed water stage data is used to find the BMA weights for each model member. In order to estimate the parameter w_k and σ_k, on the basis of training data set, the log likelihood function (Equation 3.7) is maximized.

$$f_t = \begin{cases} \frac{f^{\lambda}-1}{\lambda}, & \lambda \neq 0 \\ \ln(f), & \lambda = 0 \end{cases} \qquad (3.6)$$

$$l(\theta) = \sum_{s,t} \log\left(\sum_{k=1}^{K} w_k g_k\left((d_{st}|M_{kst})\right)\right) \qquad (3.7)$$

Where λ is the power that used to transform the data to a normal distribution; f is the data in its original space and f_t is the data after transformation. The summation in Equation 3.7 is over locations (s) and time (t) in the training data. Generally, water stage data are correlated in time, but the correlation in prediction errors is assumed to be weakly correlated, and is not expected to have significant impact on the results. Additionally, Equation 3.7 computes the conditional distribution/expectation for a scalar observation given forecasts, rather than for several observations simultaneously (Raftery et al., 2005). Equation 3.7 is difficult to solve analytically

or numerically (Vrugt and Robinson, 2007; Zhu et al., 2016), and thus an Expectation Maximization (EM) algorithm is used to find out the maximum likelihood. The EM algorithm is iterative and alternates between two steps, the E (or expectation) step, and the M (or maximization) step by using a latent variable *z*. In the E step, z is estimated given the current estimates of the model weight w_k and σ_k (Equation 3.8).

$$z_{kst}^{j} = \frac{w_k^{j-1} g\left(d_{st} \middle| M_{kst}, \sigma_k^{j-1}\right)}{\sum_{i=1}^{K} w_i^{j-1} g\left(d_{st} \middle| M_{ist}, \sigma_i^{j-1}\right)} \tag{3.8}$$

Where the superscript *j* refers to the jth iteration of the EM algorithm and $g\left(d_{st} \middle| M_{kst}, \sigma_k^{j-1}\right)$ is a normal density with mean M_{kst} and standard deviation σ_k^{j-1}. In the M step, the weight w_k and standard deviation σ_k are calculated with the current estimate of z_{kst} (Equations 9-10).

$$w_k^{j} = \frac{1}{n} \sum_{s,t} z_{kst}^{j} \tag{3.9}$$

$$\sigma_k^{2(j)} = \frac{\sum_{s,t} z_{kst}^{j} (d_{st} - M_{st})^2}{\sum_{s,t} z_{kst}^{j}} \tag{3.10}$$

Where *n* is the number of observations including distinct values of locations and time. The E step and M step are iterated to convergence when the changes of weight, variance, latent variable *z* and log-likelihood are smaller than some small tolerances (10^{-6} in this study). The log-likelihood is guaranteed to increase at each EM iteration, therefore, finally it converges to a local maximum (Wu, 1983).

3.4.4 Sources of uncertainty and LISFLOOD-FP configurations

Epistemic uncertainty in the LISFLOOD-FP model comes from a variety of sources including the assumed cross-section shape of the channel, channel width, channel roughness, DEM resolution, streamflow input and exclusion of hydraulic structures (Bales and Wagner, 2009; Merwade et al., 2008; Papaioannou et al., 2017; Pappenberger et al., 2006b). Among these

sources, uncertainty in cross sectional shape, channel width, model parameter and streamflow input are considered in this study. It is acknowledged that topography plays an important role in hydraulic modeling, but its impact varies depending on the data source and quality (Fernandez et al., 2016). The U.S. National Elevation Dataset (NED) DEM used in this study is derived from LIDAR with relatively higher accuracy so it is excluded from this study. Uncertainty arising from hydraulic structures such as bridges/culverts is also excluded due to the lack of detailed data at the stream network scale.

The uncertainty in streamflow estimates at USGS gaging stations can be as high as 10% for stable natural channels (Harmel et al., 2006a; Slade, 2004). Thus, three streamflow estimates including 0.9Q, Q and 1.1Q are used. Hydraulic models require the description of channel banks to separate the floodplain from the main channel. In LISFLOOD-FP, this is typically done by specifying the channel width along the centerline. Generally, the channel banks or width data are obtained based on aerial images captured during non-flooding conditions. This process adds substantial uncertainty in estimating the channel width depending on the time of the year when the aerial images are captured and the number of locations used in getting the average channel width. The uncertainty in channel width is incorporated by creating a distribution of channel widths from manually digitized banks and cross sections for four major rivers in the Black River Watershed, and then picking the 5%, mean and 95% values from this distribution (Table 3.2).

Both channel and floodplain roughness play an important role in the hydrodynamic simulations, but it is assumed that the floodplain roughness in this study, which is derived from the land use data is less uncertain compared to the channel roughness. A recent study by Liu (Liu et al., 2018) also found the channel roughness to affect the inundation extent more than the channel roughness. Hence, only the uncertainty from channel roughness is included in this study. The uncertainty in

channel roughness, Manning's n value, is incorporated by selecting three values ranging from 0.01-0.05. This range is selected based on previous studies (Horritt, 2006; Yu et al., 2013; Yu et al., 2015) to reflect the roughness of most natural channels characterized from a clean, straight channel to a winding, grassy and graveled channel. Finally, the uncertainty in cross-sectional shape is incorporated by assuming three shapes, namely triangular, parabola and rectangle. Considering the aforementioned four uncertainty sources and three varied types for each source, 81 (3×3×3×3) LISFLOOD-FP model configurations are created by using different combinations of these uncertainty sources (Table 3.3).

Table 3.2 Widths of 5% (low), 50%(average), and 95% (high) thresholds for each river

Width (m)	Low (5%)	Average (50%)	High (95%)
Spring River	14	35	71
Eleven Point River	28	51	96
Current River	34	87	154
Black River	74	181	355

Table 3.3 Four sources of uncertainty and their varied types

Uncertainty source	Channel shape	Channel width	Channel roughness	Input flow
Type 1	Rectangle	95% high value	0.05	1.1 USGS data
Type 2	Parabola	Average	0.03	USGS data
Type 3	Triangle	5 % low value	0.01	0.9 USGS data

3.4.5 Evaluation of BMA prediction in training and validation periods

Two types of BMA predictions are considered in this study. In the first type, named BMA station (BMA_S), the simulation outputs are trained with the observed water stage data at each station separately to obtain the corresponding weights for model members. Specifically, the BMA weights are determined by maximizing the likelihood function (Eq. 3.7) over time for each location separately (only considering the temporal component). Thus, one set of weights is derived separately for each of the four validation stations. In the second type, named BMA global (BMA_G), the BMA weights are determined by maximizing the likelihood function by including simulation and observed data from all four validation stations to give one set of weights (considering the spatial-temporal components). To validate the BMA predictive ability, the BMA_S and BMA_G weights are applied to model simulation outputs of another flood event, and the predictions are evaluated with the observed flood stage. In addition, BMA predictions are also compared with an ensemble mean prediction for each station, and the root mean square error values of the water stage (Equation 3.11) are calculated and ranked for each BMA model prediction and the ensemble mean prediction. Besides water stage, flood inundation extent is also compared with observed data using F and C indices (Equation 3.12-3.13,(Horritt and Bates, 2001a; Sangwan and Merwade, 2015)). Finally, a probabilistic flood map is created by applying the BMA_G weights to each raster grid cell (Equation 3.14) based on ensemble flood inundation extent predictions. The result is compared with the probabilistic flood map derived from the ensemble mean prediction (Equation 3.15) in which all models have equal weights. These two types of probabilistic flood inundation extents are also evaluated using F and C indices.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_{obs})^2}{n}} \qquad (3.11)$$

$$F = \frac{A_{om}}{A_o + A_m - A_{om}} \qquad (3.12)$$

$$C = \frac{A_{om}}{A_o} \quad (3.13)$$

$$Prob_{BMA_G} = \sum_{i=1}^{k} w_i \times 1 \quad (3.14)$$

$$Prob_{ensemble\ mean} = \frac{k}{n} \quad (3.15)$$

Where y_i stands for model predicted water stage, y_{obs} is the observed water stage, n is the total number of model configurations, k is the number of model configurations that predict the cell is flooded and w_i is the weight for model. i. A_{om} refers to area that is both observed and simulated as inundation, A_o refers to the observed area of inundation, A_m is the model simulated flood inundation area. F and C indices stand for fit and correctness, respectively, and range from 0 (worst) to 1 (best).

3.5 Results and discussion

3.5.1 Convergence of Expectation Maximization iterations and final model weights

The four indicators of the Expectation Maximization algorithm, including the maximum absolute change of weight, maximum absolute change of logarithmic variance, maximum absolute change of latent variable z and maximum absolute change of logarithmic likelihood converge after about 200 iterations (Fig. 3.2). The convergence time for both station based weights and global weights is comparable, although global based calculations required a few more iterations. The BMA_S and BMA_G weights for each model member in the ensemble at iteration convergence are shown in Fig. 3.3. The model configurations (x axis) in Fig. 3.3 are described in Table 3.4. Fig. 3.3 is a stacked bar plot and it shows that several model implementations give their best predictions with higher weights for each validation station. In some cases, a specific model implementation may obtain vary high weight (above 0.5) compared with other model implementations at one specific validation station. This indicates that this "specific" model

implementation may mostly fit for the real condition with respect to channel geometry, roughness and input forcing along the rivers. Comparatively, global data assigns relatively evenly distributed weights to all the model configurations and no one model could acquire a very high weight. BMA_S weights vary across the basin and one set of weights obtained from one station may not be suitable and used to find out the BMA predictions for another station. Whereas BMA_G weights could be applied to predictions for the entire basin consistently including the ungauged locations. Fig. 3.3 also shows that the first 40 configurations, which use lower channel width, have lower weights compared to the next 40. This means that selection of proper channel width is crucial and a smaller channel width will produce less accurate results in LISFLOOD-FP. LISFLOOD-FP relates the channel width with channel depth, in general, a wider channel also results in a deeper cross section. When 5% threshold low channel width value is used, the channel cross sectional area is significantly underestimated, thus affecting the flood inundation depth and extent.

Table 3. 1 Model configurations and their corresponding implementation of uncertainty sources

Model configuration	Channel shape	Channel width	Channel roughness	Input flow
1	Triangle	5%	0.01	0.9USGS
2	Triangle	5%	0.01	USGS
3	Triangle	5%	0.01	1.1USGS
4	Triangle	5%	0.03	0.9USGS
5	Triangle	5%	0.03	USGS
6	Triangle	5%	0.03	1.1USGS
7	Triangle	5%	0.05	0.9USGS
8	Triangle	5%	0.05	USGS
9	Triangle	5%	0.05	1.1USGS
10	Rectangle	5%	0.01	0.9USGS
11	Rectangle	5%	0.01	USGS
12	Rectangle	5%	0.01	1.1USGS
13	Rectangle	5%	0.03	0.9USGS
14	Rectangle	5%	0.03	USGS
15	Rectangle	5%	0.03	1.1USGS
16	Rectangle	5%	0.05	0.9USGS
17	Rectangle	5%	0.05	USGS
18	Rectangle	5%	0.05	1.1USGS
19	Parabola	5%	0.01	0.9USGS
20	Parabola	5%	0.01	USGS
21	Parabola	5%	0.01	1.1USGS
22	Parabola	5%	0.03	0.9USGS
23	Parabola	5%	0.03	USGS
24	Parabola	5%	0.03	1.1USGS
25	Parabola	5%	0.05	0.9USGS
26	Parabola	5%	0.05	USGS
27	Parabola	5%	0.05	1.1USGS
28	Triangle	Average	0.01	0.9USGS
29	Triangle	Average	0.01	USGS
30	Triangle	Average	0.01	1.1USGS
31	Triangle	Average	0.03	0.9USGS
32	Triangle	Average	0.03	USGS
33	Triangle	Average	0.03	1.1USGS
34	Triangle	Average	0.05	0.9USGS
35	Triangle	Average	0.05	USGS
36	Triangle	Average	0.05	1.1USGS
37	Rectangle	Average	0.01	0.9USGS
38	Rectangle	Average	0.01	USGS
39	Rectangle	Average	0.01	1.1USGS
40	Rectangle	Average	0.03	0.9USGS
41	Rectangle	Average	0.03	USGS

Table 3.4 continued

Model configuration	Channel shape	Channel width	Channel roughness	Input flow
42	Rectangle	Average	0.03	1.1USGS
43	Rectangle	Average	0.05	0.9USGS
44	Rectangle	Average	0.05	USGS
45	Rectangle	Average	0.05	1.1USGS
46	Parabola	Average	0.01	0.9USGS
47	Parabola	Average	0.01	USGS
48	Parabola	Average	0.01	1.1USGS
49	Parabola	Average	0.03	0.9USGS
50	Parabola	Average	0.03	USGS
51	Parabola	Average	0.03	1.1USGS
52	Parabola	Average	0.05	0.9USGS
53	Parabola	Average	0.05	USGS
54	Parabola	Average	0.05	1.1USGS
55	Triangle	95%	0.01	0.9USGS
56	Triangle	95%	0.01	USGS
57	Triangle	95%	0.01	1.1USGS
58	Triangle	95%	0.03	0.9USGS
59	Triangle	95%	0.03	USGS
60	Triangle	95%	0.03	1.1USGS
61	Triangle	95%	0.05	0.9USGS
62	Triangle	95%	0.05	USGS
63	Triangle	95%	0.05	1.1USGS
64	Rectangle	95%	0.01	0.9USGS
65	Rectangle	95%	0.01	USGS
66	Rectangle	95%	0.01	1.1USGS
67	Rectangle	95%	0.03	0.9USGS
68	Rectangle	95%	0.03	USGS
69	Rectangle	95%	0.03	1.1USGS
70	Rectangle	95%	0.05	0.9USGS
71	Rectangle	95%	0.05	USGS
72	Rectangle	95%	0.05	1.1USGS
73	Parabola	95%	0.01	0.9USGS
74	Parabola	95%	0.01	USGS
75	Parabola	95%	0.01	1.1USGS
76	Parabola	95%	0.03	0.9USGS
77	Parabola	95%	0.03	USGS
78	Parabola	95%	0.03	1.1USGS
79	Parabola	95%	0.05	0.9USGS
80	Parabola	95%	0.05	USGS
81	Parabola	95%	0.05	1.1USGS

*5%, average, 95% represents the threshold values of sampled channel widths

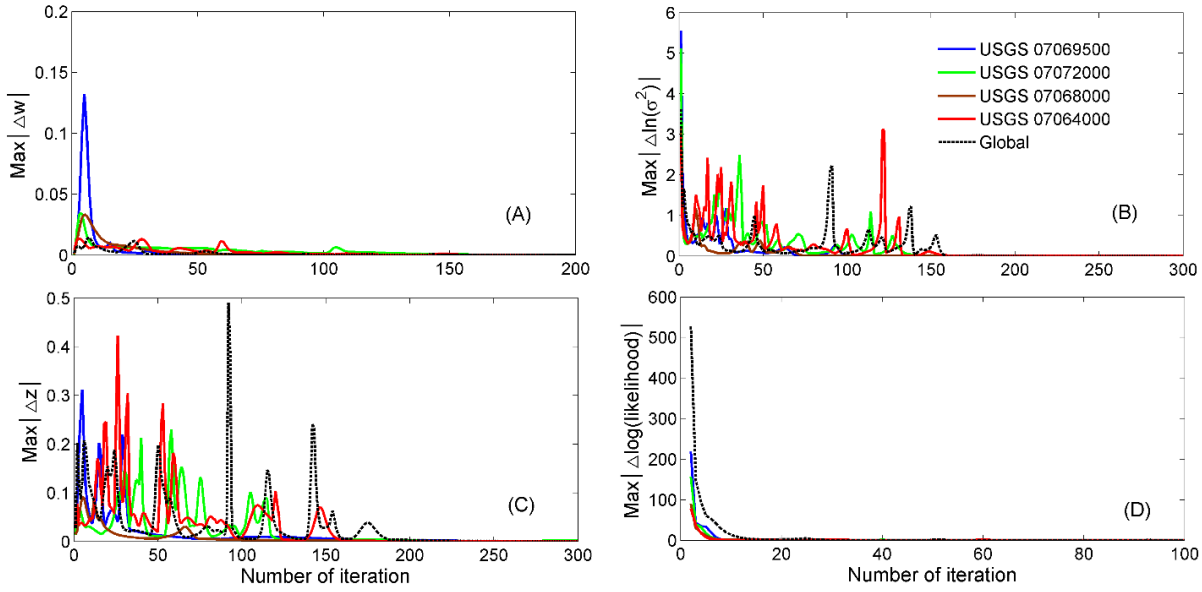

Figure 3.2 Convergence of four indicators after EM iterations at four USGS validation stations during training period

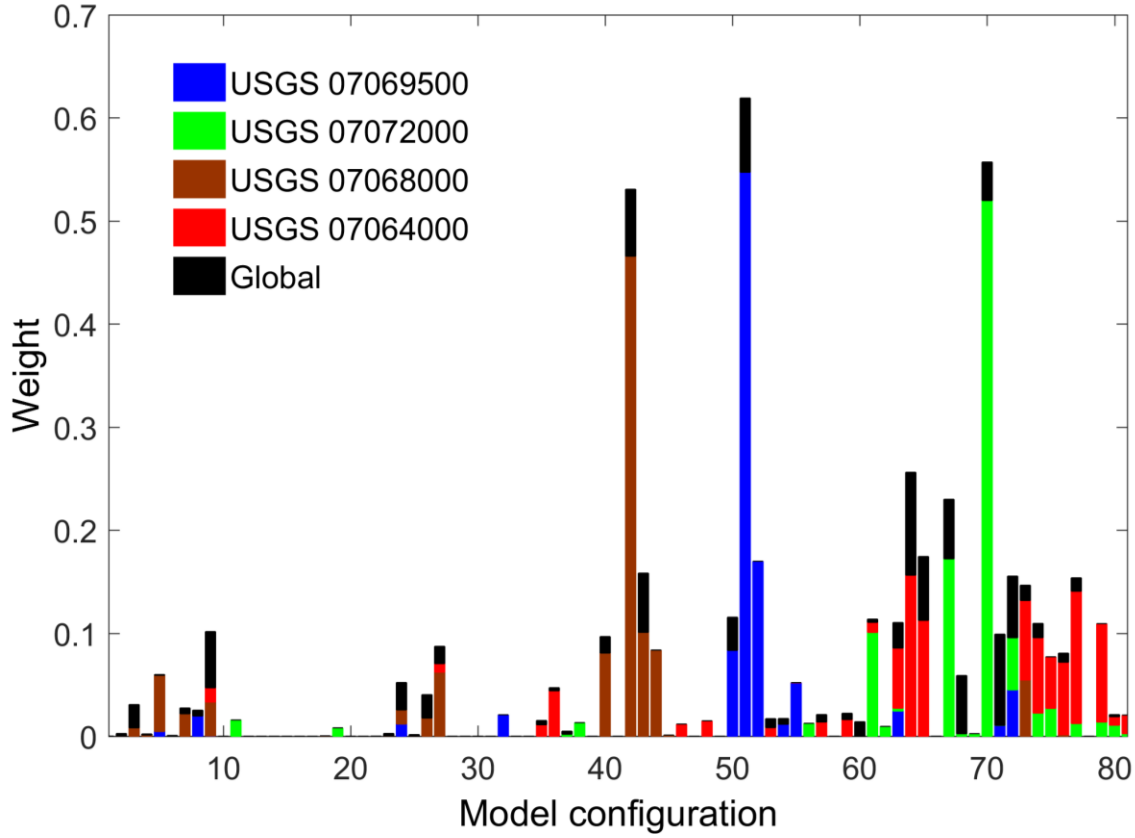

Figure 3.3 BMA_S and BMA_G Weights for each model configuration at convergence based on training data

3.5.2 Validation of BMA prediction performance

Overall, the water stage range in Figure 3.4 provided by the model ensemble predictions at four stations includes the observed water stage, ensemble mean prediction as well as BMA predictions. The BMA_S performs the best followed by BMA_G and ensemble mean predictions in the validation period when the same weights derived from the training period are used. In Figure 3.4B&C, the simulations from the model ensemble overestimate water levels when compared with observation data at low flow condition. This indicates that assuming a high flow input forcing might cause prediction bias before and after the flood peaks. Additionally, Figure 3.4 also points out that the quality of DEM from NED that is derived from LIDAR could satisfy the requirement of this study and it won't cause huge bias for the predicted water stages. The difference in the performance of these models is more visible in the plot of water stage RMSE in Figure 3.5. Additionally, the difference of RMSEs between these predictions varies from station to station. Fig. 3.5A and Fig. 3.5C show that ensemble mean predictions in these two stations are closer to BMA predictions. This is because the observation data is generally in the middle of the prediction ensemble, which can be found in Fig. 3.4. Besides BMA, for these two validation stations, ensemble mean can also produce desirable output. However, in other cases as shown in Fig. 3.5B and Fig. 3.5D, observed water stage data is closer to the upper or lower bound of the prediction ensemble. For these two cases, BMA predictions are still among those best predictions whereas ensemble mean predictions are much worse. Comparing with BMA_S, BMA_G performance is slightly worse for each specific station and this is because BMA_G weights consider a wider spatial range of training data from different stations. When predictions for locations other than those four validation stations needed, BMA_G is trustworthy. BMA_G is especially useful for those data sparse watersheds where there is not enough training data for all the locations interested. As shown in Fig. 3.5, BMA_S may not outperform all the models in the

ensemble. However, without knowing which model implementation will yield better outcome beforehand, it is good to use BMA_S and BMA_G predictions instead of trusting only one specific model implementation or the ensemble mean prediction. This is because one model may perform well for one flood event in the training period but may not perform equally well for another. In order to test the impact of temporal scale, a shortened period (40days around the flood peak) is used for training. However, BMA prediction still has its advantage, and performs better compared with most of the model member predictions. Therefore, to some extent, the temporal scale may affect the results, but not to change the conclusions. Moreover, the weight obtained from one station is also used as global weights to test how it works at other locations. One test done by applying the BMA weight trained from USGS 07072000 to USGS 07069500 (from validation station on Eleven point River to validation station on Spring river) indicates that the BMA performance becomes worse/more biased compared with the observations. It overestimates the water level about 0.5m over the validation period on average whereas it behaves very well when the weights obtained from USGS 07069500 itself is used (Figure 3.4A). This also points out that the spatial dependency for this study area is not very strong.

