UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

CHARACTERIZING UNCERTAINTY OF A HYDROLOGIC MODELING SYSTEM FOR OPERATIONAL FLOOD FORECASTING OVER THE CONTERMINOUS UNITED STATES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

HUMBERTO JOSE VERGARA ARRIETA
Norman, Oklahoma
2015

CHARACTERIZING UNCERTAINTY OF A HYDROLOGIC MODELING SYSTEM FOR OPERATIONAL FLOOD FORECASTING OVER THE CONTERMINOUS UNITED STATES

A DISSERTATION APPROVED FOR THE
SCHOOL OF CIVIL ENGINEERING AND ENVIRONMENTAL SCIENCE

BY

_____
Dr. Yang Hong, Chair

_____
Dr. Jonathan Gourley

_____
Dr. Randall Kolar

_____
Dr. Robert Knox

_____
Dr. Xuguang Wang

I dedicate this work to my family. To my wife Diana and to my children Nicolas, Sara and Samuel, for their love, support and patience. To my parents and to my parents-in-law for their constant support and help.

Acknowledgements

I would like to thank Dr. Yang Hong, my academic advisor from the school of Civil Engineering and Environmental Science (CEES), for his encouragement during the past four years of the Doctoral degree program. Likewise, thanks to Dr. Jonathan Gourley, from the National Severe Storms Laboratory (NSSL) for his advising and support of my work with the Flooded Locations And Simulated Hydrographs (FLASH) project. Also, thanks to the other members of the committee Dr. Randall Kolar, Dr. Robert Knox and Dr. Xuguang Wang for their contribution throughout the development of this work. Special thanks to Dr. Pierre Kirstetter for all his contribution and mentoring during the development of this work.

Some of the computing for this project was performed at the OU Supercomputing Center for Education & Research (OSCER) at the University of Oklahoma (OU). The Disaster Relief Appropriations Act of 2013 (P.L. 113-2) supported this work, which funded NOAA research grant NA14OAR4830100. Partial support was provided by the Hydrometeorology and Remote Sensing (HyDROS) Lab at the National Weather Center, Norman, OK.

Table of Contents

List of Tables

Table of Figures

Abstract

The purpose of this work was to study the macro scale patterns of simulated streamflow errors in order to characterize uncertainty in a hydrologic modeling system and establish the basis for a probabilistic forecasting framework. The particular application of this endeavor is on flood and flash flood forecasting in an operational context. The hydrologic modeling system has been implemented at 1-km/5-min resolution to generate estimates of streamflow over the Conterminous United States (CONUS). The parameterization of the hydrologic model was prepared using spatially distributed information on soil characteristics, land cover/land use, and topography alone. An innovative method to estimate parameter values for the physics-based flow routing model was developed for the purpose of this research. Unlike the standard practice in hydrologic modeling exercises, no calibration of the hydrologic model was performed following its initial configuration. This calibration-free approach guarantees the spatiotemporal consistency of uncertainty and model biases, which is key for the methodology explored herein.

Data from the CONUS-wide stream gauge network of the United States' Geological Survey (USGS) were used as a reference to evaluate the discrepancies with the hydrological model predictions. Only stream gauges with drainages less than or equal to 1,000 km^2 were employed. Streamflow errors were studied at the event scale with particular focus on the peak flow magnitude and timing. A total of 2,680 catchments and 75,496 events were used for the error analysis. A methodology based on automatic processing algorithms was developed to deal with this large sample for model diagnostics.

Associations between streamflow errors and geophysical factors were explored and modeled. It was found that hydro-climatic factors and radar coverage could explain significant underestimation of peak flow in regions of complex terrain. Furthermore, the statistical modeling of peak flow errors showed that other geophysical factors such as basin geomorphometry and pedology could also provide explanatory information. Results from this research demonstrate the potential of uncertainty characterization in providing feedback for model improvement and its utility in enabling probabilistic flood forecasting that can be extended to ungauged locations.