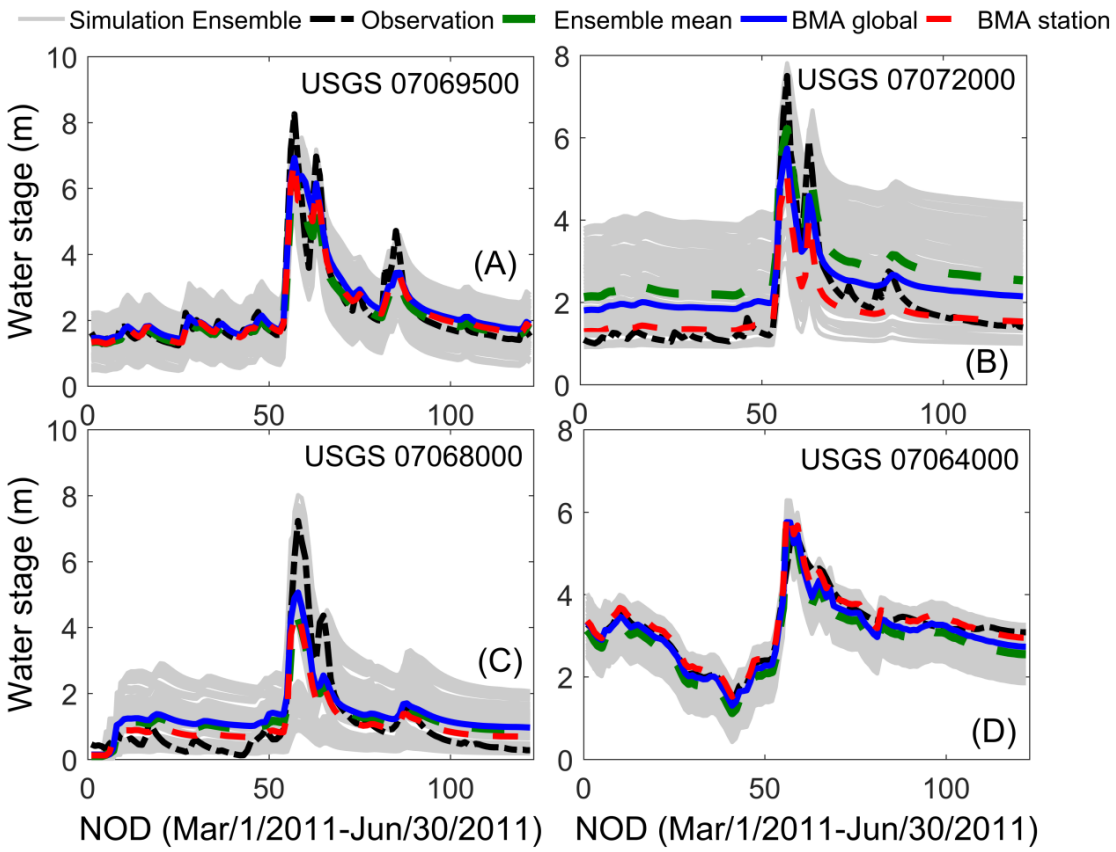

Figure 3.4 Predicted water stages from model ensemble, BMA_S and BMA_G, ensemble mean as well as observed water stage at four USGS stations in the validation period

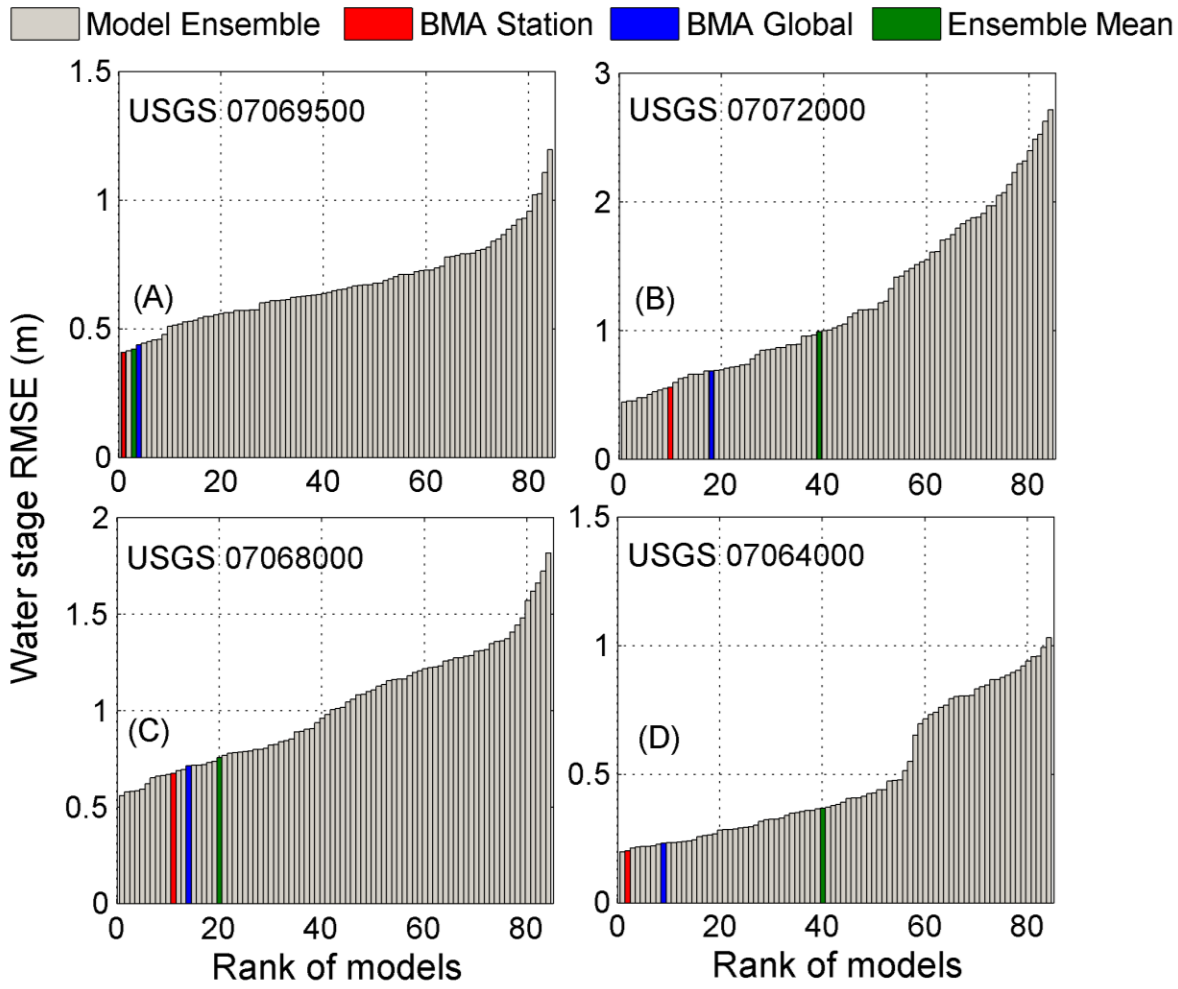

Figure 3.5 Rank of daily average RMSEs of water stage from model ensemble, BMA_S, BMA_G, ensemble mean as well as observed water stage at four USGS stations during the validation period

3.5.3 Comparison of probabilistic flood maps based on BMA_G and ensemble mean

The probabilistic flood map obtained from the BMA and Ensemble mean is shown in Fig. 3.6 for an area where satellite image derived observed inundation extent exists. Although both BMA and ensemble mean derived flood inundation cover the same area, the probability of flooding predicted by these two methods is different, especially for the areas with high probability of flooding. By assigning BMA_G weights to all model configurations, the probability of flooding is higher for the upper right region in Fig. 3.6B compared with the ensemble mean probabilistic

flood map. Since the upper right part is observed to be flooded (black outline), the higher probability as shown by the BMA_G probabilistic flood map is more reliable compared to the ensemble mean prediction. Fig. 3.7 evaluates the performances of these two types of probabilistic flood maps with observed flood extent by calculating *F* and *C* indices. Specifically, the predicted flood extents from raster grids in which flooding probability are greater than certain thresholds are evaluated with observed inundation extent respectively. It indicates that the BMA_G probabilistic flood map has better performance than ensemble mean based approach when higher thresholds are selected (chance of flood is greater than 60%) for evaluation. As the threshold level decreases below 60%, there is no significant difference between these two approaches. This is caused by the difference in model behaviors in the ensemble. Considering that the observed flood region is captured by some of the models in the ensemble while missed by others, BMA picks better performing models (capturing the flood) to assign higher weights. Accordingly, this produces higher chance of flood in this region. On the contrary, the ensemble mean method assigns equal weights to all the models, which reduces the probability of flooding in the region.

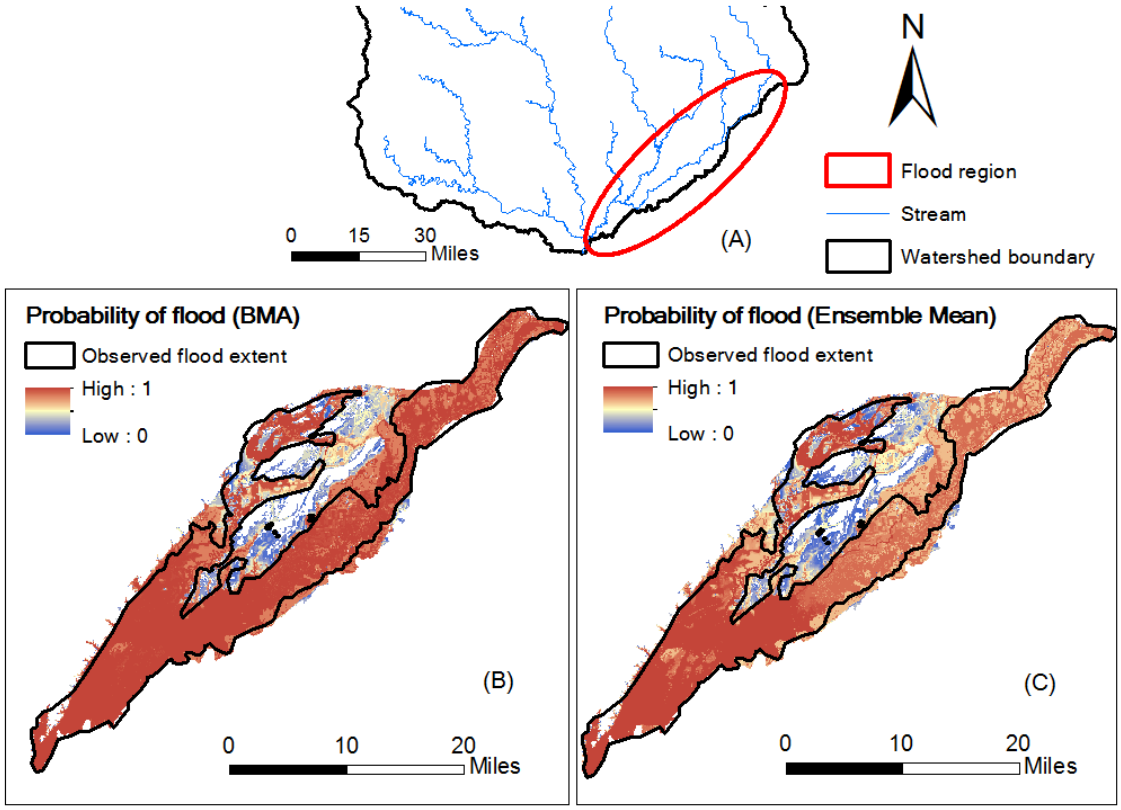

Figure 3.6 Probabilistic flood maps based on BMA_G and ensemble mean

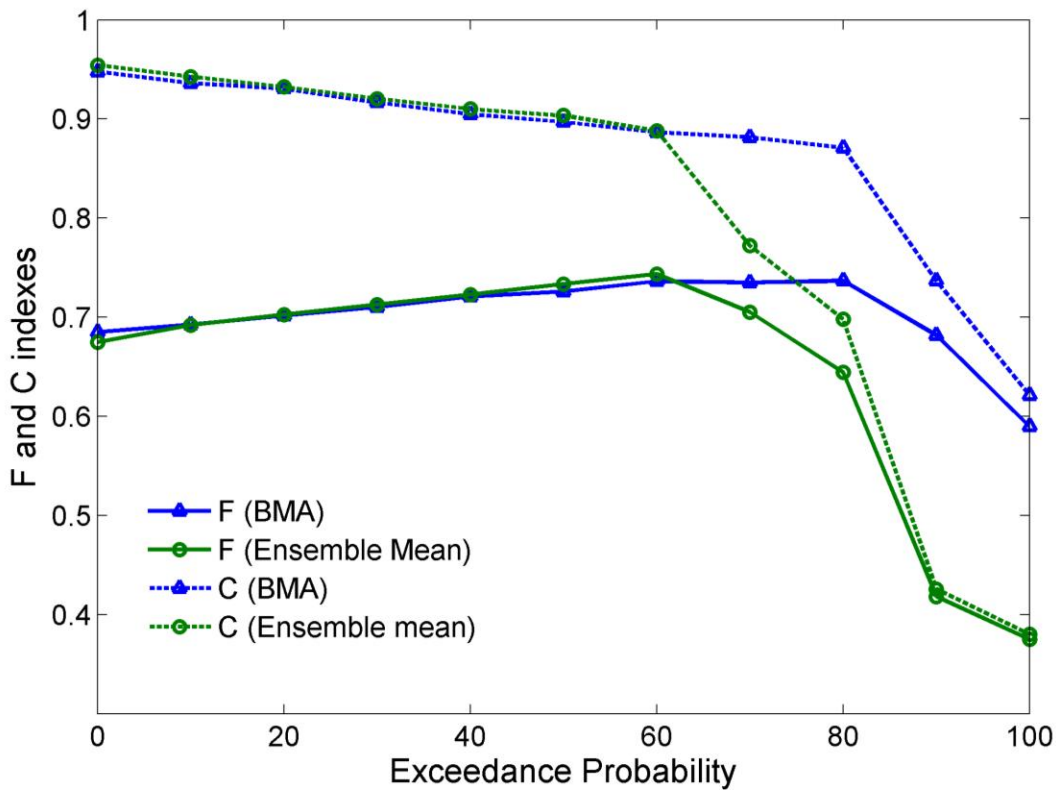

Figure 3.7 Evaluation of the performance of two types of probabilistic flood maps

3.5.4 Evaluation of uncertainty associated with BMA deterministic prediction

Fig. 3.8 shows the 95% confidence intervals of BMA deterministic water stage predictions. The bandwidth (space between upper interval and lower interval) reflects uncertainty of the deterministic prediction. The difference of bandwidths for BMA_S predictions, among stations, is caused by the different variances σ_k^2 and weights w_k associated with each model member varying significantly across stations. Fig. 3.8A and 3.8B show that the BMA prediction uncertainty of these two stations are less than the other stations. One reason is that these two validation stations (3.8A and 3.8B) are closer to the flow station upstream (Table 3.5). Additionally, the change of width along these two reaches (3.8A and 3.8B) is not as significant as the other reaches (3.8C and 3.8D, Table 3.5). Thus, the model configurations have relatively less variance on the BMA simulation output for these two stations (3.8A and 3.8B).

Comparatively, BMA_G prediction bands have the same width across the four stations since the same variances and weights are applied to the model ensemble. Moreover, Figs. 3.8C and 3.8D indicate that though BMA_S deterministic water stage prediction has better performance than BMA_G prediction, the uncertainty bounds associated with BMA_S prediction could be higher.

Table 3.4 Maximum distance between input (upstream) and validation (downstream) stations along each river

River	Distance (km)
Spring River	33
Eleven Point River	36
Current River	118
Black River	128

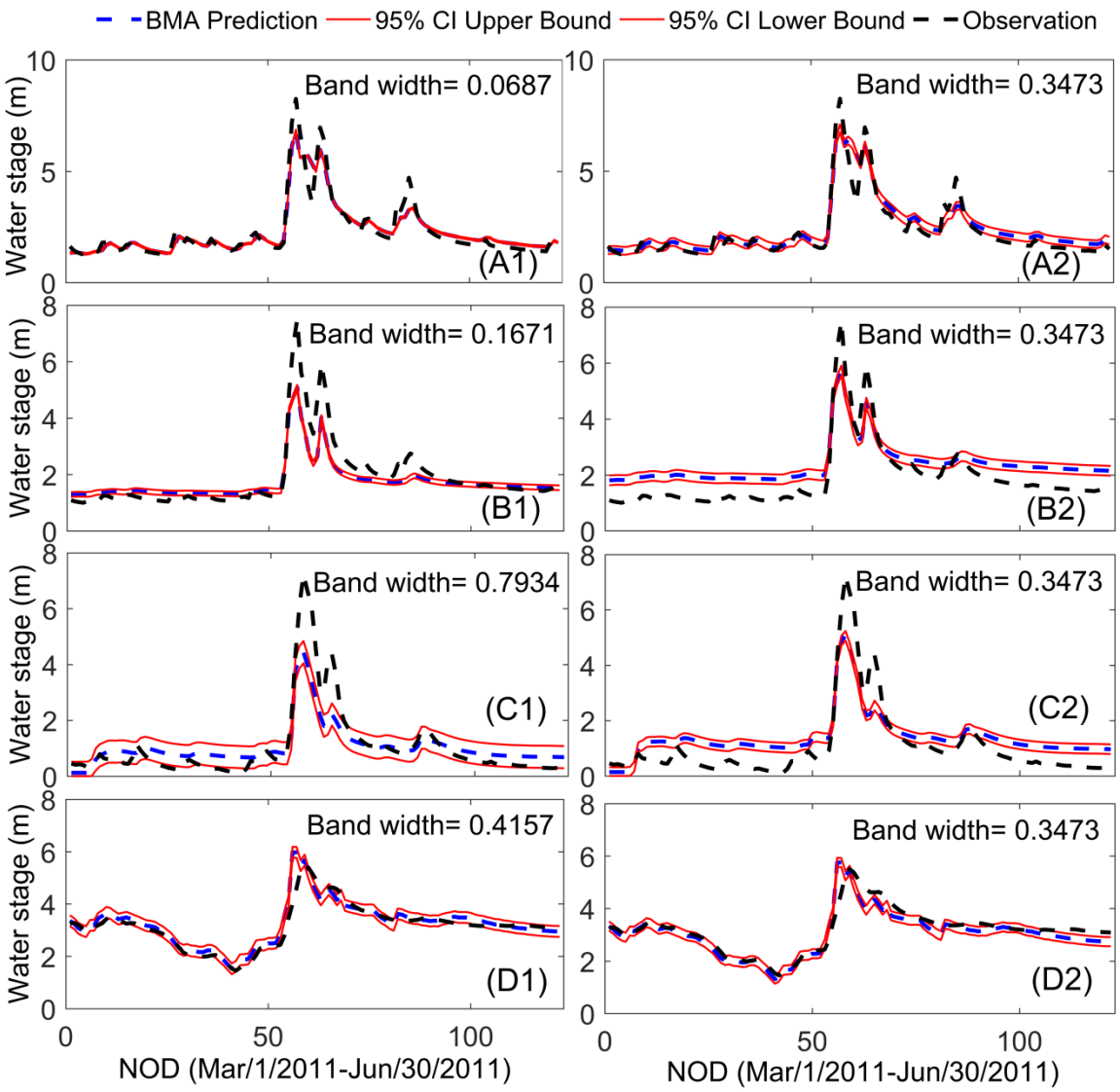

Figure 3.8 Uncertainty associated with BMA deterministic predictions of water stage under 95% confidence interval. A1, B1, C1 and D1 are station based prediction ranges while A2, B2, C2 and D2 are global based prediction ranges at USGS 07069500, USGS 07072000, USGS 07068000 and USGS 07064000 respectively

3.6 Conclusion

Mitigating and managing flood risk depends on reliable prediction of flood inundation and water stage. Considering a wide range of epistemic uncertainty sources associated with the flood inundation modeling, it is important to understand the combined effects of these uncertainties on the prediction of flood inundation extents. Instead of trusting prediction from only one specific

model implementation, the ensemble prediction trajectories are generated in this study by considering the uncertainty from channel width, channel cross-section shape, roughness and bias in flow input data. The Bayesian model averaging approach is applied as the multi-model combing method to provide a deterministic prediction as well as its associated prediction uncertainty bounds. Two types of BMA (BMA_S and BMA_G) predicted deterministic water stages are evaluated with observed water stage data at four USGS validation stations and compared with ensemble mean prediction. In addition, BMA_G based probabilistic flood map is created and compared with the ensemble mean based probabilistic inundation extent. Uncertainty associated with the BMA deterministic prediction is also investigated. The following conclusions are drawn from this study.

1. Although BMA deterministic prediction may not always outperform all the model members in the ensemble, this approach is able to provide a relatively robust water stage prediction at each validation station. Thus, the BMA approach can be used as the multi-model combining method for the ensemble flood inundation modeling.
2. Typically, BMA_S deterministic prediction behaves better than most of the member predictions in the ensemble. Additionally, BMA_S has better prediction than BMA_G, which is better or at least similar to ensemble mean prediction.
3. In spite of BMA_G prediction is not as good as BMA_S by considering a wider spatial range of training data from different stations, it could provide relatively reliable and robust prediction for an ungauged location in the watershed. Therefore, BMA_G is especially useful for those data sparse watersheds where the training data of the interested locations is insufficient. Comparatively, applying the weights obtained from one station in the training period as global weights to another validation station usually yields poorer result than

adopting the weights derived from BMA_G. Additionally, in general, models implementing 5% low channel width perform poorer due to the underestimation of channel cross sectional areas.

4. In terms of probabilistic flood inundation extents, BMA_G has better performance than ensemble mean prediction for high-chance flood regions at Black River watershed. On the other hand, there is no significant difference between two types of probabilistic flood maps for low-chance flood regions.
5. The uncertainty bands of prediction based on 95% confidence intervals vary from location to location for BMA_S prediction. This reflects the change of weights and variances associated with the normal distribution for models at different locations. At USGS 07069500 and USGS 07072000, the bandwidths of confidence interval are relatively narrower. This is because these stations are closer to the upstream flow input stations and it results in less variance from different model implementations. Whereas the bandwidths are the same for BMA_G prediction at all validation stations since the weights and variances associated with the normal distribution for models remain unchanged across the basin. Despite BMA_S deterministic prediction typically performing better than BMA_G, the uncertainty associated with BMA_S might be higher.

The overall goal of this study was to use the BMA method for quantifying the combined uncertainty from different sources in flood inundation modeling to create a reliable prediction from model ensembles. Additionally, this study sets the stage for the next phase where the hierarchical BMA can be used to quantify the relative contribution from each uncertainty source. While the results show that the BMA method is appropriate for providing robust predictions from an ensemble of model results incorporating multiple uncertainty sources, the methodology used in this study has some limitations that merit attention. First, this study assumed that the flow

depths are independent in time and space. Although, the data are strongly correlated in time, the correlation in the prediction errors is relatively smaller. While this may not change the findings of this study, it would be good to investigate the relative performance of the BMA methodology when the temporal dependence is taken into consideration.

Limitations related to the assumption of normally distributed predictive PDFs of transformed data and the adoption of local optimization technique, e.g., Expectation Maximization, have also been mentioned in other studies (Madadgar and Moradkhani, 2014; Parrish et al., 2012; Vrugt and Robinson, 2007). Although Expectation Maximization algorithm is suggested to find out the solution for the log likelihood function, it is a local optimization approach. The normal distribution is only valid for limited variables such as sea level pressure and temperature, but other variables, e.g., streamflow used in this study, needs to be transformed to the normal space before BMA is applied(Raftery et al., 2005). Although the Box-Cox transformation can adjust the skewness of the data, the PDF still remains bounded. This is especially true for this study since as the same transformation power, λ, is used for all the simulations in the ensemble and observation data for each validation station. This may have slightly affect the PDF of the transformed data. A different transformation power could be used when individual simulation is considered for analysis. To take this into account in future studies, sequential data assimilation could be incorporated into BMA to release the Gaussian assumption of the likelihood function (Parrish et al., 2012). Also, a group of copula functions could be integrated with BMA to estimate the posterior distribution of model forecast, which removes the normal distribution assumption of prediction data as well as the transformation process (Madadgar and Moradkhani, 2014).

CHAPTER 4. SEPARATION AND PRIORITIZATION OF UNCERTAINTY SOURCES IN A RASTER BASED FLOOD INUNDATION MODEL USING HIERARCHICAL BAYESIAN MODEL AVERAGING

4.1 Abstract

Uncertainty in hydrodynamic model comes from input data, model structure and parameters. In order to provide the robust model predictions, Bayesian model averaging (BMA) approach could be used as a multi-model combining method to account for the compound effects from various uncertainty sources. However, BMA cannot provide a clear picture of the impact from individual uncertainty sources including those from model structure, model parameter and model input, which are crucial for people to understand. Hierarchical Bayesian model averaging (HBMA) is a recently developed approach to study the relative impact of different uncertainty sources, which explicitly considers various sources of uncertainty in the hierarchical structure (BMA tree) for analyzing uncertainty propagation. In this study, HBMA is tested over Black River watershed in Missouri and Arkansas based on water stage predictions from 243 LISFLOOD-FP model configurations that integrate five sources of uncertainty including channel shape, channel width, channel roughness, topography and flow input. The results indicate that without considering the model weight, channel width and topographical data resolution have the largest impact on the hydrodynamic model predictions followed by flow forcing, which has relatively greater influence than channel cross-sectional shape and model parameter. However, when model weights are taken into account, model input (topography and flow forcing) and model parameter have larger impact on prediction variance than model structure (channel width and cross-sectional shape). Moreover, as level moves up from base level towards the hierarch level in the

BMA tree, accuracy of the deterministic mean prediction also increases in general whereas the 95% confidence interval associated with the deterministic mean prediction might become larger.

4.2 Introduction

Floods are the most frequently occurring natural disasters in the United States that cause huge amount of property loss and damage (Deniz et al., 2017; Wing et al., 2018). Appropriate flood risk management decision is underpinned by the reliable water stage and inundation extent predicted by a hydrodynamic model (Alfieri et al., 2014; Bates et al., 2010; Bermúdez et al., 2017; Dimitriadis et al., 2016; Hunter et al., 2008; Merwade et al., 2008; Neal et al., 2012b; Pappenberger et al., 2005a). There is growing interest to simulate floods at multiple spatial scales ranging from a single reach to continental stream networks to meet the needs of local, state and federal agencies (Alho and Aaltonen, 2008; de Paiva et al., 2013; Liu et al., 2018; Schumann et al., 2013; Teng et al., 2017a; Wilson et al., 2007; Wood et al., 2016b). However, regardless of whether a one- dimensional (1D) or a two- dimensional (2D) hydrodynamic model is selected for simulation, there always exist a lot of uncertainty that affect the quality of model predictions (Aronica et al., 1998; Duong and Gourbesville, 2016; Merwade et al., 2008). A 1D model such as HEC-RAS 1D (Horritt and Bates, 2002b; Pappenberger et al., 2005b) extracts information using discrete cross sections whereas a 2D model such as HEC-RAS 2D, LISFLOOD-FP, Mike Flood and Interconnected Channel and Pond Routing Model (ICPR) model applies more complex grids or meshes to simulate the flow transportation in the channel and floodplain (Apel et al., 2009; Bates et al., 2010; Neal et al., 2012a; Patro et al., 2009; USACE, 2015). The different dimensionalities belong to the model structure uncertainty. Besides this, model structure uncertainty also includes the form of governing equations and their assumptions, and the river geometry representations in the model. The most common calibration parameter of the

hydrodynamic model is channel roughness (Aronica et al., 1998; Liu et al., 2018). The calibrated channel roughness value for one event with a certain flow magnitude may not be suitable for another event with a different flow magnitude, thus adding uncertainty to the model prediction (Liu et al., 2018; Wood et al., 2016b). Finally there are two major inputs for the flood inundation models including topography data and input forcing. Depending on accuracy of the two major inputs used, the results from a hydrodynamic simulation is significantly affected (Saksena and Merwade, 2015).

According to the concept of "equifinality", different model structures, parameters and input/boundary conditions may make the predictions achieve the similar level of fit to the observations. To interpret this in another way: a one specific model implementation that works well for one flood event might not perform well for another since it may have reached equifinality as the best model implementation for the first flood event. Therefore, the prediction based on single model approach is subjected to significant bias if the model is not properly implementated. Comparatively, multi-model combining approach is a more robust way to extract information from a group of existing model implementations to provide reliable predictions. Among the current existing multi-model combining approaches such as simple model average, weighted model average and multi-model super ensemble (Ajami et al., 2006; Chowdhury and Sharma, 2009; Liu et al., 2018; Najafi and Moradkhani, 2015b; Wood et al., 2016b), Bayesian model averaging is one of the most commonly and widely used approach to generate a reliable deterministic model prediction and its prediction distribution from model ensemble based on model performance in the past (Madadgar and Moradkhani, 2014; Morales et al., 2006; Notaro et al., 2014; Raftery et al., 2005; Vrugt and Robinson, 2007; Yan and Moradkhani, 2014). Although the BMA method aggregates all plausible models to obtain one optimal prediction, it only

evaluates the combination effects from all uncertainty sources. The role of individual uncertainty source is still not well understood (Yen, 2012). The uncertainty source that could significantly change the prediction outcome or prediction variance should be identified. Therefore, separation and evaluation of relative importance of model structure, model parameter and model input forcing on the final prediction and prediction variance needs to be investigated.

Hamill (2001) provides a perturbation approach to study the impact of individual factors on model predictions. In the perturbation approach, one individual factor is added or modified each time to evaluate the change in the model prediction. Each model implementation is treated equally in the perturbation approach. However, in a multi-model combining method such as BMA, models are weighted and poor predictions from some model implementations in the ensemble are eliminated. Therefore, the impact of each uncertainty source when weights are considered is different from that found using model perturbation approach since instead of model prediction, model prediction variance is evaluated (Chitsazan and Tsai, 2015). Hierarchical Bayesian model averaging (HBMA) approach considers both situations by separating and locating each uncertainty source on one level of a BMA tree structure (Chitsazan and Tsai, 2014; Chitsazan and Tsai, 2015; Tsai and Elshall, 2013). On one hand, the original predictions from different model implementations are the root of the BMA tree, which are located at a hidden level (base+1 level, discussed later) and no weight has been counted. Model perturbation method can be applied to evaluate the impact of each uncertainty source on model prediction based on these original predictions. On the other hand, as level moves up along the hierarchical structure, HBMA also provides an insight into the impact of different uncertainty sources on prediction variance and how they propagate over the BMA tree when weights are counted.

Considering the need to understand the influence of individual uncertainty source in flood inundation modeling, the objectives of this study are to: (1) prioritize the impact of different uncertainty sources on model prediction and prediction variance using the HBMA approach; (2) investigate how the model performance changes at different hierarchical levels in the BMA tree; (3) quantify the uncertainty associated with the mean predictions at each hierarchical level and its propagation along the BMA tree.

4.3 Study area and Data

The 20,000 km^2 Black River watershed located across Arkansas and Missouri states in the U.S. is selected for this study. Historical records from the United States Geological Survey (USGS) stations indicate that this region has experienced numerous flood events in the past. Additionally, Black River watershed has four major rivers including Black river, Current River, Eleven Point River and Spring River, which drain towards the watershed outlet in Arkansas. These four major rivers provide distinct topographical and geomorphic settings as well as varying reach lengths and sinuosity, thus making the Black River watershed a good test bed for this study. The daily streamflow (input to the hydraulic model) and stage data (for validation) are obtained from the USGS gauging stations (Figure 4.1 and Table 4.1). The topography and land use dataset are obtained in the form of National Elevation Dataset (NED) 90m DEM (http://ned.usgs.gov) and NLCD 2011 land use data (http://www.mrlc.gov/nlcd2011.php), respectively. Although the prediction accuracy improves by using higher resolution DEM, 90-m resolution is used in this study to strike a balance between predication accuracy and computational demand for a large number of simulations needed for uncertainty analysis. Additionally, many large scale models use 90m or coarser DEM for flood modeling (Neal et al., 2012a; Schumann et al., 2013). Inundation extents for selected storms, which will be used for the validation of results, are

derived by classifying Landsat images from the USGS earth explorer (http://earthexplorer.usgs.gov/). The Landsat images are classified into water and non-water area with a supervised classification technique using ArcGIS classification toolbar.

Table 4.1 USGS gauge IDs for the streamflow input and validation stations

Streamflow Input Station	USGS Station ID
S1	07069305
S2	07071500
S3	07065495
S4	07064533
S5	07068510
S6	07061900
S7	07061500
Validation Station	
A	07069500
B	07072000
C	07068000
D	07064000

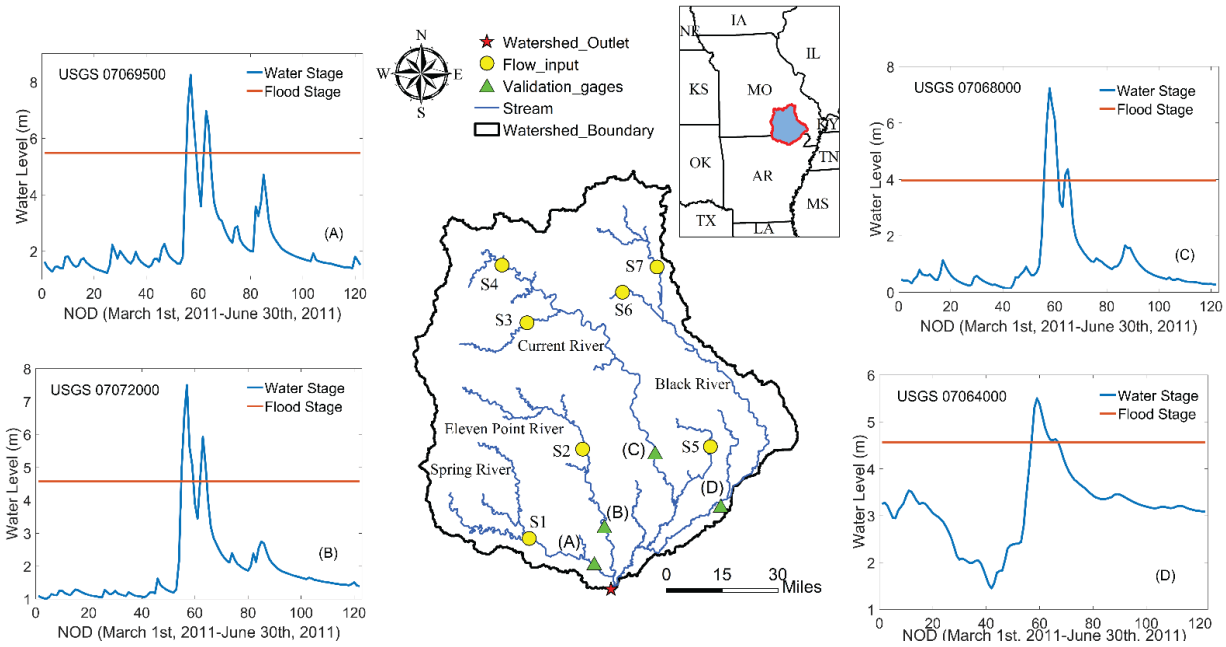

Figure 4.1 Layout map of Black River watershed and input/validation USGS stations

4.4 Methodology

4.4.1 Hydraulic modeling

Considering the amount of efforts in model set-up, computational demand and the results from past studies, LISFLOOD-FP is selected in this study. The latest subgrid version of LIFLOOD-FP is easy to implement for large scale flood simulations and it usually yields accurate predictions (Bates and De Roo, 2000; Neal et al., 2012a). LISFLOOD-FP solves continuity and momentum equations to obtain the information of streamflow and water depth in each raster cell (Bates et al., 2010). The input data for LISFLOOD-FP model includes the DEM, flow forcing time series, stream centerline, and channel geometry information including average river width, slope, surface roughness and channel cross sectional shape (Merwade et al., 2018). The floodplain roughness values used in this study are obtained by relating each land use category in the NLCD land use dataset to Manning's n value (Kalyanapu et al., 2010). A six-month daily time-step simulation, from 01/2011 to 06/2011, is used for this research. This period is selected based on historical flood events and the availability of input data. A total number of 243 LISFLOOD-FP model configurations, described in the next section, are created and simulated using High Performance Computing (HPC) resources at Purdue University. After trial runs, the first two months are used as warm-up periods to avoid unrealistic predictions in the BMA analysis.

4.4.2 Uncertainty sources and model implementations

Uncertainty in LISFLOOD-FP model comes from a variety of sources including model structure, model parameter, input forcing and boundary conditions, among others (Demeritt et al., 2007; Merwade et al., 2008; Merz et al., 2004; Merz and Thieken, 2009; Todini, 2004). Five common uncertainty sources including the cross-sectional shape of the channel, channel width, channel roughness, DEM resolution, and streamflow input are considered in this study. Although there

are other uncertainty sources such as the configuration of hydraulic structures including bridges and culverts, they are excluded due to the lack of detailed data at the stream network scale. The uncertainty in streamflow estimates at the USGS gaging stations can be as high as 10% for stable natural channels (Harmel et al., 2006b; Slade, 2004). Thus, three streamflow estimates including 0.9Q, Q and 1.1Q are used. Similarly, uncertainty in channel width is incorporated by picking 5%, average and 95% from a distribution of channel widths (Table 2). The uncertainty in channel roughness, Manning's n value, is selected from 0.01-0.05. This range is used based on previous studies to reflect the roughness conditions of most natural channels characterized from clean, straight channel to winding, grassy and graveled channel (Liu et al., 2018). Three different DEM resolutions including 60m, 90m and 120m are used for topography data, and these resolutions are selected to match the computational demand for the large number of simulations in the Black River watershed. Finally, the uncertainty in cross-sectional shape is incorporated by assuming three shapes, namely triangular, parabola and rectangle.

Considering the aforementioned five uncertainty sources and three values for each source (Table 3), 243 (3×3×3×3×3) LISFLOOD-FP model configurations are created by using different combinations of these uncertainty sources (Table A1). To distinguish the model parameter uncertainty (since change of model parameter is not considered as a new model), the 243 LISFLOOD-FP model configurations are further reconstructed to 81 base models for the base level of BMA tree by only considering the change in model structure and model input. Each of these 81 base model has three within model configurations located at a hidden "base+1" level by considering the uncertainty in channel roughness parameter (Chitsazan and Tsai, 2015).

Table 4.2 Widths of 5% (low), 50% (average), and 95% (high) thresholds for each river

Width (m)	Low (5%)	Average (50%)	High (95%)
River 1	14	35	71
River 2	28	51	96
River 3	34	87	154
River 4	74	181	355

Table 4.3 Five sources of uncertainty and their varied types

Uncertainty source	Channel shape	Channel width	Input flow	Topography	Channel roughness
Type 1	Rectangle	95% high value	1.1 USGS data	60m	0.05
Type 2	Parabola	Average	USGS data	90m	0.03
Type 3	Triangle	5 % low value	0.9 USGS data	120m	0.01

4.4.3 Hierarchical Bayesian Model Averaging (HBMA)

4.4.3.1 Bayesian model averaging

Bayesian model averaging (BMA) is a statistical approach to combine estimations from individual models and produce reliable prediction (Madadgar and Moradkhani, 2014; Raftery et al., 2005). Using the BMA, the mean and variance of predicted water stage using the 243 model implementations can be expressed as Equation 4.1 and 4.2.

$$E(d|M_1, \dots, M_k) = \sum_k E\,(d|M_k) Pr(M_k|D) \tag{4.1}$$

$$Var(d|M_k) = \sum_k Var(d|M_k)\, Pr(M_k|D)$$

$$+ \sum_k [E(d|M_k) - E(d|M_1, \dots, M_k)]^2\, Pr(M_k|D) \tag{4.2}$$

In Equation 4.1, $E(d|M_1, \dots, M_k)$ is the mean of the predicted water stage d from the kth model implementations. $E(d|M_k)$ is the prediction from the kth model. $Pr(M_k|D)$ is the posterior model probability of model M_k and D is the observed water stage data. In Equation 4.2, $Var(d|M_k)$ is the variance of predicted water stage using model M_k. The first term on the right side of Equation 4.2 is the within-model variance of the predicted water stage that averages model prediction variances. The second term on the right side of Equation 4.2 is the between-model variance that accounts for the spreading of mean predicted water stage by different models.

For non-Gaussian data, a power transformation is needed to map the variables from their original space to the Gaussian space. The observed and simulated water stage data are normalized using the Box-Cox power transformation (Box and Cox, 1964; Osborne, 2010). This method is used to transform the skewed data into normal distribution using Equation 4.3. When the power λ is zero, the natural logarithm is applied to the data. Different powers are tested and the best power that yields the closest normal distribution of transformed water stage data is found and applied. The transformed water stage data is used to find BMA weights for each model member. In order to estimate the parameter w_k and σ_k, on the basis of training data set, the log likelihood function (Equation 4.4) is maximized.

$$\mathrm{f}^{\lambda} = \begin{cases} \frac{\mathrm{f}^{\lambda}-1}{\lambda}, & \lambda \neq 0 \\ \ln(\mathrm{f}), & \lambda = 0 \end{cases} \tag{4.3}$$

$$\mathrm{l}(\theta) = \sum_{s,t} \log\left(\sum_{k=1}^{K} \mathrm{w_k g_k}\left((\mathrm{d_{st}}|\mathrm{M_{kst}})\right)\right) \tag{4.4}$$

Where λ is the power that is used to transform the data to a normal distribution and the summation in Equation 4.4 is over locations (s) and time (t) in the training data. Equation 4.4 is difficult to solve analytically or numerically (Vrugt and Robinson, 2007; Zhu et al., 2016), and thus an expectation-maximization (EM) algorithm is used to find out the maximum likelihood.

The EM algorithm is iterative and alternates between two steps, the E (or expectation) step, and the M (or maximization) step by using a latent variable *z*. In the E step, z is estimated given the current estimates of the model weight w_k and σ_k (Equation 4.5).

$$z_{kst}^{j} = \frac{w_k^{j-1} g\left(d_{st} \middle| M_{kst}, \sigma_k^{j-1}\right)}{\sum_{i=1}^{K} w_i^{j-1} g\left(d_{st} \middle| M_{ist}, \sigma_i^{j-1}\right)} \tag{4.5}$$

Where the superscript *j* refers to the jth iteration of the EM algorithm and $g\left(d_{st} \middle| M_{kst}, \sigma_k^{j-1}\right)$ is a normal density with mean M_{kst} and standard deviation σ_k^{j-1}. In the Mth step, weight w_k and standard deviation σ_k are calculated with the current estimate of z_{kst} (Equations 4.6-4.7).

$$w_k^{j} = \frac{1}{n} \sum_{s,t} z_{kst}^{j} \tag{4.6}$$

$$\sigma_k^{2(j)} = \frac{\sum_{s,t} z_{kst}^{j} (d_{st} - M_{st})^2}{\sum_{s,t} z_{kst}^{j}} \tag{4.7}$$

Where *n* is the number of observations including distinct values of locations and time. The E step and M step are iterated to convergence when the changes of weight, variance, latent variable *z* and log-likelihood are smaller than a specified small tolerances (10^{-6} in this study). The log-likelihood is guaranteed to increase at each EM iteration, therefore, finally it converges to a local maximum (Wu, 1983).

Although BMA is a robust method to integrate a series of model outputs for uncertainty analysis, it can not provide detailed information about the relative importance of each uncertainty source. Additionally, the contribution of individual source of uncertainty to the total prediction variance is not clear. Therefore, hierarchical BMA method is applied to address these issues.

4.4.3.2 Hierarchical Bayesian model averaging

Hierarchical Bayesian model averaging (HBMA) uses a BMA tree structure, and the schematic of BMA tree structure developed by the five uncertainty sources in this study is shown in Figure

4.2. Each level in the BMA tree represents one uncertainty source, and contains competing model predictions from this source in a discrete form. The 81 models at the base level are called the base models, which are derived by averaging predictions from the hidden level. All models above the base level are BMA models that are developed by averaging the models one level below them. In this study, the letters (C_t, C_p, C_r) are used to represent the triangular, parabolic and rectangular channel shapes respectively. The letters (W_5, W_{ave}, W_{95}) are applied to describe 5% low channel width, average channel width and 95% channel width. The letters (F_{-10}, F_0, F_{10}) are employed to distinguish 0.9 to 1.1 of the USGS flow observation. Similarly, the letters (T_{60}, T_{90}, T_{120}) are named to represent 60m, 90m and 120m DEM data respectively. Based on this description, "C_t_W_5_F_{-10}_T_{60}" represents a model implementing triangular channel shape, 5% low channel width, 0.9 USGS flow and 60m resolution DEM. In the tree structure, a parent model is the model at a vertex of a level with respect to their child models exactly one level below. For instance, "C_t_W_5_F_{-10}" is a parent model at third level and its child models are "C_t_W_5_F_{-10}_T_{60}", "C_t_W_5_F_{-10}_T_{90}", "C_t_W_5_F_{-10}_T_{120}" respectively. BMA is performed to average the child models to obtain their parent models at the current level. The mean and variance of predicted water stage for a BMA model at level *n* are shown in Equation 4.8 and 4.9.

$$E(d|M^n) = \sum_k E\,(d|M_k^{n+1}) Pr(M_k^{n+1}|D, M^n) \tag{4.8}$$

$$Var(d|M^n) = \sum_k Var(d|M_k^{n+1})\, Pr(M_k^{n+1}|D, M^n)$$

$$+ \sum_k [E(d|M_k^{n+1}) - E(d|M^n)]^2\, Pr(M_k^{n+1}|D, M_n) \tag{4.9}$$

Where M^n are the BMA models at level n, M_k^{n+1} are the child models at level n+1 under their parent models, and $Pr(M_k^{n+1}|D, M^n)$ is the conditional model probability of model under their

parent models M^n. The first term at the right hand side of Equation 4.4 is the within-model variance of predicted water stage using model M^n, which is the mean prediction variance using the child models of M^n. The second term represents the between-model variance of predicted water stage that accounts for the spreading of mean predicted water stages by different M_k^{n+1} models. Using Equations 4.8 and 4.9, the HBMA provides insights into model averaging for analysts in order to evaluate all possible averaged predicted water stages and variances at different levels.

4.4.3.3 Conditional model probability and posterior model probability

In order to calculate the conditional model probability (or model weight) $Pr(M_k^{n+1}|D, M^n)$, the posterior model probability at level n+1 needs to be calculated using Equation 4.10.

$$Pr(M^n|D) = \sum_k Pr\,(M_k^{n+1}|D) \tag{4.10}$$

$$Pr(M_k^{n+1}|D, M^n) = \frac{Pr\left(M_k^{n+1}|D\right)}{Pr(M^n|D)} \tag{4.11}$$

Equation 4.10 is a recursive equation. Therefore, the posterior model probability of models at level n can be obtained by adding up the posterior model probabilities recursively from the base level to level n+1 using Equation 4.10. The conditional posterior model probability at level n+1 under a model at level n is computed using Equation 4.11.

4.4.3.4 Prediction mean and variance of hierarch model

The mean of predicted water stage at a given hierarchy level can be obtained by averaging the mean predicted water stage of the models using Equation 4.8 recursively from the base level (level 4) to level 1. Therefore, the mean of the hierarchy model in a four-level BMA tree can be

described using Equation 4.12. Similarly, the variance of predicted water stage at the hierarchy level can be obtained by using Equation 4.13 recursively from the base level to level 1.

$$E(d|D) = E_{M1}E_{M2}E_{M3}E_{M4}[E(d|D, M_4)] \tag{4.12}$$

$$Var(d|D) = E_{M1}E_{M2}E_{M3}E_{M4}[Var(d|D, M_4)] + E_{M1}E_{M2}E_{M3}Var_{M4}[E(d|D, M_4)]$$

$$+E_{M1}E_{M2}Var_{M3}[E(d|D, M_3)] + E_{M1}Var_{M2}[E(d|D, M_2)] + Var_{M1}[E(d|D, M_1)] \tag{4.13}$$

Where $E(d|D)$ is the mean predicted water level. E_{Mn} is the expectation operator to average a quantity given by model M_n. $Var(d|D, M_4)$ is the variance of predicted water stage. Var_{Mn} is the variance operator to average a quantity given by models M_n. The first term on the right side of Equation 4.13 shows the contribution of the within-model variance of the base models to the total prediction variance. In this study, the within-model variance of the base models stems from uncertainty in channel roughness. Given the hierarchy in Figure 4.2, the second, third, fourth, and fifth terms in Equation 4.13 are the contributions of variances of the channel cross sectional shape, channel width, flow accuracy and topographical data resolution respectively.

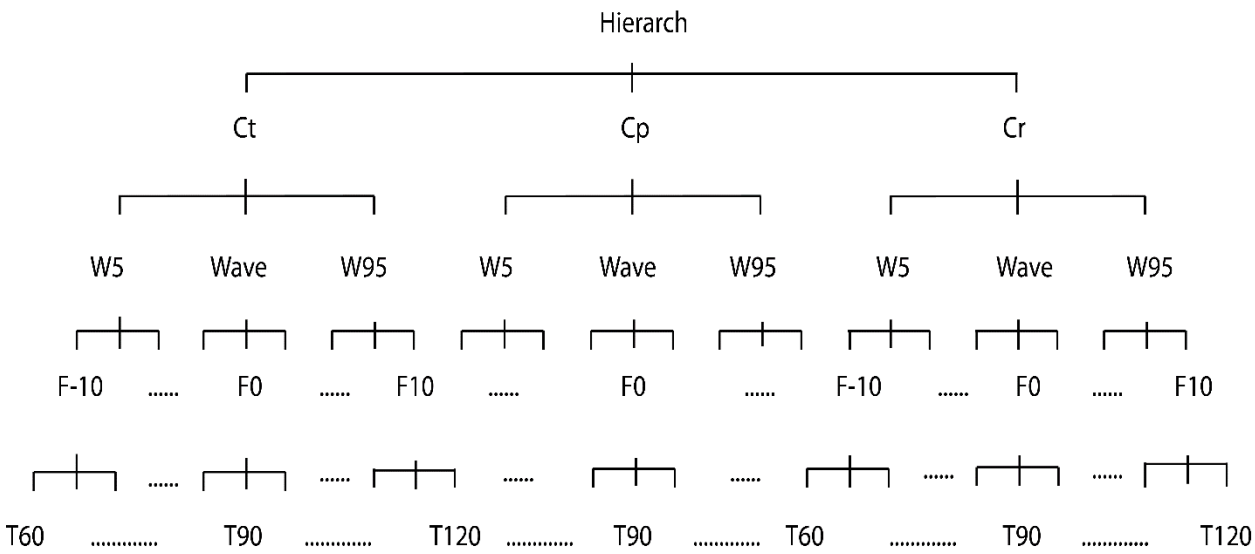

Figure 4.2 The schematic diagram of hierarchical structure (BMA tree)

The impact of uncertainty sources on water stage prediction and prediction variance

In order to evaluate the impact of each type of uncertainty source on the simulated water stage, model perturbation approach is applied to the 243 predictions at base+1 level of BMA tree. First, the 243 predictions are assigned into 81 groups and each group has three predictions which share the same implementations except the one that is being evaluated. For instance, when channel cross section shape is being evaluated, each group has three members that have the same channel width, channel roughness, flow magnitude and topography values whereas the channel cross section shapes are different. Next, the difference between any two predictions are found and averaged for each group. Finally, the averaged values from each group are averaged again among all 81 groups. This process is expressed using Equation 4.14.

$$S = \frac{1}{81}\sum_{i=1}^{81}\frac{1}{3} * \left\{\left[E(d)_{peak}^{p1} - E(d)_{peak}^{p2}\right] + \left[E(d)_{peak}^{p2} - E(d)_{peak}^{p3}\right] + \left[E(d)_{peak}^{p1} - E(d)_{peak}^{p3}\right]\right\}_i$$

(4.14)

Where S is the value to evaluate the impact of one source on the predicted water stage. $\mathrm{E(d)}_{\mathrm{peak}}^{\mathrm{p1}}$ is the one prediction in one group on peak day and *i* is the group number. The impact of each uncertainty source found here is based on the predictions from 243 model implementations. However, this rank excludes the weights of models. When employing the BMA approach for the prediction ensemble, each model implementation obtains its own weight. Therefore, the rank of uncertainty sources may change accordingly. In this case, the relative influence of each uncertainty source can be reflected by the model prediction variance (Chitsazan and Tsai, 2015). Equation 4.13 is used and second to fifth terms represent the influence of channel cross sectional shape, channel width, flow accuracy and topographical data resolution on model prediction variance, respectively.

4.5 Results and discussion

4.5.1 BMA trees on flood peak day

Three BMA trees including BMA tree of model weight/conditional model weight, BMA tree of mean water stage prediction and BMA tree of prediction variance at USGS 07069500 for peak flow are shown in Fig. 4.3-4.5, respectively. Figure 4.3 shows the BMA tree of model weights (in parentheses) and conditional model weights. The model weights reflect the comparative importance of all the competitive modeling predictions in one level whereas the conditional model weights represent the relative importance of different predictions under the same parent model.

The base level in the BMA tree of model weight corresponds to different topography resolutions. It indicates that DEM data with 60m resolution provides the best prediction. Comparatively, 90m DEM yields slightly poorer but acceptable result. However, when the

topography resolution reduces to 120m, the prediction accuracy is seriously affected. Therefore, in general, a higher resolution topography dataset provides better prediction, but the prediction accuracy does not improve monotonically with resolution. The third level corresponds to different flow forcing magnitudes. It demonstrates that the original USGS flow record and the high flow magnitude usually provide better results. On the contrary, the model implemented with low flow forcing provides poorer results. This indicates that some of the USGS flow records tend to slightly underestimate the flow magnitudes. The second level is developed by averaging water stage predictions from their child BMA models that use different flow values. This shows that the average width value is most suitable for this validation station. Similarly, at first level, parabolic cross-sectional shape proves to be the best fit for the model prediction at validation station USGS 07069500. This shows that the combination of average channel width and parabolic channel cross-sectional shape are the closest channel geometry to the reality.

Fig. 4.4 shows the BMA tree of mean water stage predictions during the flow peak (4/26/2011). This figure demonstrates the mean water stage predictions over the accumulation of sources of uncertainty along the BMA tree. At the base level, the mean water stage predictions for peak flow range between 6.41m and 8.41m among different model implementations. As the level goes up in the hierarchical structure, the difference between the largest and smallest prediction becomes smaller. This shows that integrating predictions on different levels using BMA approach will reduce the prediction variability. Fig. 4.5 shows the BMA tree of prediction variance on flood peak day at USGS 07069500. At the base level, the prediction variances result from the hidden level (base+1 level) that reflect the change of model prediction caused by model parameters. As level goes up, the prediction variances at different branches seem to increase in general, and it demonstrates the cumulative effect of variances from all lower levels. When

compared with Fig. 4.3, it indicates that that better performing models (higher weights) do not necessarily have the lowest prediction variances.

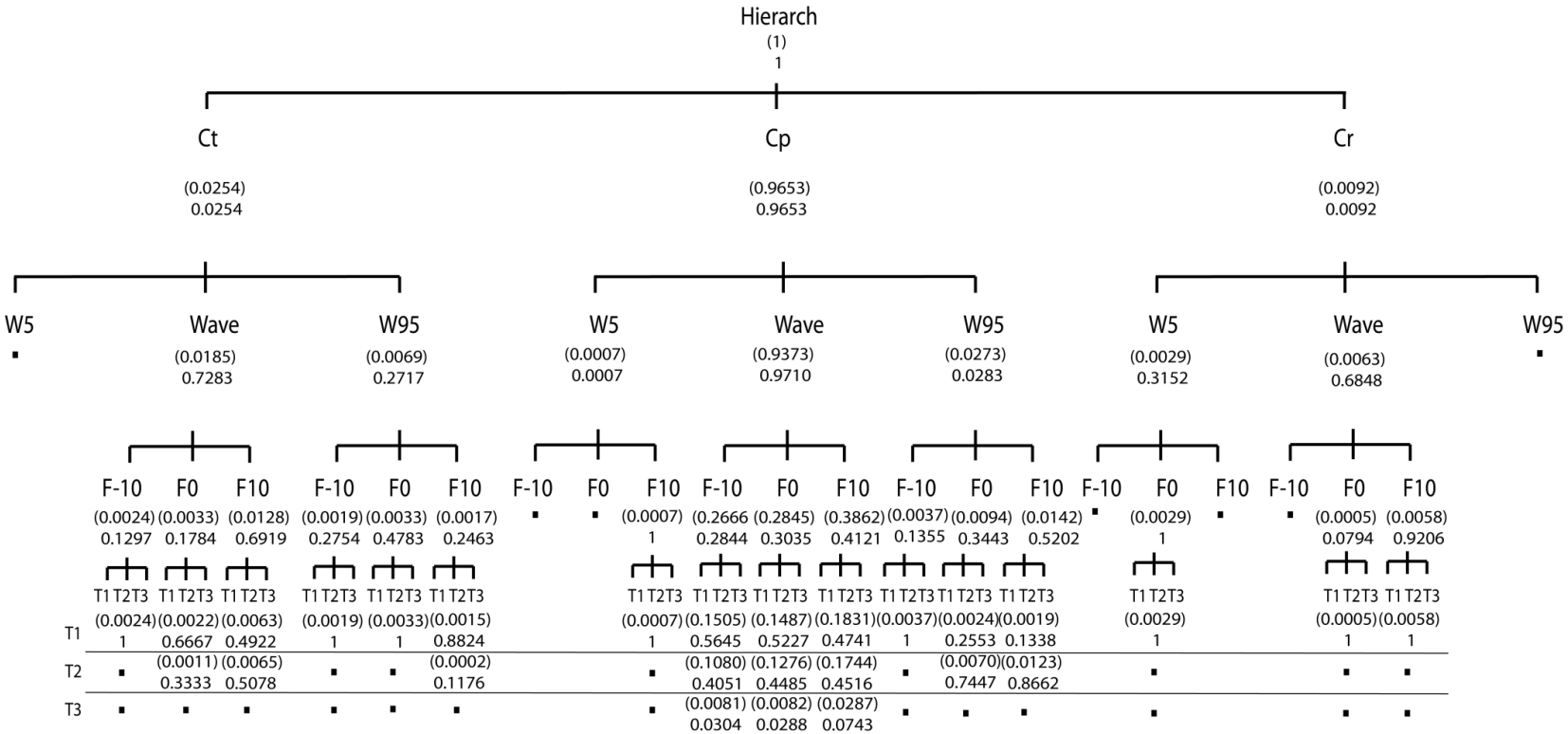

Figure 4.3 BMA tree of model weights and conditional model weights on flood peak day (April 26, 2011) at validation station USGS 07069500

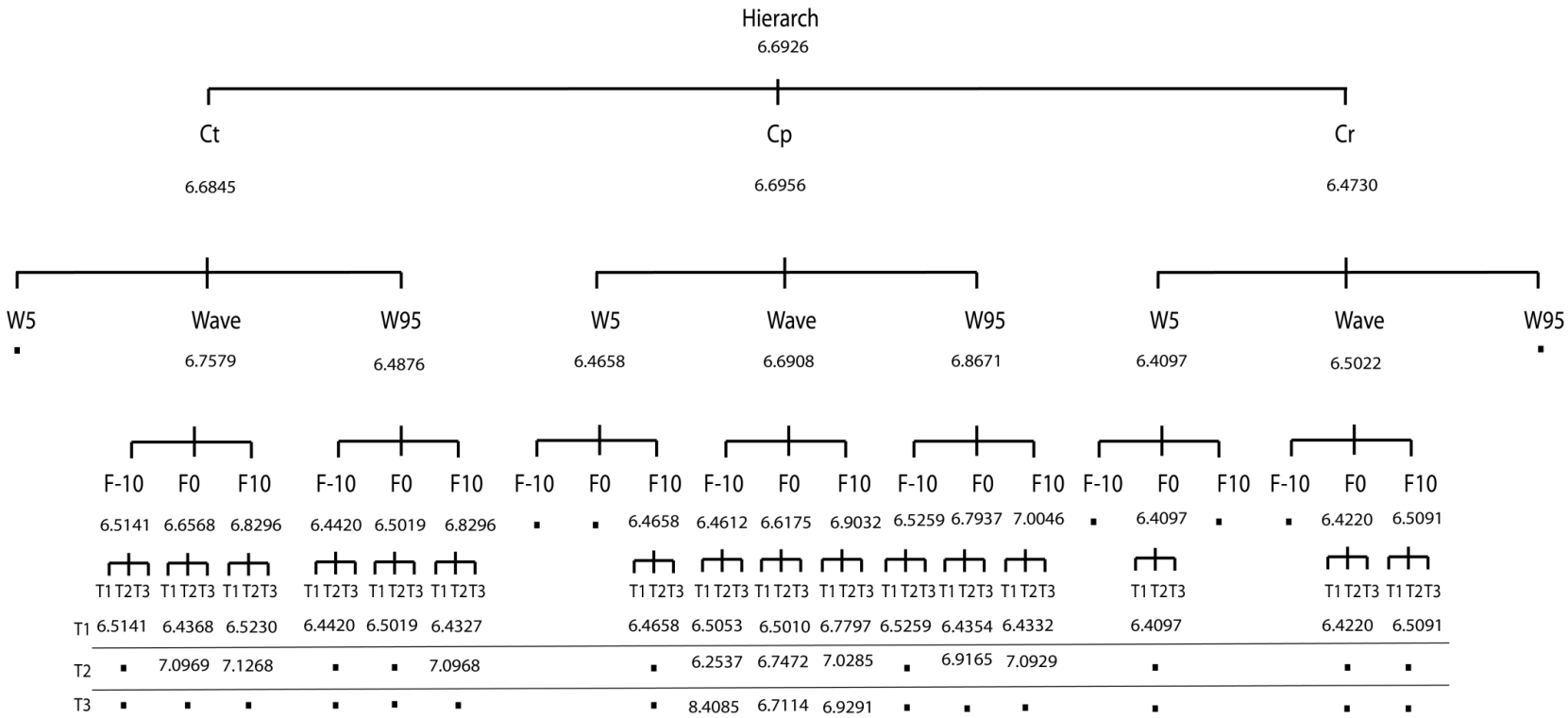

Figure 4.4 BMA tree of mean water stage predictions on peak day (April 26, 2011) at validation station USGS 07069500

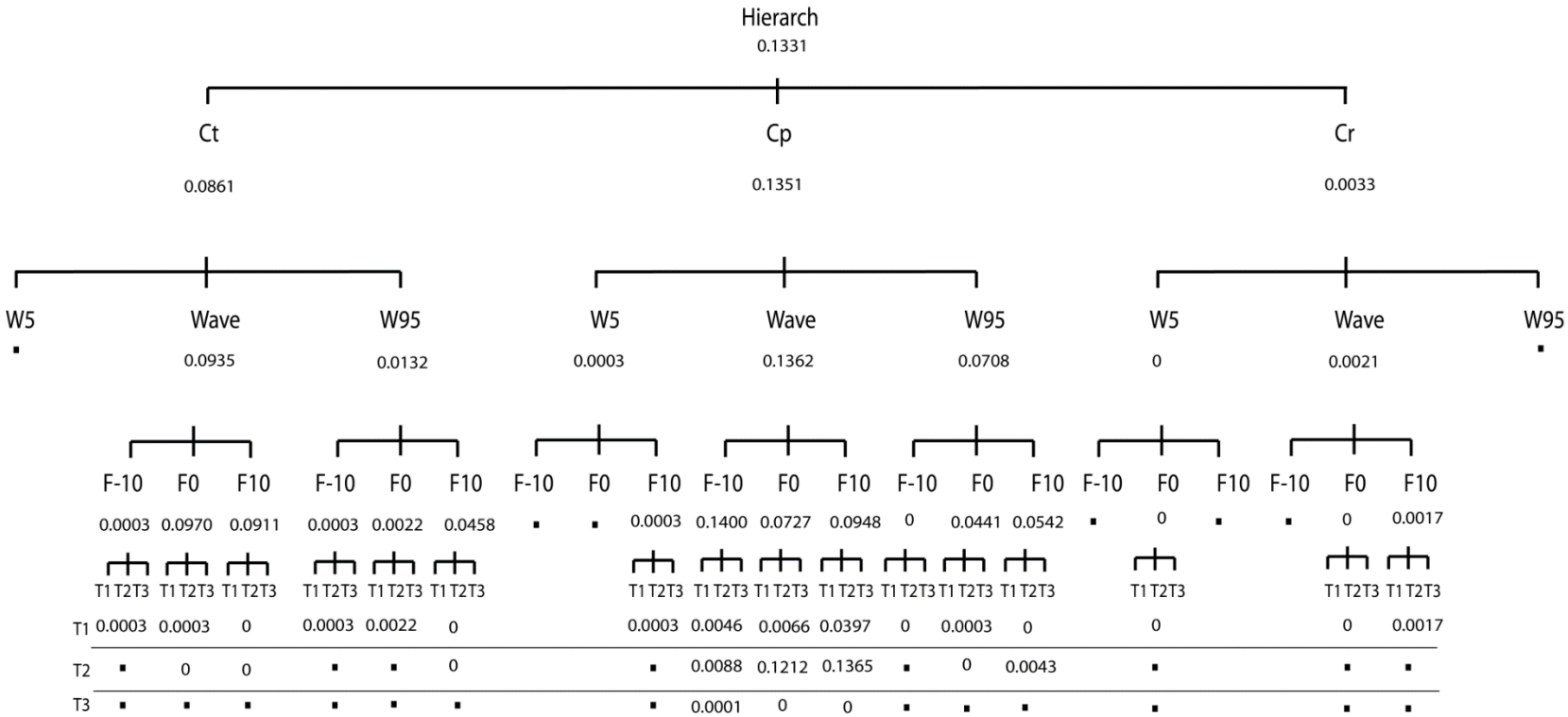

Figure 4.5 BMA tree of prediction variances on peak day (April 26, 2011) at validation station USGS 07069500

4.5.2 The relative importance of each uncertainty sources

Fig. 4.6 shows the relative impact of different uncertainty sources on the mean model prediction on peak flow (4/26/2011). This is achieved through model perturbation approach when BMA weights are not taken into consideration. Fig. 4.6 shows that topography and channel width play more significant roles than flow forcing which has relatively larger impact than the model parameter and cross-sectional shape. Many past studies(Saksena and Merwade, 2015; Sanders, 2007) have shown that the model prediction accuracy is sensitive to DEM data resolution in most cases. This study shows that specifying accurate channel width is also important, and it may have even larger impact on model predictions than topographical data (Fig. 4.6 A&C). Additionally, as cross-sectional area is also affected by width, implementing a smaller channel width reduces the amount of water transported through the cross-section. Since only 10% uncertainty is considered in the flow data, it has relatively less impact compared with other factors. Moreover, channel roughness value is selected from a small range (0.01-0.05) and the model is not always very sensitive to it (Liu et al., 2018).

Fig. 4.7 shows the relative influence of different uncertainty sources on prediction variance when BMA weights are taken into account. In this case, the relative impact is reflected by the model prediction variance(Chitsazan and Tsai, 2015) and this is different from model perturbation method which has an underlying assumption that each model is treated equally. Results show that model inputs (topography and flow forcing) and model parameters have larger influence on model prediction variance compared to the model structure (channel cross-sectional shape and channel width). The impact of channel width on prediction variance is not as significant as its impact on model prediction as shown in Fig. 4.6. This is due to the fact that in most cases one unique width value obtains much higher weight and the model prediction does not vary

significantly after implementing this width value. On the contrary, the impact of roughness parameter and flow are amplified since high weights are obtained by different values of these two sources and the variance of predicted water stage caused by these two sources are large..

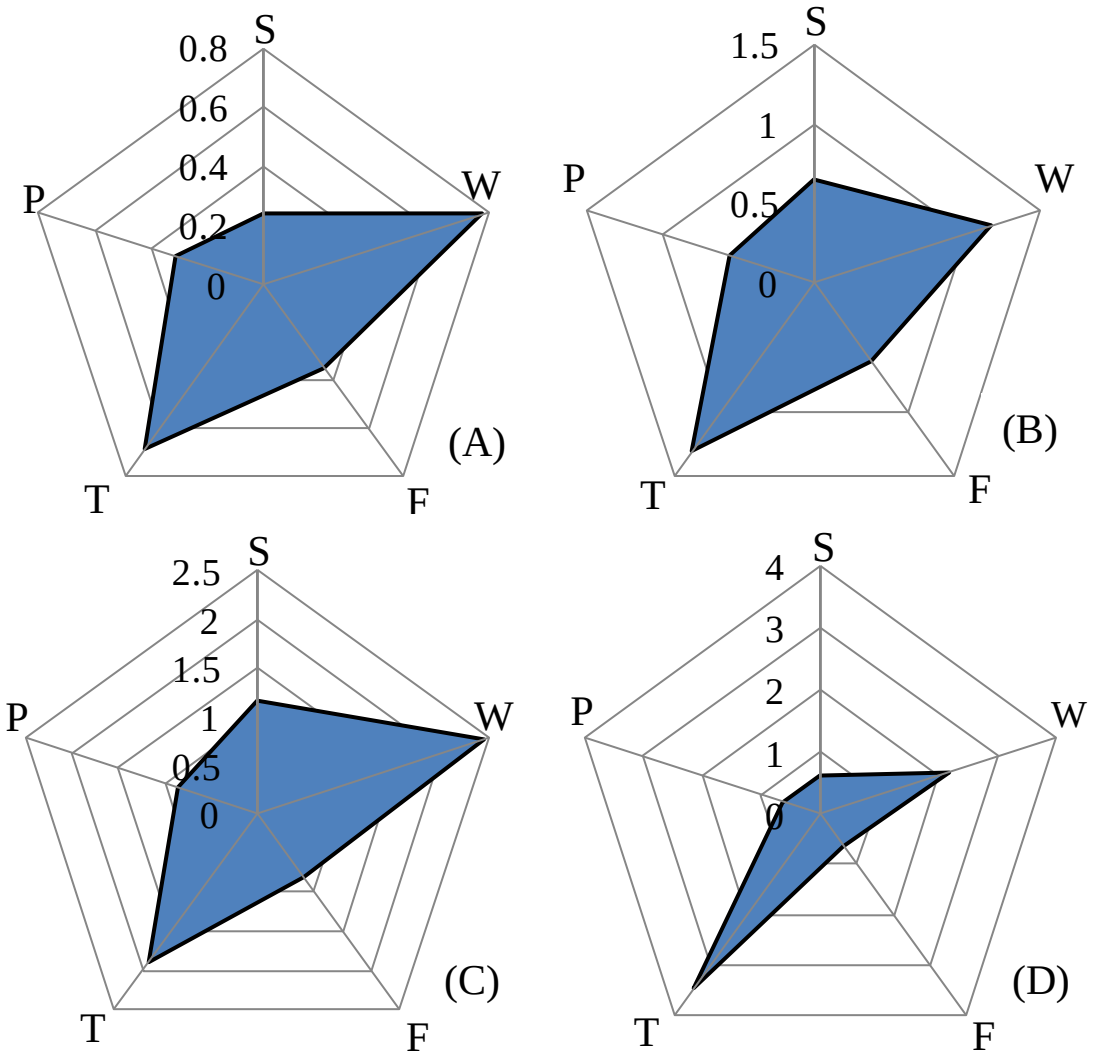

Figure 4.6 Relative impact of uncertainty sources on predicted water stage using model perturbation approach at four validation USGS stations (S is channel cross-sectional shape, P is parameter, T is topography resolution, F is flow magnitude and W is channel width).

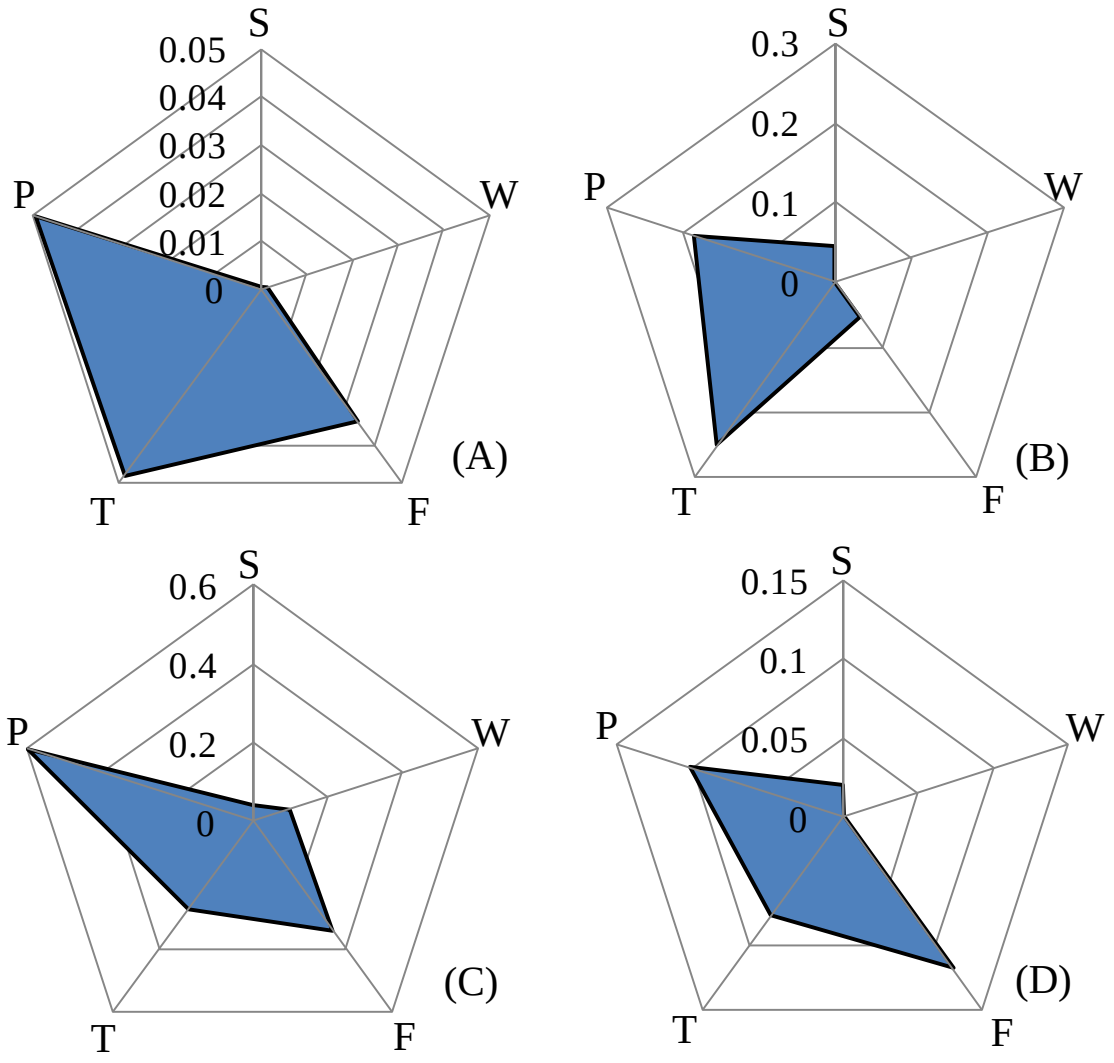

Figure 4.7 Relative impact of uncertainty sources on water stage prediction variance considering BMA weights at four validation USGS stations (S is channel cross-sectional shape, P is parameter, T is topography resolution, F is flow magnitude and W is channel width).

4.5.3 Model mean prediction accuracy in the hierarchical structure

In order to evaluate model mean prediction at each level of the BMA tree, two branches including a high weight model branch and a relatively low weight model branch are evaluated. Predictions of the high weight model, the low weight model and ensemble of other good models (weights are above 0.0001) are shown in Fig. 4.8 for four validation stations. To see a clear trend,

the mean model predictions are compared with observed water stage data for each level at USGS 07069500 as shown in Fig. 4.9-4.10. Results show that R^2 values generally increase from the base level up to the higher level. Additionally, Root mean square errors (RMSE) between model simulations and observed water stages are also calculated at different levels. The results, presented in Table 4.4, show that the model mean prediction is closer to the observation as level goes up. The improvement of RMSE is remarkable for the worst performing branch whereas it is not very significant for the best branch. This is because BMA prediction may not always outperform the best model. Similarly, the change of R^2 and RMSE at all four validation stations, presented in Fig. 4.11 and Table 4.4, show similar trend. Since models with better performance are trusted more than others in the BMA process, as the model integrating more uncertainty sources move up in the hierarchy, the mean prediction gets better.

Table 4.4 Root mean square error (RMSE) between predicted and observed water stage at all levels in the BMA tree for four validation stations

RMSE		Base+1	Base	3rd	2nd	1st	Hierarch
Validation Station (A)	Best Branch	0.415	0.426	0.409	0.428	0.425	0.426
	Worst Branch	0.601	0.601	0.511	0.502	0.457	0.426
Validation Station (B)	Best Branch	0.483	0.486	0.485	0.483	0.480	0.476
	Worst Branch	0.867	0.796	0.730	0.583	0.523	0.476
Validation Station (C)	Best Branch	0.651	0.618	0.581	0.511	0.502	0.511
	Worst Branch	0.753	0.660	0.642	0.519	0.515	0.511
Validation Station (D)	Best Branch	0.362	0.345	0.312	0.268	0.268	0.251
	Worst Branch	0.463	0.363	0.336	0.268	0.268	0.251

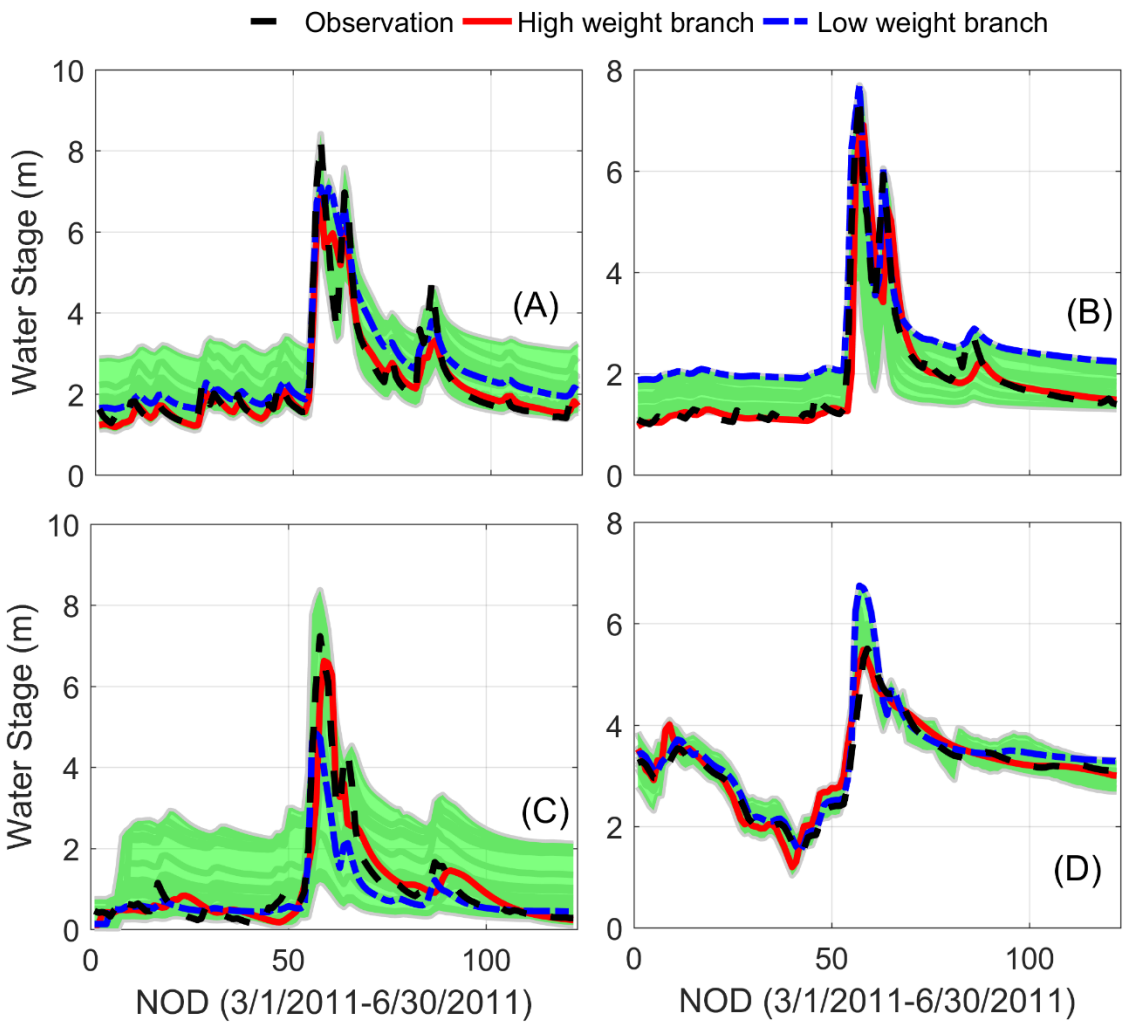

Figure 4.8 Model (weight>=0.0001) prediction ensemble, high weight model, low weight model as well as observed water stage at four validation USGS stations.

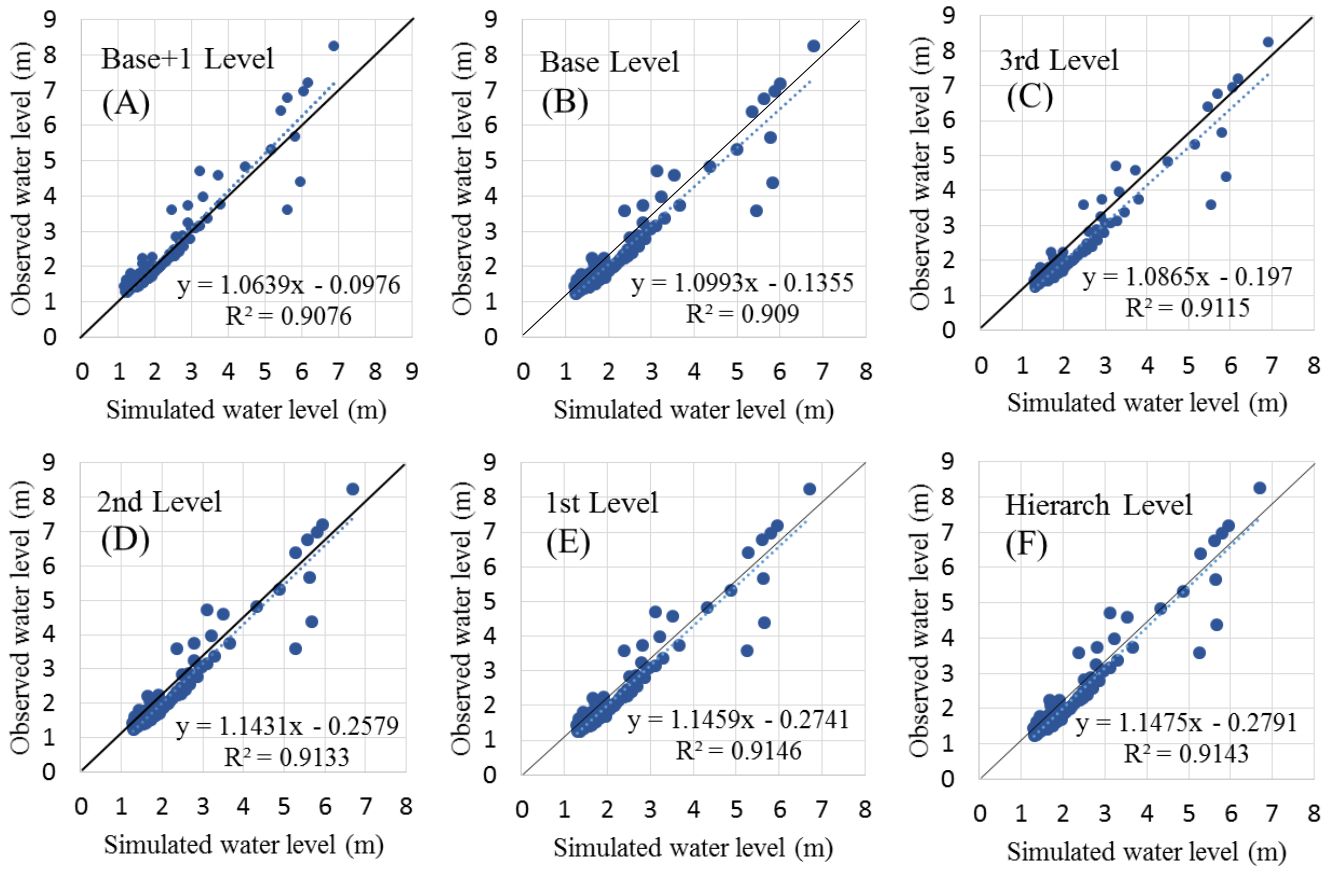

Figure 4.9 Simulated and observed water stages of high weight model branch at different levels of BMA tree for USGS 07069500

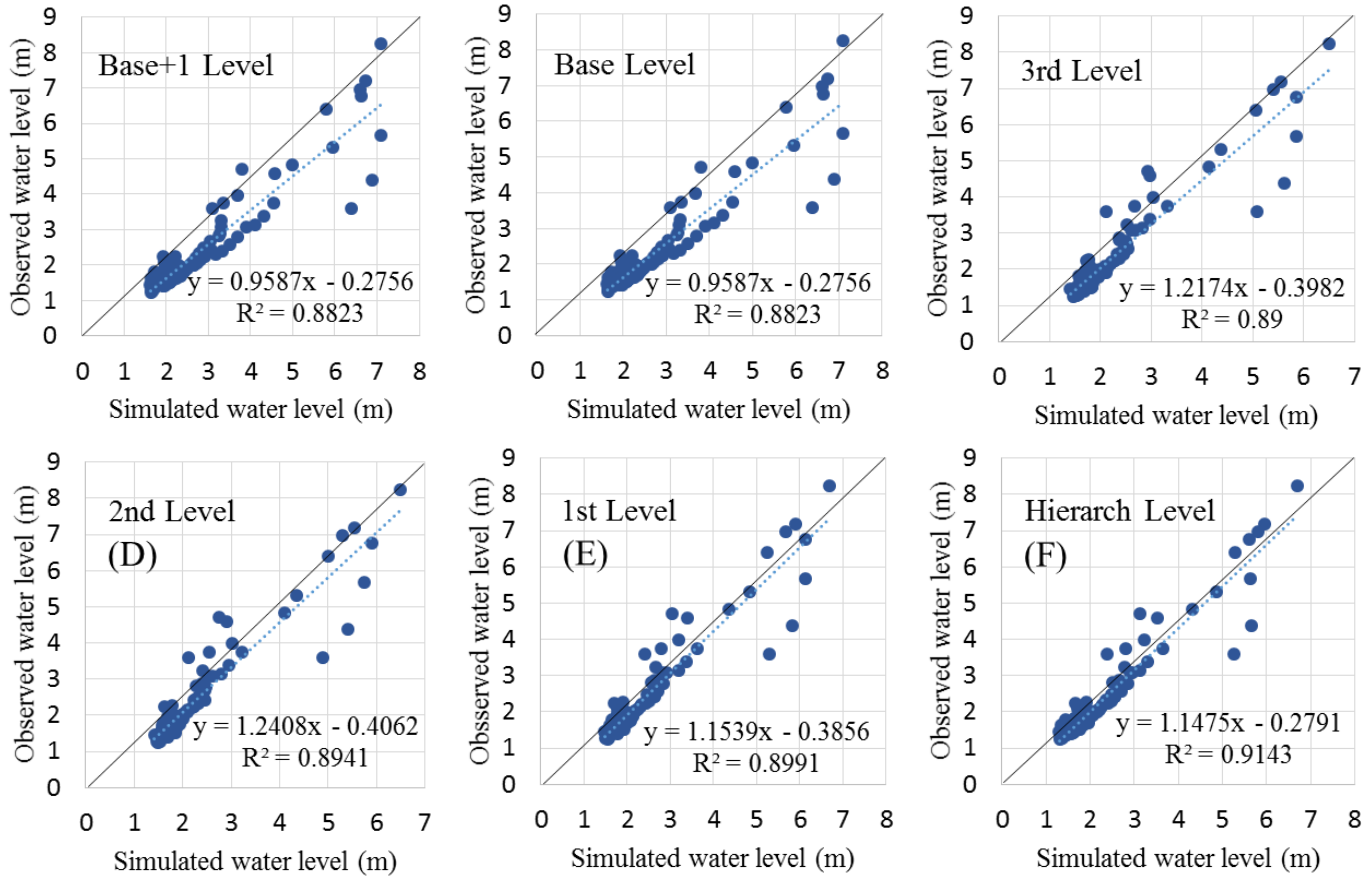

Figure 4.10 Simulated and observed water stages of low weight model branch at different levels of BMA tree for USGS 07069500

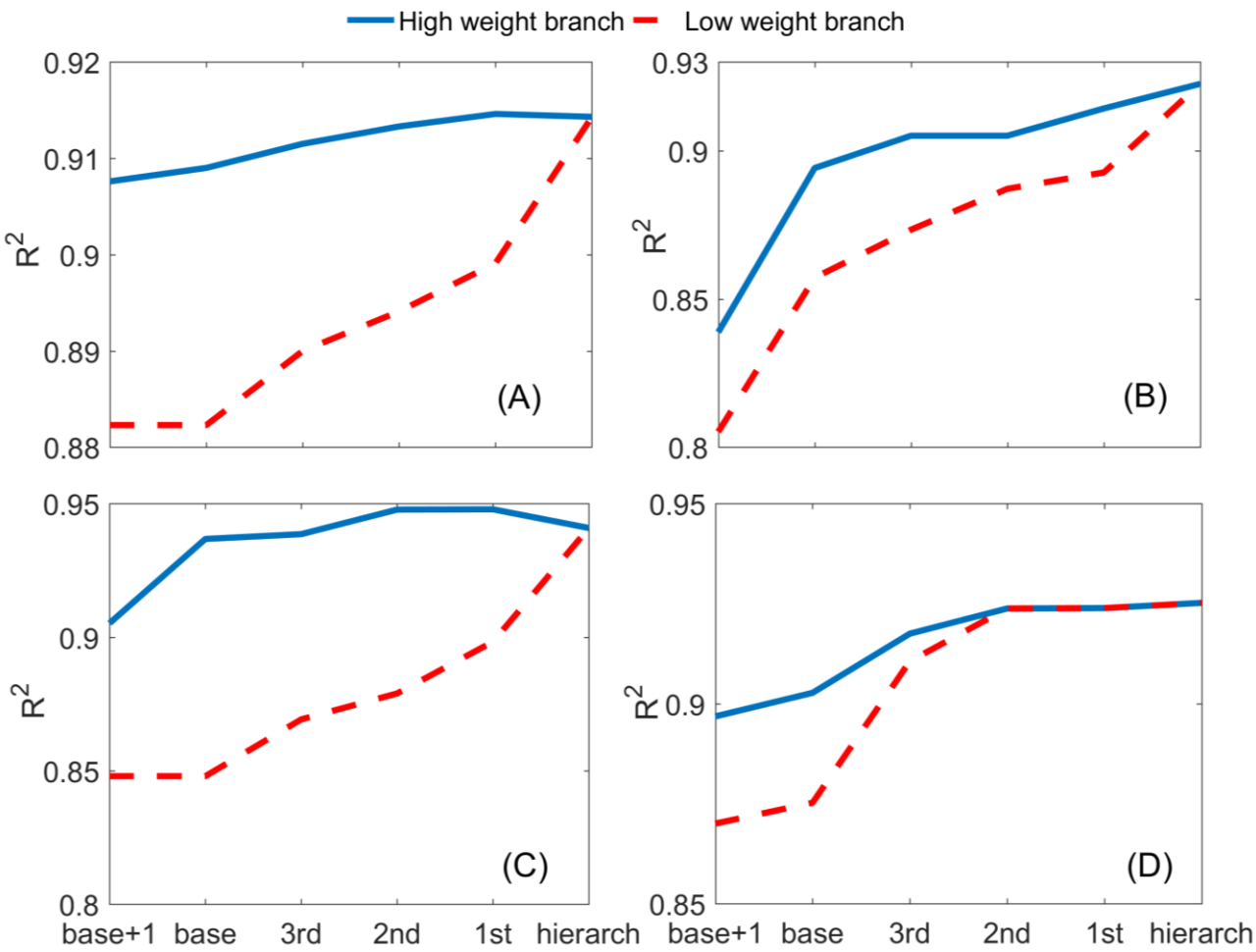

Figure 4.11 R^2 of predicted water stages on different levels of BMA tree for four validation USGS stations

4.5.4 Uncertainty associated with mean prediction at each level

Although the deterministic mean water stage predictions at each level of BMA tree are determined, there exists uncertainty associated with them. To evaluate this uncertainty, 95% confidence intervals associated with the deterministic predictions at each level are calculated. Fig. 4.12-4.13 display the 95% confidence intervals for mean water stage prediction at different levels for the high and low weight branches of USGS 07069500. The band widths are calculated and it indicates that the uncertainty associated with mean prediction increases as the level goes

up in the hierarchical structure. Although the model mean prediction performs better as the level moves up in the BMA tree in general, the uncertainty associated with the mean predictions becomes larger. Fig. 4.13 also demonstrates that mean prediction of the low weight model deviates obviously from the observation data at base level and the 95% confidence interval is very narrow which could not cover the observed water stage. However, when level moves up in the BMA tree, the band width could mask most of the observed water stage. This demonstrates that the deterministic prediction at hierarchical level and its uncertainty bounds are more trustworthy for decision makers.

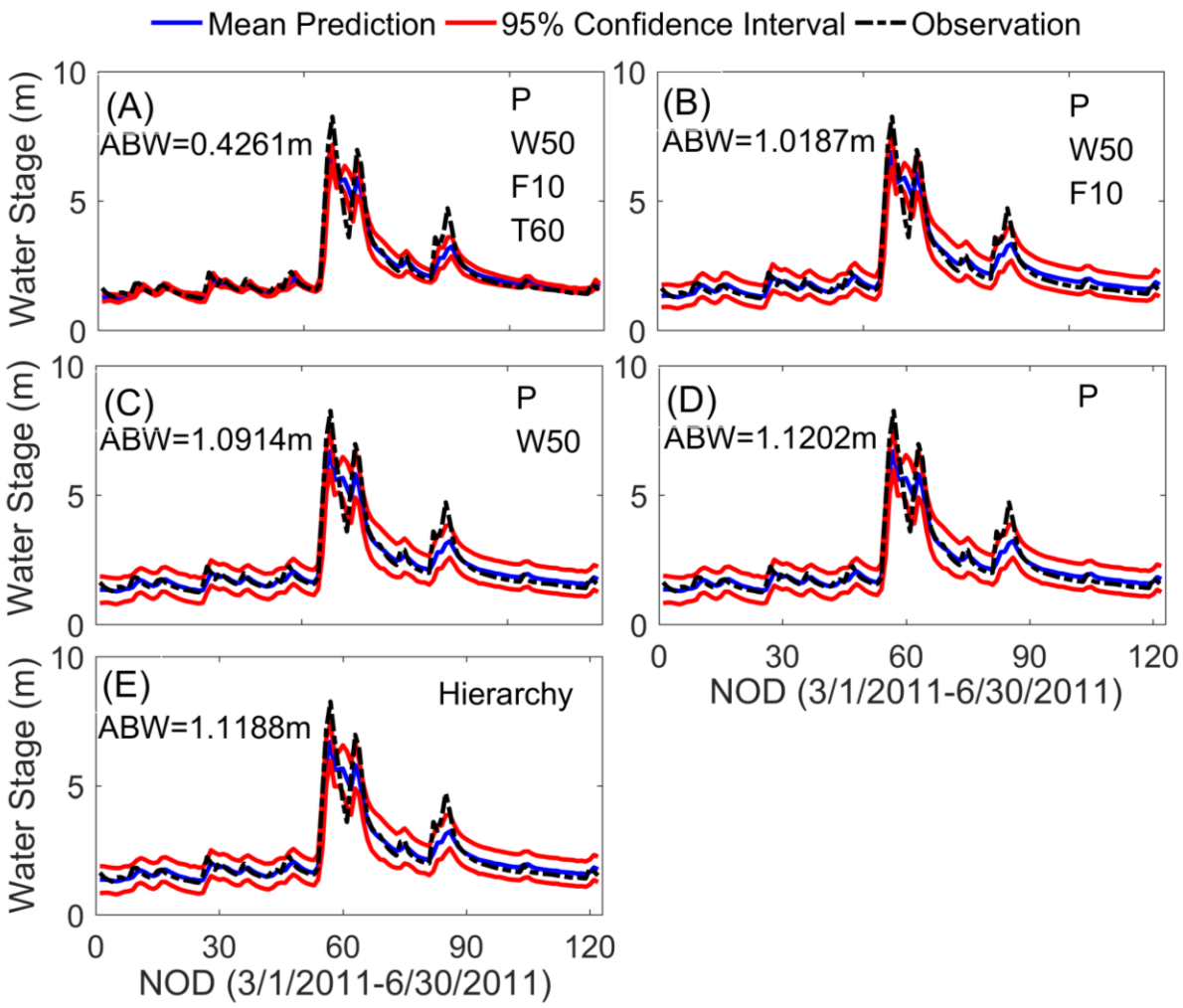

Figure 4.12 Mean water stage predictions at different levels of BMA tree and their 95% confidence interval bounds for the high weight branch at USGS 07069500

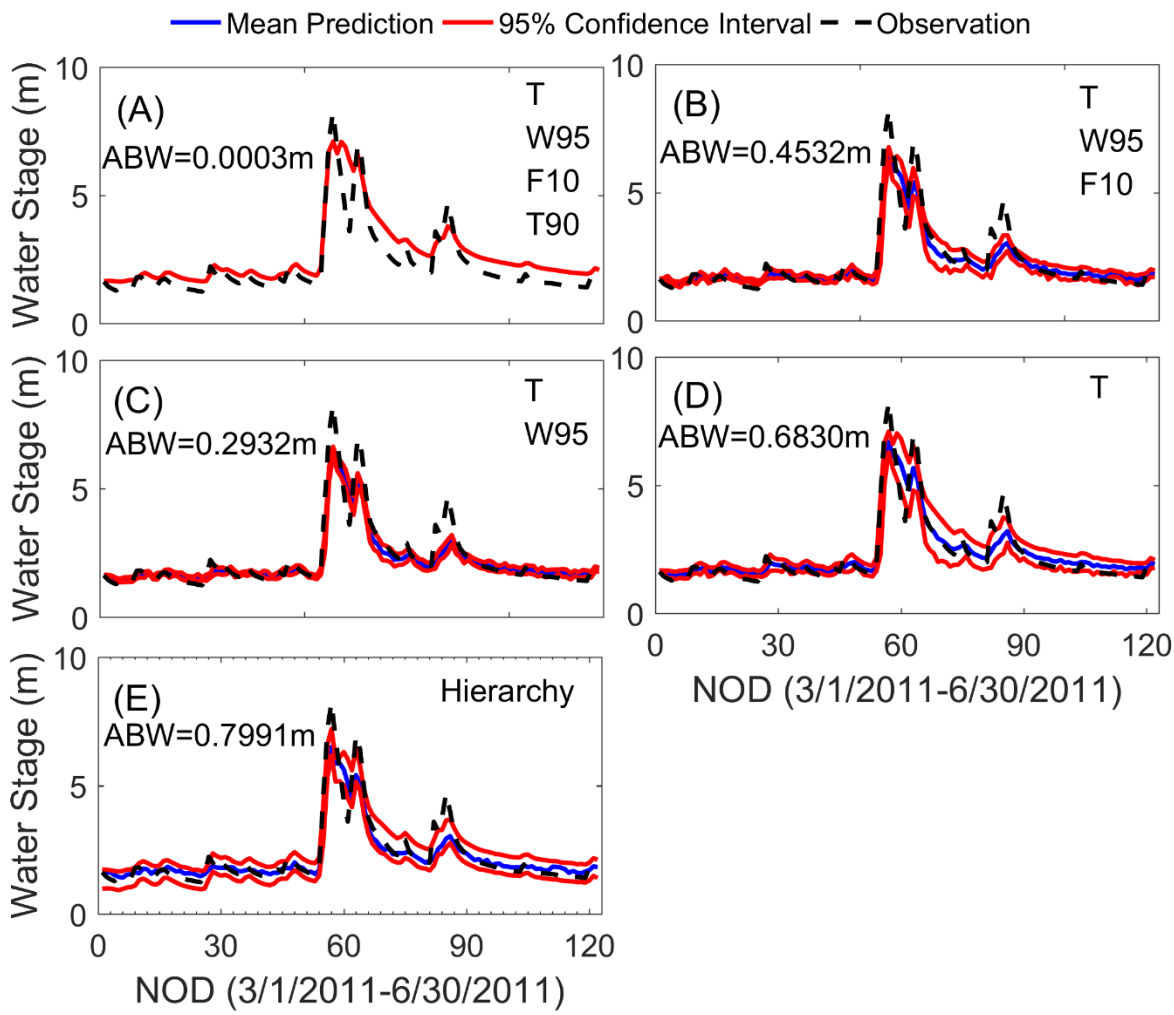

Figure 4.13 Mean water stage predictions at different levels of BMA tree and their 95% confidence interval bounds for the low weight branch at USGS 07069500

4.6 Conclusion

Flood risk management and mitigation require reliable predictions of water stage and flood inundation extent. Instead of trusting one specific model implementation with respect to model structure, parameter and input forcing, Bayesian model averaging approach can be used as a multi-model combining method to investigate the combined effects of uncertainty from these sources and provide a reliable deterministic prediction as well as the prediction distribution. However, evaluating the cumulative impact from a group of major uncertainty sources alone is

not adequate, the relative influence of different uncertainty sources also needs to be investigated. The hierarchical Bayesian model averaging approach (HBMA) extends the application of the BMA method by explicitly considering different sources of uncertainty in the hierarchical structure (BMA tree) for analyzing model uncertainty and uncertainty propagation. Each uncertainty source is located at one level of the hierarchical structure and three BMA trees corresponding to model weight, mean prediction and prediction variance are created respectively. BMA procedure is conducted upwards from base level to the hierarch level for each branch in the BMA tree of mean prediction. The relative impact of different uncertainty sources on model prediction (not under BMA frame) and prediction variance (under BMA frame) are prioritized. A high weight and a relative low weight branches in the BMA tree are evaluated at each level to find out how model performances are changed and propagated along the hierarchical structure. Moreover, the uncertainty associated with mean predictions at each level is also investigated. The following conclusions are drawn from this study:

1. In general, channel width and topographical data resolution have largest impact on the hydrodynamic model predictions. These two sources are followed by flow forcing, which has relatively greater influence than channel cross-sectional shape and model parameter. This result is derived from model perturbation approach when model weights are not taken into consideration. However, when dealing with the ensemble forecasting systems, multi-model combing approach is usually used and model weights are considered. In this situation, the relative influence of different uncertainty sources could be reflected by model prediction variance. Results from this study indicate that model inputs (topography and input forcing) and model parameter (roughness) have larger impact on prediction variance than model structure (channel shape and width).

2. BMA trees of model weights/conditional model weights evaluate the comparative predictions in one level. In general, a higher resolution topography data provides better prediction, however, the improvement may not be significant after the resolution improves to a certain level. Some of the USGS station records tend to slightly underestimate the flow value since assuming a 10% higher flow sometimes provide better outcome. For the validation station USGS 07069500, implementing average width value and parabolic cross-sectional shape produces much better prediction compared with other types of model setup. This might be because these two features are best fitted for the real condition for the site.

3. The BMA tree of mean water stage predictions provides an understanding of mean prediction variability over the accumulation of sources of uncertainty along the BMA tree. As the level goes up from base level towards the hierarchical level, the difference between the largest and smallest predictions becomes smaller. BMA tree of prediction variance demonstrates that those better performed models (highest weights) do not necessarily have the lowest prediction variances.

4. As level moves up in the BMA tree, the accuracy of mean prediction also increases in general. The R^2 and RMSE values of both high weight branch and low weight branch for a given validation station show improvement at a higher level. The improvement is more significant for the worst branch whereas it may not be very remarkable for the best branch in many cases, and eventually they will merge together in the hierarchical level. This is because the BMA approach weights and averages all the models to provide a relatively reliable prediction, which may not be always better than the best model prediction.

5. The uncertainty associated with the mean predictions of high weight model and low weight model are investigated at each level. The band widths of 95% confidence interval display a increasing trend as the level goes higher. This indicates that although the mean prediction of

hierarch level is more reliable than lower levels, the uncertainty associated with it could be larger. Therefore, observed water stage is more likely to fall within this range.

There exist certain limitations for this study. The evaluation of relative influence of different uncertainty sources is under a pre-defined range. For instance, flow is considered with only 10% bias based on previous studies and this resulted in showing relatively less impact on model prediction. A higher bias in flow may give different results. Although the same box-cox transformation coefficient is applied to all the model predictions to adjust the skewness, the transformed PDF remains bounded. Regardless of the limitations, findings from this study provide useful information for flood modelers, especially for those who use the ensemble model forecasting systems. For modelers that use BMA method or other multi-model combining approach to derive a deterministic prediction from prediction ensemble, they should pay more attention to the data quality for model input and parameter instead of spending lots of efforts on finding the right model structure because the the best fitted model can be obtained during the BMA process.

CHAPTER 5. SYNTHESIS

The results from this work show that epistemic uncertainty sources play a major role in water stage and flood inundation extent predictions from hydrodynamic models. Understanding the role of individual uncertainty source as well as their cumulative impacts on model performance can help modelers and decision makers to reduce epistemic uncertainty, generate reliable predictions and make proper decisions. The major findings of this dissertation are described below.

5.1 Impact of model structure uncertainty on flood predictions

Model structure uncertainty can be reflected by the dimensionality of the model, the form of governing equations and the geometry description in the model. Results from this dissertation clearly indicate that the proper selection of a hydrodynamic model structure based on the needs is very crucial. Overall, a 2D model provides better predictions than a 1D model. However, a 1D model does not always underperform compared to a 2D model. At low channel roughness, HEC-RAS 1D performs better than the 2D version for four different sites with various geophysical conditions. Given that a 1D model is computationally much faster than a 2D model, selection of 1D model is preferable for situations that require quick simulation and decision. For a given dimensionality (1D or 2D), the form of governing equations also affect simulation outcomes. An earlier version of LISFLOOD-FP model solves the diffusive wave equations that simplified from the full Saint-Venant equations. This model is computationally fast, but it overestimates the flood inundation extents. Results from this study also show that channel width is one of the dominant factors that affect model predictions. Also, cross-sectional shape which is related to the water transportation capacity is influenced by the channel width. Implementing a low channel

width usually produces poor results. Due to the lack of channel width data especially for lower order rivers worldwide, the accuracy and reliability of hydrodynamic model prediction are reduced. In the future, a database of channel width including both major rivers and lower order rivers should be set up in order to fulfil the needs of highly accurate flood predictions.

5.2 Impact of model parameter and input uncertainty on flood predictions

All flood inundation models are calibrated using channel roughness and floodplain roughness. In a large scale study, due to the variability of floodplain conditions across the region, modelers usually adopt distributed roughness values that are derived from the landuse information. However, in reach scale study, the traditional practice requires the use of single channel roughness value and single floodplain roughness value for model calibration. A comparison of applying distributed floodplain roughness versus unique floodplain roughness at four reach scale study sites with different geographical settings show that adopting distributed floodplain roughness does not really improve the model performance. Additionally, it is found that adopting relatively higher channel and floodplain roughness values in their common range usually generate better results. The major inputs for a flood inundation model are topography and flow forcing. Results from this study show that topography data resolution and channel width play important roles in creating accurate hydrodynamic model simulations whereas the influence from streamflow input and channel roughness is much smaller. However, in the model selection process such as HBMA, when weights are taken into account, streamflow input and channel roughness have large impact on the prediction variance.

5.3 Impact of individual and combined uncertainty sources on flood predictions

Flood inundation models are affected by multiple epistemic uncertainty sources, including model structure, parameter and input forcing, among others. Each of these uncertainty sources has their own influence on the model performance and they also interact with each other to affect the final model prediction. Different combinations of model implementation may produce the same prediction (equifinality). This study attempts to find a way to produce reliable predictions considering the combined influence from various major epistemic uncertain sources as well as separating and analysing the role of individual uncertainty sources. Bayesian model averaging is a promising approach to be applied to the flood inundation modeling field for robust prediction considering the combined impact from different uncertainty sources. Hierarchical Bayesian model averaging and model perturbation approaches is used to analyze the impact of individual uncertainty sources and uncertainty propagation. Results show that channel width and topographical data resolution have largest impact on the hydrodynamic model predictions whereas model input (topography and input forcing) and model parameter (roughness) have larger impact on prediction variance than model structure (channel shape and width) when model weights are taken into account.

5.4 Limitations and future work

There exist certain limitations that merit attention in this dissertation. Channel width information is derived based on google images in this research, however, the width derived this way may not be the bankful width. Also, the topographical resolution we considered is horizontal resolution whereas the accuracy of inundation extents are affected by the vertical accuracy of topography data as well. Additionally, when the four models' performances are compared, inundation extent alone may not be enough to decide whether a specific model structure is good or not. Besides

this, the Box-Cox transformation technique is not a perfect method to transform data from non-Gaussian space to Gaussian space because the data after transformation are still bounded. In the future, a new set of bankful width data could be derived based on the LiDAR DEM. Also, the vertical resolution of topographical data could be considered. Besides inundation extents, more indicators such as water depth, velocity could be evaluated to produce a thorough picture among different model structures. A group of copula futures can be applied to remove the normal distribution assumption for the BMA approach.

REFERENCES

Ackerman, C.T., 2005. HEC-GeoRAS; GIS Tools for support of HEC-RAS using ArcGIS. United States Army Corps of Engineers, Davis.

Ajami, N.K., Duan, Q., Gao, X., Sorooshian, S., 2006. Multimodel combination techniques for analysis of hydrological simulations: Application to distributed model intercomparison project results. Journal of Hydrometeorology, 7(4): 755-768.

Ajami, N.K., Duan, Q., Sorooshian, S., 2007. An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resources Research, 43(1).

Alfieri, L. et al., 2014. Advances in pan-European flood hazard mapping. Hydrological Processes, 28(13): 4067-4077.

Alho, P., Aaltonen, J., 2008. Comparing a 1D hydraulic model with a 2D hydraulic model for the simulation of extreme glacial outburst floods. Hydrological Processes, 22(10): 1537-1547.

Altenau, E.H., Pavelsky, T.M., Bates, P.D., Neal, J.C., 2017. The effects of spatial resolution and dimensionality on modeling regional-scale hydraulics in a multichannel river. Water Resources Research, 53(2): 1683-1701.

Apel, H., Aronica, G., Kreibich, H., Thieken, A., 2009. Flood risk analyses—how detailed do we need to be? Natural Hazards, 49(1): 79-98.

Apel, H., Merz, B., Thieken, A.H., 2008. Quantification of uncertainties in flood risk assessments. International Journal of River Basin Management, 6(2): 149-162.

Apel, H., Thieken, A.H., Merz, B., Blöschl, G., 2004. Flood risk assessment and associated uncertainty. Natural Hazards and Earth System Science, 4(2): 295-308.

Arnell, N.W., Gosling, S.N., 2016. The impacts of climate change on river flood risk at the global scale. Climatic Change, 134(3): 387-401.

Aronica, G., Bates, P., Horritt, M., 2002. Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. Hydrological Processes, 16(10): 2001-2016.

Aronica, G., Hankin, B., Beven, K., 1998. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. Advances in Water Resources, 22(4): 349-365.

Bales, J., Wagner, C., 2009. Sources of uncertainty in flood inundation maps. Journal of Flood Risk Management, 2(2): 139-147.

Bates, P., Horritt, M., Smith, C., Mason, D., 1997. Integrating remote sensing observations of flood hydrology and hydraulic modelling. Hydrological Processes, 11(14): 1777-1795.

Bates, P., Marks, K., Horritt, M., 2003. Optimal use of high-resolution topographic data in flood inundation models. Hydrological Processes, 17(3): 537-557.

Bates, P., Neal, J., Dabrowa, A., 2013. LISFLOOD-FP User Manual. University of Bristol, Bristol.

Bates, P.D., 2012. Integrating remote sensing data with flood inundation models: how far have we got? Hydrological Processes, 26(16): 2515-2521.

Bates, P.D., De Roo, A., 2000. A simple raster-based model for flood inundation simulation. Journal of hydrology, 236(1): 54-77.

Bates, P.D., Horritt, M.S., Fewtrell, T.J., 2010. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. Journal of Hydrology, 387(1): 33-45.

Beck, J., 2016. Comparison of Three Methodologies for Quasi-2D River Flood Modeling with SWMM5. Journal of Water Management Modeling.

Bermúdez, M. et al., 2017. Quantifying local rainfall dynamics and uncertain boundary conditions into a nested regional-local flood modeling system. Water Resources Research.

Beven, K., Hall, J., 2014. Applied uncertainty analysis for flood risk management. World Scientific.

Blöschl, G., Reszler, C., Komma, J., 2008. A spatially distributed flash flood forecasting model. Environmental Modelling & Software, 23(4): 464-478.

Bladé, E. et al., 2012. Integration of 1D and 2D finite volume schemes for computations of water flow in natural channels. Advances in Water Resources, 42: 17-29.

Box, G.E., Cox, D.R., 1964. An analysis of transformations. Journal of the Royal Statistical Society. Series B (Methodological): 211-252.

Bren, L., Gibbs, N., 1986. Relationships between flood frequency, vegetation and topography in a river red gum forest. Australian Forest Research (Australia).

Brunner, G.W., 2002. HEC-RAS River Analysis System: User's Manual. US Army Corps of Engineers, Institute for Water Resources, Hydrologic Engineering Center.

Chang, L.-C., Shen, H.-Y., Wang, Y.-F., Huang, J.-Y., Lin, Y.-T., 2010. Clustering-based hybrid inundation model for forecasting flood inundation depths. Journal of hydrology, 385(1): 257-268.

Chitsazan, N., Tsai, F.T.-C., 2014. Uncertainty segregation and comparative evaluation in groundwater remediation designs: a chance-constrained hierarchical Bayesian model averaging approach. Journal of Water Resources Planning and Management, 141(3): 04014061.

Chitsazan, N., Tsai, F.T.C., 2015. A hierarchical Bayesian model averaging framework for groundwater prediction under uncertainty. Groundwater, 53(2): 305-316.

Chowdhury, S., Sharma, A., 2009. Multisite seasonal forecast of arid river flows using a dynamic model combination approach. Water resources research, 45(10).

Cook, A., Merwade, V., 2009. Effect of topographic data, geometric configuration and modeling approach on flood inundation mapping. Journal of Hydrology, 377(1): 131-142.

Cullen, A.C., Frey, H.C., 1999. Probabilistic techniques in exposure assessment: a handbook for dealing with variability and uncertainty in models and inputs. Springer Science & Business Media.

Das, T., Maurer, E.P., Pierce, D.W., Dettinger, M.D., Cayan, D.R., 2013. Increases in flood magnitudes in California under warming climates. Journal of Hydrology, 501: 101-110.

de MOEL, H., Aerts, J., 2011. Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates. Natural Hazards, 58(1): 407-425.

de Paiva, R.C.D. et al., 2013. Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin. Water Resources Research, 49(3): 1226-1243.

Demeritt, D. et al., 2007. Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting. Environmental Hazards, 7(2): 115-127.

Deniz, D., Arneson, E.E., Liel, A.B., Dashti, S., Javernick-Will, A.N., 2017. Flood loss models for residential buildings, based on the 2013 Colorado floods. Natural Hazards, 85(2): 977-1003.

Dimitriadis, P. et al., 2016. Comparative evaluation of 1D and quasi-2D hydraulic models based on benchmark and real-world applications for uncertainty assessment in flood mapping. Journal of Hydrology, 534: 478-492.

Dottori, F., Di Baldassarre, G., Todini, E., 2013. Detailed data is welcome, but with a pinch of salt: Accuracy, precision, and uncertainty in flood inundation modeling. Water Resources Research, 49(9): 6079-6085.

Duan, Q., Ajami, N.K., Gao, X., Sorooshian, S., 2007. Multi-model ensemble hydrologic prediction using Bayesian model averaging. Advances in Water Resources, 30(5): 1371-1386.

Duong, V.N., Gourbesville, P., 2016. Model Uncertainty in Flood Modelling. Case Study at Vu Gia Thu Bon Catchment-Vietnam. Procedia Engineering, 154: 450-458.

Dysarz, T., Wicher-Dysarz, J., Sojka, M., 2014. Two approaches to forecasting of sedimentation in the Stare Miasto reservoir, Poland. Taylor & Francis Group, London: 119-127.

Elshorbagy, A. et al., 2017. Topography-and nightlight-based national flood risk assessment in Canada. Hydrology and Earth System Sciences, 21(4): 2219-2232.

Fernandez, C., Ley, E., Steel, M.F., 2001. Benchmark priors for Bayesian model averaging. Journal of Econometrics, 100(2): 381-427.

Finaud-Guyot, P., Delenne, C., Guinot, V., Llovel, C., 2011. 1D–2D coupling for river flow modeling. Comptes Rendus Mécanique, 339(4): 226-234.

Follum, M.L., Tavakoly, A.A., Niemann, J.D., Snow, A.D., 2017. AutoRAPID: A Model for Prompt Streamflow Estimation and Flood Inundation Mapping over Regional to Continental Extents. JAWRA Journal of the American Water Resources Association, 53(2): 280-299.

Freer, J. et al., 2013. Flood risk and uncertainty.

Fu, J., Hsu, M., Duann, Y., 2016. Development of roughness updating based on artificial neural network in a river hydraulic model for flash flood forecasting. Journal of Earth System Science, 125(1): 115-128.

Gobeyn, S. et al., 2017. Impact of the timing of a SAR image acquisition on the calibration of a flood inundation model. Advances in Water Resources, 100: 126-138.

Guidolin, M. et al., 2016. A weighted cellular automata 2D inundation model for rapid flood analysis. Environmental Modelling & Software, 84: 378-394.

Hall, J., Solomatine, D., 2008. A framework for uncertainty analysis in flood risk management decisions. International Journal of River Basin Management, 6(2): 85-98.

Hamill, T.M., 2001. Interpretation of rank histograms for verifying ensemble forecasts. Monthly Weather Review, 129(3): 550-560.

Harmel, R., Cooper, R., Slade, R., Haney, R., Arnold, J., 2006a. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. Transactions-American society of agricultural engineers, 49(3): 689.

Harmel, R., Cooper, R., Slade, R., Haney, R., Arnold, J., 2006b. Cumulative uncertainty in measured streamflow and water quality data for small watersheds.

Hirabayashi, Y. et al., 2013. Global flood risk under climate change. Nature Climate Change, 3(9): 816-821.

Homer, C. et al., 2015. Completion of the 2011 National Land Cover Database for the conterminous United States–representing a decade of land cover change information. Photogrammetric Engineering & Remote Sensing, 81(5): 345-354.

Horritt, M., 2006. A methodology for the validation of uncertain flood inundation models. Journal of Hydrology, 326(1): 153-165.

Horritt, M., Bates, P., 2001a. Effects of spatial resolution on a raster based model of flood flow. Journal of Hydrology, 253(1): 239-249.

Horritt, M., Bates, P., 2001b. Predicting floodplain inundation: raster-based modelling versus the finite-element approach. Hydrological processes, 15(5): 825-842.

Horritt, M., Bates, P., 2002a. Evaluation of 1D and 2D numerical models for predicting river flood inundation. Journal of Hydrology, 268(1): 87-99.

Horritt, M., Bates, P., 2002b. Evaluation of 1D and 2D numerical models for predicting river flood inundation. Journal of hydrology, 268(1-4): 87-99.

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

Huang, S., Hattermann, F.F., 2018. Coupling a global hydrodynamic algorithm and a regional hydrological model for large-scale flood inundation simulations. Hydrology Research, 49(2): 438-449.

Huang, S., Rauberg, J., Apel, H., Disse, M., Lindenschmidt, K.-E., 2007. The effectiveness of polder systems on peak discharge capping of floods along the middle reaches of the Elbe River in Germany. Hydrology and Earth System Sciences Discussions, 11(4): 1391-1401.

Hunter, N. et al., 2008. Benchmarking 2D hydraulic models for urban flood simulations, Proceedings of the Institution of Civil Engineers: Water Management. Thomas Telford (ICE publishing), pp. 13-30.

Jafarzadegan, K., Merwade, V., 2017. A DEM-based approach for large-scale floodplain mapping in ungauged watersheds. Journal of hydrology, 550: 650-662.

Jiang, S. et al., 2017. Quantifying multi-source uncertainties in multi-model predictions using the Bayesian model averaging scheme. Hydrology Research: nh2017272.

Jiang, W., Deng, L., Chen, L., Wu, J., Li, J., 2009. Risk assessment and validation of flood disaster based on fuzzy mathematics. Progress in Natural Science, 19(10): 1419-1425.

Jonkman, S.N., Kelman, I., 2005. An analysis of the causes and circumstances of flood disaster deaths. Disasters, 29(1): 75-97.

Jung, Y., Merwade, V., 2011. Uncertainty quantification in flood inundation mapping using generalized likelihood uncertainty estimate and sensitivity analysis. Journal of Hydrologic Engineering, 17(4): 507-520.

Kalyanapu, A.J., Burian, S.J., McPherson, T.N., 2010. Effect of land use-based surface roughness on hydrologic model output. Journal of Spatial Hydrology, 9(2).

Kay, A., Davies, H., Bell, V., Jones, R., 2009. Comparison of uncertainty sources for climate change impacts: flood frequency in England. Climatic Change, 92(1-2): 41-63.

Knebl, M., Yang, Z.-L., Hutchison, K., Maidment, D., 2005. Regional scale flood modeling using NEXRAD rainfall, GIS, and HEC-HMS/RAS: a case study for the San Antonio River Basin Summer 2002 storm event. Journal of Environmental Management, 75(4): 325-336.

Komi, K., Neal, J., Trigg, M.A., Diekkrüger, B., 2017. Modelling of flood hazard extent in data sparse areas: a case study of the Oti River basin, West Africa. Journal of Hydrology: Regional Studies, 10: 122-132.

Leandro, J., Chen, A.S., Djordjević, S., Savić, D.A., 2009. Comparison of 1D/1D and 1D/2D coupled (sewer/surface) hydraulic models for urban flood simulation. Journal of hydraulic engineering, 135(6): 495-504.

Leskens, J., Brugnach, M., Hoekstra, A., Schuurmans, W., 2014. Why are decisions in flood disaster management so poorly supported by information from flood models? Environmental modelling & software, 53: 53-61.

Liu, Z., Mehran, A., Phillips, T.J., AghaKouchak, A., 2014. Seasonal and regional biases in CMIP5 precipitation simulations. Climate Research, 60(1): 35.

Liu, Z., Merwade, V., Jafarzadegan, K., Investigating the role of model structure and surface roughness in generating flood inundation extents using 1D and 2D hydraulic models. Journal of Flood Risk Management: e12347.

Liu, Z., Merwade, V., Jafarzadegan, K., 2018. Investigating the role of model structure and surface roughness in generating flood inundation extents using 1D and 2D hydraulic models. Journal of Flood Risk Management: e12347.

Ma, Y. et al., 2018. Performance of optimally merged multisatellite precipitation products using the dynamic Bayesian model averaging scheme over the Tibetan Plateau. Journal of Geophysical Research: Atmospheres, 123(2): 814-834.

Madadgar, S., Moradkhani, H., 2014. Improved Bayesian multimodeling: Integration of copulas and Bayesian model averaging. Water Resources Research, 50(12): 9586-9603.

Manfreda, S. et al., 2015. Flood-prone areas assessment using linear binary classifiers based on flood maps obtained from 1D and 2D hydraulic models. Natural Hazards, 79(2): 735-754.

Mark, O., Weesakul, S., Apirumanekul, C., Aroonnet, S.B., Djordjević, S., 2004. Potential and limitations of 1D modelling of urban flooding. Journal of Hydrology, 299(3): 284-299.

Marsooli, R., Orton, P.M., Georgas, N., Blumberg, A.F., 2016. Three-dimensional hydrodynamic modeling of coastal flood mitigation by wetlands. Coastal Engineering, 111: 83-94.

McCuen, R.H., 1989. Hydrologic analysis and design. Prentice-Hall Englewood Cliffs, NJ.

Merlise, A., 1999. Bayesian model averaging and model search strategies. Bayesian statistics, 6: 157.

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

Merwade, V., Rajib, M.A., Liu, Z., 2018. An Integrated Approach for Flood Inundation Modeling on Large Scales, Bridging Science and Policy Implication for Managing Climate Extremes. World Scientific, pp. 133-155.

Merz, B., Kreibich, H., Thieken, A., Schmidtke, R., 2004. Estimation uncertainty of direct monetary flood damage to buildings. Natural Hazards and Earth System Science, 4(1): 153-163.

Merz, B., Thieken, A., Gocht, M., 2007. Flood risk mapping at the local scale: concepts and challenges, Flood risk management in Europe. Springer, pp. 231-251.

Merz, B., Thieken, A.H., 2005. Separating natural and epistemic uncertainty in flood frequency analysis. Journal of Hydrology, 309(1): 114-132.

Merz, B., Thieken, A.H., 2009. Flood risk curves and uncertainty bounds. Natural hazards, 51(3): 437-458.

Morales-Hernández, M., García-Navarro, P., Burguete, J., Brufau, P., 2013. A conservative strategy to couple 1D and 2D models for shallow water flow simulation. Computers & Fluids, 81: 26-44.

Morales, K.H., Ibrahim, J.G., Chen, C.-J., Ryan, L.M., 2006. Bayesian model averaging with applications to benchmark dose estimation for arsenic in drinking water. Journal of the American Statistical Association, 101(473): 9-17.

MRLC, 2011. National Land Cover Dataset (http://www.mrlc.gov/nlcd2011.php).

Mueller, N. et al., 2016. Water observations from space: Mapping surface water from 25years of Landsat imagery across Australia. Remote Sensing of Environment, 174: 341-352.

Mukolwe, M.M., Yan, K., Di Baldassarre, G., Solomatine, D.P., 2016. Testing new sources of topographic data for flood propagation modelling under structural, parameter and observation uncertainty. Hydrological Sciences Journal, 61(9): 1707-1715.

Næss, L.O., Bang, G., Eriksen, S., Vevatne, J., 2005. Institutional adaptation to climate change: flood responses at the municipal level in Norway. Global Environmental Change, 15(2): 125-138.

Najafi, M.R., Moradkhani, H., 2015a. Ensemble combination of seasonal streamflow forecasts. Journal of Hydrologic Engineering, 21(1): 04015043.

Najafi, M.R., Moradkhani, H., 2015b. Multi-model ensemble analysis of runoff extremes for climate change impact assessments. Journal of Hydrology, 525: 352-361.

Neal, J., Schumann, G., Bates, P., 2012a. A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. Water Resources Research, 48(11).

Neal, J. et al., 2012b. How much physical complexity is needed to model flood inundation? Hydrological Processes, 26(15): 2264-2282.

Neuhold, C., Stanzel, P., Nachtnebel, H., 2009. Incorporating river morphological changes to flood risk assessment: uncertainties, methodology and application. Natural Hazards and Earth System Sciences, 9(3): 789-799.

Neuman, S.P., 2003. Maximum likelihood Bayesian averaging of uncertain model predictions. Stochastic Environmental Research and Risk Assessment, 17(5): 291-305.

Nguyen, P., Thorstensen, A., Sorooshian, S., Hsu, K., AghaKouchak, A., 2015. Flood forecasting and inundation mapping using HiResFlood-UCI and near-real-time satellite precipitation data: The 2008 Iowa flood. Journal of Hydrometeorology, 16(3): 1171-1183.

NOAA, 2015. Hydrological Prediction Service (http://water.weather.gov/ahps2).

Nobre, A.D. et al., 2011. Height Above the Nearest Drainage–a hydrologically relevant new terrain model. Journal of Hydrology, 404(1-2): 13-29.

Notaro, V., Fontanazza, C.M., Freni, G., 2014. Bayesian Model Averaging Approach For Reducing Urban Flooding Damage Estimation Uncertainty.

Osborne, J.W., 2010. Improving your data transformations: Applying the Box-Cox transformation. Practical Assessment, Research & Evaluation, 15(12): 1-9.

Papaioannou, G., Vasiliades, L., Loukas, A., Aronica, G.T., 2017. Probabilistic flood inundation mapping at ungauged streams due to roughness coefficient uncertainty in hydraulic modelling. Advances in Geosciences, 44: 23-34.

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

Pappenberger, F., Beven, K., Horritt, M., Blazkova, S., 2005b. 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. et al., 2005c. Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). Hydrology and Earth System Sciences Discussions, 9(4): 381-393.

Pappenberger, F., Matgen, P., Beven, K.J., Henry, J.-B., Pfister, L., 2006a. Influence of uncertain boundary conditions and model structure on flood inundation predictions. Advances in Water Resources, 29(10): 1430-1449.

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

Pappenberger, F. et al., 2013. Visualizing probabilistic flood forecast information: expert preferences and perceptions of best practice in uncertainty communication. Hydrological Processes, 27(1): 132-146.

Pappenberger, F., Thielen, J., Del Medico, M., 2011. The impact of weather forecast improvements on large scale hydrology: analysing a decade of forecasts of the European Flood Alert System. Hydrological Processes, 25(7): 1091-1113.

Parinussa, R.M., Lakshmi, V., Johnson, F.M., Sharma, A., 2016. A new framework for monitoring flood inundation using readily available satellite data. Geophysical Research Letters.

Parrish, M.A., Moradkhani, H., DeChant, C.M., 2012. Toward reduction of model uncertainty: Integration of Bayesian model averaging and data assimilation. Water Resources Research, 48(3).

Patro, S., Chatterjee, C., Mohanty, S., Singh, R., Raghuwanshi, N., 2009. Flood inundation modeling using MIKE FLOOD and remote sensing data. Journal of the Indian Society of Remote Sensing, 37(1): 107-118.

Pourali, S., Arrowsmith, C., Chrisman, N., Matkan, A., Mitchell, D., 2016. Topography wetness index application in flood-risk-based land use planning. Applied Spatial Analysis and Policy, 9(1): 39-54.

Prestininzi, P., Di Baldassarre, G., Schumann, G., Bates, P., 2011. Selecting the appropriate hydraulic model structure using low-resolution satellite imagery. Advances in Water Resources, 34(1): 38-46.

Quiroga, V.M., Kure, S., Udo, K., Mano, A., 2016. Application of 2D numerical simulation for the analysis of the February 2014 Bolivian Amazonia flood: Application of the new HEC-RAS version 5. RIBAGUA-Revista Iberoamericana del Agua.

Raftery, A.E., Gneiting, T., Balabdaoui, F., Polakowski, M., 2005. Using Bayesian model averaging to calibrate forecast ensembles. Monthly weather review, 133(5): 1155-1174.

Raftery, A.E., Madigan, D., Hoeting, J.A., 1997. Bayesian model averaging for linear regression models. Journal of the American Statistical Association, 92(437): 179-191.

Rings, J., Vrugt, J.A., Schoups, G., Huisman, J.A., Vereecken, H., 2012. Bayesian model averaging using particle filtering and Gaussian mixture modeling: Theory, concepts, and simulation experiments. Water Resources Research, 48(5).

Ross, J.L., Ozbek, M.M., Pinder, G.F., 2009. Aleatoric and epistemic uncertainty in groundwater flow and transport simulation. Water Resources Research, 45(12).

Rosser, J.F., Leibovici, D., Jackson, M., 2017. Rapid flood inundation mapping using social media, remote sensing and topographic data. Natural Hazards, 87(1): 103-120.

Saksena, S., Merwade, V., 2015. Incorporating the effect of DEM resolution and accuracy for improved flood inundation mapping. Journal of Hydrology, 530: 180-194.

Sanders, B.F., 2007. Evaluation of on-line DEMs for flood inundation modeling. Advances in Water Resources, 30(8): 1831-1843.

Sangwan, N., Merwade, V., 2015. A Faster and Economical Approach to Floodplain Mapping Using Soil Information. JAWRA Journal of the American Water Resources Association, 51(5): 1286-1304.

Schumann, G. et al., 2016. Unlocking the full potential of Earth observation during the 2015 Texas flood disaster. Water Resources Research, 52(5): 3288-3293.

Schumann, G.P. et al., 2013. A first large-scale flood inundation forecasting model. Water Resources Research, 49(10): 6248-6257.

Şensoy, A., Uysal, G., Şorman, A., 2016. Developing a decision support framework for real-time flood management using integrated models. Journal of Flood Risk Management.

Shamseldin, A.Y., O'Connor, K.M., Liang, G., 1997. Methods for combining the outputs of different rainfall–runoff models. Journal of Hydrology, 197(1): 203-229.

Shields Jr, F.D., Coulton, K.G., Nepf, H., 2017. Representation of Vegetation in Two-Dimensional Hydrodynamic Models. Journal of Hydraulic Engineering, 143(8): 02517002.

Slade, R., 2004. General methods, information, and sources for collecting and analyzing water-resources data. CD-ROM. Copyright.

Spence, A., Poortinga, W., Butler, C., Pidgeon, N.F., 2011. Perceptions of climate change and willingness to save energy related to flood experience. Nature climate change, 1(1): 46-49.

Sun, S., Fu, G., Djordjević, S., Khu, S.-T., 2012. Separating aleatory and epistemic uncertainties: Probabilistic sewer flooding evaluation using probability box. Journal of hydrology, 420: 360-372.

Tavakoly, A., 2015. Hyper-Resolution Large Scale Flood Inundation Modeling: Development of AutoRAPID Model, 2015 AGU Fall Meeting. Agu.

Tayefi, V., Lane, S., Hardy, R., Yu, D., 2007. A comparison of one-and two-dimensional approaches to modelling flood inundation over complex upland floodplains. Hydrological Processes, 21(23): 3190-3202.

Te Chow, V., 1959. Open channel hydraulics.

Teng, J. et al., 2017a. Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. Environmental Modelling & Software, 90: 201-216.

Teng, J. et al., 2017b. Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. Environmental Modelling & Software, 90: 201-216.

Tingsanchali, T., 2012. Urban flood disaster management. Procedia engineering, 32: 25-37.

Tiwari, M.K., Chatterjee, C., 2010. Development of an accurate and reliable hourly flood forecasting model using wavelet–bootstrap–ANN (WBANN) hybrid approach. Journal of Hydrology, 394(3): 458-470.

Todini, E., 2004. Role and treatment of uncertainty in real-time flood forecasting. Hydrological Processes, 18(14): 2743-2746.

Tsai, F.T.C., Elshall, A.S., 2013. Hierarchical Bayesian model averaging for hydrostratigraphic modeling: Uncertainty segregation and comparative evaluation. Water Resources Research, 49(9): 5520-5536.

Tung, Y.K., Mays, L.W., 1981. Optimal risk-based design of flood levee systems. Water Resources Research, 17(4): 843-852.

USACE, 2015. HEC-RAS River Analysis System, Version 5.0, Hydrologic Engineering Center User Manual. U.S. Army Corps of Engineers.

USGS, 1998. Digital Elevation Model (http://ned.usgs.gov/).

USGS, 2008. 2008 floods in Central and Southern Indiana (http://pubs.usgs.gov/of/2008/1322/).

USGS, 2015. Earth Exploer Landsat Archive (http://earthexplorer.usgs.gov/).

Vojinovic, Z., Tutulic, D., 2009. On the use of 1D and coupled 1D-2D modelling approaches for assessment of flood damage in urban areas. Urban water journal, 6(3): 183-199.

Vrugt, J.A., Robinson, B.A., 2007. Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging. Water Resources Research, 43(1).

Wilby, R.L., Beven, K.J., Reynard, N., 2008. Climate change and fluvial flood risk in the UK: more of the same? Hydrological processes, 22(14): 2511-2523.

Wilson, M. et al., 2007. Modeling large-scale inundation of Amazonian seasonally flooded wetlands. Geophysical Research Letters, 34(15).

Wing, O.E. et al., 2018. Estimates of present and future flood risk in the conterminous United States. Environmental Research Letters, 13(3): 034023.

Wintle, B.A., McCarthy, M.A., Volinsky, C.T., Kavanagh, R.P., 2003. The use of Bayesian model averaging to better represent uncertainty in ecological models. Conservation Biology, 17(6): 1579-1590.

Wood, M. et al., 2016a. Calibration of channel depth and friction parameters in the LISFLOOD-FP hydraulic model using medium-resolution SAR data and identifiability techniques. Hydrology and Earth System Sciences, 20(12): 4983.

Wood, M. et al., 2016b. Calibration of channel depth and friction parameters in the LISFLOOD-FP hydraulic model using medium resolution SAR data. Hydrol. Earth Syst. Sci. Discuss., 2016: 1-24.

Wu, C.J., 1983. On the convergence properties of the EM algorithm. The Annals of statistics: 95-103.

Wu, X. et al., 2017. Scenario-based projections of future urban inundation within a coupled hydrodynamic model framework: A case study in Dongguan City, China. Journal of Hydrology, 547: 428-442.

Xiong, L., Shamseldin, A.Y., O'connor, K.M., 2001. A non-linear combination of the forecasts of rainfall-runoff models by the first-order Takagi–Sugeno fuzzy system. Journal of hydrology, 245(1): 196-217.

Yan, H., Moradkhani, H., 2014. Bayesian model averaging for flood frequency analysis, World Environmental and Water Resources Congress 2014, pp. 1886-1895.

Yen, H., 2012. Confronting input, parameter, structural, and measurement uncertainty in multi-site multiple-response watershed modeling using Bayesian inferences, Colorado State University.

Yeung, K.Y., Bumgarner, R.E., Raftery, A.E., 2005. Bayesian model averaging: development of an improved multi-class, gene selection and classification tool for microarray data. Bioinformatics, 21(10): 2394-2402.

Yu, J., Qin, X., Larsen, O., 2013. Joint Monte Carlo and possibilistic simulation for flood damage assessment. Stochastic environmental research and risk assessment, 27(3): 725-735.

Yu, J., Qin, X., Larsen, O., 2015. Uncertainty analysis of flood inundation modelling using GLUE with surrogate models in stochastic sampling. Hydrological Processes, 29(6): 1267-1279.

Zhang, J., Zhou, C., Xu, K., Watanabe, M., 2002. Flood disaster monitoring and evaluation in China. Global Environmental Change Part B: Environmental Hazards, 4(2): 33-43.

Zhang, T., Feng, P., Maksimović, Č., Bates, P.D., 2016a. Application of a three-dimensional unstructured-mesh finite-element flooding model and comparison with two-dimensional approaches. Water Resources Management, 30(2): 823-841.

Zhang, X., Srinivasan, R., Bosch, D., 2009. Calibration and uncertainty analysis of the SWAT model using Genetic Algorithms and Bayesian Model Averaging. Journal of Hydrology, 374(3): 307-317.

Zhang, X., Yan, X., Chen, Z., 2016b. Reconstructed Regional Mean Climate with Bayesian Model Averaging: A Case Study for Temperature Reconstruction in the Yunnan–Guizhou Plateau, China. Journal of Climate, 29(14): 5355-5361.

Zhou, J.G., Causon, D.M., Mingham, C.G., Ingram, D.M., 2004. Numerical prediction of dam-break flows in general geometries with complex bed topography. Journal of Hydraulic Engineering, 130(4): 332-340.

Zhu, G. et al., 2016. Multi-model ensemble prediction of terrestrial evapotranspiration across north China using Bayesian model averaging. Hydrological Processes, 30(16): 2861-2879.

APPENDIX A

Table A.1 Model configurations and their corresponding implementation of uncertainty sources

Model configuration	Channel shape	Channel width	Input flow	Topography	Channel roughness
1	Triangle	5%	0.9USGS	60m	0.01
2	Triangle	5%	USGS	60m	0.01
3	Triangle	5%	1.1USGS	60m	0.01
4	Triangle	5%	0.9USGS	60m	0.03
5	Triangle	5%	USGS	60m	0.03
6	Triangle	5%	1.1USGS	60m	0.03
7	Triangle	5%	0.9USGS	60m	0.05
8	Triangle	5%	USGS	60m	0.05
9	Triangle	5%	1.1USGS	60m	0.05
10	Rectangle	5%	0.9USGS	60m	0.01
11	Rectangle	5%	USGS	60m	0.01
12	Rectangle	5%	1.1USGS	60m	0.01
13	Rectangle	5%	0.9USGS	60m	0.03
14	Rectangle	5%	USGS	60m	0.03
15	Rectangle	5%	1.1USGS	60m	0.03
16	Rectangle	5%	0.9USGS	60m	0.05
17	Rectangle	5%	USGS	60m	0.05
18	Rectangle	5%	1.1USGS	60m	0.05
19	Parabola	5%	0.9USGS	60m	0.01
20	Parabola	5%	USGS	60m	0.01
21	Parabola	5%	1.1USGS	60m	0.01
22	Parabola	5%	0.9USGS	60m	0.03
23	Parabola	5%	USGS	60m	0.03
24	Parabola	5%	1.1USGS	60m	0.03

Table A.1 continued

25	Parabola	5%	0.9USGS	60m	0.05
26	Parabola	5%	USGS	60m	0.05
27	Parabola	5%	1.1USGS	60m	0.05
28	Triangle	Average	0.9USGS	60m	0.01
29	Triangle	Average	USGS	60m	0.01
30	Triangle	Average	1.1USGS	60m	0.01
31	Triangle	Average	0.9USGS	60m	0.03
32	Triangle	Average	USGS	60m	0.03
33	Triangle	Average	1.1USGS	60m	0.03
34	Triangle	Average	0.9USGS	60m	0.05
35	Triangle	Average	USGS	60m	0.05
36	Triangle	Average	1.1USGS	60m	0.05
37	Rectangle	Average	0.9USGS	60m	0.01
38	Rectangle	Average	USGS	60m	0.01
39	Rectangle	Average	1.1USGS	60m	0.01
40	Rectangle	Average	0.9USGS	60m	0.03
41	Rectangle	Average	USGS	60m	0.03
42	Rectangle	Average	1.1USGS	60m	0.03
43	Rectangle	Average	0.9USGS	60m	0.05
44	Rectangle	Average	USGS	60m	0.05
45	Rectangle	Average	1.1USGS	60m	0.05
46	Parabola	Average	0.9USGS	60m	0.01
47	Parabola	Average	USGS	60m	0.01
48	Parabola	Average	1.1USGS	60m	0.01
49	Parabola	Average	0.9USGS	60m	0.03
50	Parabola	Average	USGS	60m	0.03
51	Parabola	Average	1.1USGS	60m	0.03
52	Parabola	Average	0.9USGS	60m	0.05
53	Parabola	Average	USGS	60m	0.05

Table A.1 continued

54	Parabola	Average	1.1USGS	60m	0.05
55	Triangle	95%	0.9USGS	60m	0.01
56	Triangle	95%	USGS	60m	0.01
57	Triangle	95%	1.1USGS	60m	0.01
58	Triangle	95%	0.9USGS	60m	0.03
59	Triangle	95%	USGS	60m	0.03
60	Triangle	95%	1.1USGS	60m	0.03
61	Triangle	95%	0.9USGS	60m	0.05
62	Triangle	95%	USGS	60m	0.05
63	Triangle	95%	1.1USGS	60m	0.05
64	Rectangle	95%	0.9USGS	60m	0.01
65	Rectangle	95%	USGS	60m	0.01
66	Rectangle	95%	1.1USGS	60m	0.01
67	Rectangle	95%	0.9USGS	60m	0.03
68	Rectangle	95%	USGS	60m	0.03
69	Rectangle	95%	1.1USGS	60m	0.03
70	Rectangle	95%	0.9USGS	60m	0.05
71	Rectangle	95%	USGS	60m	0.05
72	Rectangle	95%	1.1USGS	60m	0.05
73	Parabola	95%	0.9USGS	60m	0.01
74	Parabola	95%	USGS	60m	0.01
75	Parabola	95%	1.1USGS	60m	0.01
76	Parabola	95%	0.9USGS	60m	0.03
77	Parabola	95%	USGS	60m	0.03
78	Parabola	95%	1.1USGS	60m	0.03
79	Parabola	95%	0.9USGS	60m	0.05
80	Parabola	95%	USGS	60m	0.05
81	Parabola	95%	1.1USGS	60m	0.05
82	Triangle	5%	0.9USGS	90m	0.01

Table A.1 continued

83	Triangle	5%	USGS	90m	0.01
84	Triangle	5%	1.1USGS	90m	0.01
85	Triangle	5%	0.9USGS	90m	0.03
86	Triangle	5%	USGS	90m	0.03
87	Triangle	5%	1.1USGS	90m	0.03
88	Triangle	5%	0.9USGS	90m	0.05
89	Triangle	5%	USGS	90m	0.05
90	Triangle	5%	1.1USGS	90m	0.05
91	Rectangle	5%	0.9USGS	90m	0.01
92	Rectangle	5%	USGS	90m	0.01
93	Rectangle	5%	1.1USGS	90m	0.01
94	Rectangle	5%	0.9USGS	90m	0.03
95	Rectangle	5%	USGS	90m	0.03
96	Rectangle	5%	1.1USGS	90m	0.03
97	Rectangle	5%	0.9USGS	90m	0.05
98	Rectangle	5%	USGS	90m	0.05
99	Rectangle	5%	1.1USGS	90m	0.05
100	Parabola	5%	0.9USGS	90m	0.01
101	Parabola	5%	USGS	90m	0.01
102	Parabola	5%	1.1USGS	90m	0.01
103	Parabola	5%	0.9USGS	90m	0.03
104	Parabola	5%	USGS	90m	0.03
105	Parabola	5%	1.1USGS	90m	0.03
106	Parabola	5%	0.9USGS	90m	0.05
107	Parabola	5%	USGS	90m	0.05
108	Parabola	5%	1.1USGS	90m	0.05
109	Triangle	Average	0.9USGS	90m	0.01
110	Triangle	Average	USGS	90m	0.01
111	Triangle	Average	1.1USGS	90m	0.01

Table A.1 continued

112	Triangle	Average	0.9USGS	90m	0.03
113	Triangle	Average	USGS	90m	0.03
114	Triangle	Average	1.1USGS	90m	0.03
115	Triangle	Average	0.9USGS	90m	0.05
116	Triangle	Average	USGS	90m	0.05
117	Triangle	Average	1.1USGS	90m	0.05
118	Rectangle	Average	0.9USGS	90m	0.01
119	Rectangle	Average	USGS	90m	0.01
120	Rectangle	Average	1.1USGS	90m	0.01
121	Rectangle	Average	0.9USGS	90m	0.03
122	Rectangle	Average	USGS	90m	0.03
123	Rectangle	Average	1.1USGS	90m	0.03
124	Rectangle	Average	0.9USGS	90m	0.05
125	Rectangle	Average	USGS	90m	0.05
126	Rectangle	Average	1.1USGS	90m	0.05
127	Parabola	Average	0.9USGS	90m	0.01
128	Parabola	Average	USGS	90m	0.01
129	Parabola	Average	1.1USGS	90m	0.01
130	Parabola	Average	0.9USGS	90m	0.03
131	Parabola	Average	USGS	90m	0.03
132	Parabola	Average	1.1USGS	90m	0.03
133	Parabola	Average	0.9USGS	90m	0.05
134	Parabola	Average	USGS	90m	0.05
135	Parabola	Average	1.1USGS	90m	0.05
136	Triangle	95%	0.9USGS	90m	0.01
137	Triangle	95%	USGS	90m	0.01
138	Triangle	95%	1.1USGS	90m	0.01
139	Triangle	95%	0.9USGS	90m	0.03
140	Triangle	95%	USGS	90m	0.03

Table A.1 continued

141	Triangle	95%	1.1USGS	90m	0.03
142	Triangle	95%	0.9USGS	90m	0.05
143	Triangle	95%	USGS	90m	0.05
144	Triangle	95%	1.1USGS	90m	0.05
145	Rectangle	95%	0.9USGS	90m	0.01
146	Rectangle	95%	USGS	90m	0.01
147	Rectangle	95%	1.1USGS	90m	0.01
148	Rectangle	95%	0.9USGS	90m	0.03
149	Rectangle	95%	USGS	90m	0.03
150	Rectangle	95%	1.1USGS	90m	0.03
151	Rectangle	95%	0.9USGS	90m	0.05
152	Rectangle	95%	USGS	90m	0.05
153	Rectangle	95%	1.1USGS	90m	0.05
154	Parabola	95%	0.9USGS	90m	0.01
155	Parabola	95%	USGS	90m	0.01
156	Parabola	95%	1.1USGS	90m	0.01
157	Parabola	95%	0.9USGS	90m	0.03
158	Parabola	95%	USGS	90m	0.03
159	Parabola	95%	1.1USGS	90m	0.03
160	Parabola	95%	0.9USGS	90m	0.05
161	Parabola	95%	USGS	90m	0.05
162	Parabola	95%	1.1USGS	90m	0.05
163	Triangle	5%	0.9USGS	120m	0.01
164	Triangle	5%	USGS	120m	0.01
165	Triangle	5%	1.1USGS	120m	0.01
166	Triangle	5%	0.9USGS	120m	0.03
167	Triangle	5%	USGS	120m	0.03
168	Triangle	5%	1.1USGS	120m	0.03
169	Triangle	5%	0.9USGS	120m	0.05

Table A.1 continued

170	Triangle	5%	USGS	120m	0.05
171	Triangle	5%	1.1USGS	120m	0.05
172	Rectangle	5%	0.9USGS	120m	0.01
173	Rectangle	5%	USGS	120m	0.01
174	Rectangle	5%	1.1USGS	120m	0.01
175	Rectangle	5%	0.9USGS	120m	0.03
176	Rectangle	5%	USGS	120m	0.03
177	Rectangle	5%	1.1USGS	120m	0.03
178	Rectangle	5%	0.9USGS	120m	0.05
179	Rectangle	5%	USGS	120m	0.05
180	Rectangle	5%	1.1USGS	120m	0.05
181	Parabola	5%	0.9USGS	120m	0.01
182	Parabola	5%	USGS	120m	0.01
183	Parabola	5%	1.1USGS	120m	0.01
184	Parabola	5%	0.9USGS	120m	0.03
185	Parabola	5%	USGS	120m	0.03
186	Parabola	5%	1.1USGS	120m	0.03
187	Parabola	5%	0.9USGS	120m	0.05
188	Parabola	5%	USGS	120m	0.05
189	Parabola	5%	1.1USGS	120m	0.05
190	Triangle	Average	0.9USGS	120m	0.01
191	Triangle	Average	USGS	120m	0.01
192	Triangle	Average	1.1USGS	120m	0.01
193	Triangle	Average	0.9USGS	120m	0.03
194	Triangle	Average	USGS	120m	0.03
195	Triangle	Average	1.1USGS	120m	0.03
196	Triangle	Average	0.9USGS	120m	0.05
197	Triangle	Average	USGS	120m	0.05
198	Triangle	Average	1.1USGS	120m	0.05

Table A.1 continued

199	Rectangle	Average	0.9USGS	120m	0.01
200	Rectangle	Average	USGS	120m	0.01
201	Rectangle	Average	1.1USGS	120m	0.01
202	Rectangle	Average	0.9USGS	120m	0.03
203	Rectangle	Average	USGS	120m	0.03
204	Rectangle	Average	1.1USGS	120m	0.03
205	Rectangle	Average	0.9USGS	120m	0.05
206	Rectangle	Average	USGS	120m	0.05
207	Rectangle	Average	1.1USGS	120m	0.05
208	Parabola	Average	0.9USGS	120m	0.01
209	Parabola	Average	USGS	120m	0.01
210	Parabola	Average	1.1USGS	120m	0.01
211	Parabola	Average	0.9USGS	120m	0.03
212	Parabola	Average	USGS	120m	0.03
213	Parabola	Average	1.1USGS	120m	0.03
214	Parabola	Average	0.9USGS	120m	0.05
215	Parabola	Average	USGS	120m	0.05
216	Parabola	Average	1.1USGS	120m	0.05
217	Triangle	95%	0.9USGS	120m	0.01
218	Triangle	95%	USGS	120m	0.01
219	Triangle	95%	1.1USGS	120m	0.01
220	Triangle	95%	0.9USGS	120m	0.03
221	Triangle	95%	USGS	120m	0.03
222	Triangle	95%	1.1USGS	120m	0.03
223	Triangle	95%	0.9USGS	120m	0.05
224	Triangle	95%	USGS	120m	0.05
225	Triangle	95%	1.1USGS	120m	0.05
226	Rectangle	95%	0.9USGS	120m	0.01
227	Rectangle	95%	USGS	120m	0.01

Table A.1 continued

228	Rectangle	95%	1.1USGS	120m	0.01
229	Rectangle	95%	0.9USGS	120m	0.03
230	Rectangle	95%	USGS	120m	0.03
231	Rectangle	95%	1.1USGS	120m	0.03
232	Rectangle	95%	0.9USGS	120m	0.05
233	Rectangle	95%	USGS	120m	0.05
234	Rectangle	95%	1.1USGS	120m	0.05
235	Parabola	95%	0.9USGS	120m	0.01
236	Parabola	95%	USGS	120m	0.01
237	Parabola	95%	1.1USGS	120m	0.01
238	Parabola	95%	0.9USGS	120m	0.03
239	Parabola	95%	USGS	120m	0.03
240	Parabola	95%	1.1USGS	120m	0.03
241	Parabola	95%	0.9USGS	120m	0.05
242	Parabola	95%	USGS	120m	0.05
243	Parabola	95%	1.1USGS	120m	0.05

*5%, average, 95% represents the threshold values of sampled channel widths

VITA

Zhu Liu was born in Xi'an, China. He graduated with a B.S in Department of Hydro-power Engineering in 2012 from Xi'an University of Technology, Xi'an, China. He received his M.S degree in Civil Engineering from University of California-Irvine, Irvine, USA in 2013. He joined the graduate program in Civil Engineering at Purdue University in 2014 and received Doctor of Philosophy degree in August, 2018.