Chapter 1. Introductory Aspects of the Research

1.1 Introduction

Forecasting the future state of a hydrologic system, regardless of the spatio-temporal scale, is a central objective of hydrologic science and engineering. This aspect is of particular interest for risk management applications, such as in flood early warning systems, where information obtained from forecasts is used in decision-making operations that attempt to prevent loss of lives and property. In the classical paradigm of hydrology, forecasts are produced deterministically using a modeling system with particular modeler- and problem-defined settings. By definition, deterministic frameworks assume that there is a unique and sufficient solution resulting from the particular settings defining them. However, it is now widely accepted that deterministic frameworks have limitations in producing forecasts with acceptable skill in decision-making situations. The underlying limitation of deterministic frameworks is the negligence of uncertainty. The necessity to account for uncertainty in the forecasting system started to receive attention over two decades ago (Juston et al. 2013), and although it is now acknowledged as an essential aspect in any modeling exercise (Liu et al. 2012; Montanari and Koutsoyiannis 2012), a complete change in the hydrologic forecasting paradigm is yet to be seen as many challenges need to be overcome (Pappenberger and Beven 2006).

One particular challenge is in determining how to produce an adequate qualitative and quantitative description of uncertainty in the forecasts (Hrachowitz et al. 2013). This is a non-trivial task mainly because there are no observations that can be

used as reference to characterize uncertainty for some of the hydrologic processes occurring in a watershed (e.g. the amount of water infiltrated after rainfall occurs). Other difficulties associated with this challenge come from the entanglement resulting from the multiple sources of uncertainty in a modeling system, whose interaction and interdependencies are complex due to the highly non-linear dynamics of the hydrologic processes. Moreover, associating the impact of uncertainty with particular sources as it manifests in forecast errors is challenging. The latter is exacerbated with the use of aggregated metrics (e.g. sum or mean of squared errors) that reduce the dimensionality of errors (Yilmaz et al. 2008), in an attempt to simplify the quantitative evaluation of model performance. Gupta et al. (2008) have discussed this issue and the need for evaluation strategies with diagnostic power: this is, "*a well-considered approach to reconciling environmental theory with observations"* able to describe how uncertainties in the modeling system affect the forecasts and where in the modeling system these uncertainties might reside.

The issues discussed above are particularly challenging in problems involving predictions at ungauged locations. At unobserved times and locations, the only available information about the state of the hydrologic system comes from model estimates. Naturally, it becomes essential to have not only a good estimate of what the most probable state of the system will be (e.g., the mean), but also a way to forecast the error characteristics (e.g., the variance) of those estimates. To this end, an error characterization framework needs to be developed based on available information at observed locations and be designed to extend to ungauged areas.

A *large catchment sample* approach is the basis of the development of the uncertainty characterization framework. This kind of approach is possible today, not only because of the advances of computing resources and data management, but because the knowledge in hydrologic prediction has been constructed by many in-depth analysis type of studies (see for example, Schumm 1956; Nash 1957; Mockus 1961; Nash and Sutcliffe 1970; Burnash et al. 1973; Williams 1978; Pettyjohn and Henning 1979; Cosby et al. 1984; Ponce 1986; Costa 1987; Chow et al. 1988). An important aspect of the methodology of this research is finding a balance between depth and breadth of information (Gupta et al. 2013).

This research is conducted in the context of the Flooded Locations And Simulated Hydrographs (FLASH; http://blog.nssl.noaa.gov/flash/) project, whose main purpose is to provide flash flood warnings in the United States with improved accuracy, timing, and specificity, so as to mitigate the impacts on society. The FLASH project seeks to make a shift in the classical paradigm of operational flash flood forecasting in the United States through direct forward simulation of the hydrometerological phenomena at unprecedented spatiotemporal resolutions in a probabilistic framework. The present work aims to establish the foundation of the probabilistic forecasting approach in FLASH based on a hydrologic modeling system error analysis at the Conterminous United States scale. The methodology will enable FLASH's modeling system to produce forecasts with uncertainty information at all locations over the Conterminous United States, regardless of whether observations are available or not.

1.2 Why does uncertainty need to be characterized?

It is widely accepted in scientific hydrology that there is no single representation of the physical system able to reproduce all hydrologic processes for all possible conditions and at all meaningful scales. Evidence of this fact can be found in the literature, which has been overwhelmed with a diversity of models varying in complexity, degree of physical consistency and applicability (for a detailed review see Kampf and Burges 2007). Moreover, there is an ongoing debate on whether hydrologic models should be developed strictly based on physical laws and how the appropriateness of model structure needs to be defined. Beven (2006), Clark et al. (2011) and Gupta et al. (2012) have discussed these issues from a philosophical viewpoint, raising concerns about the direction hydrological sciences have been going, and proposing new approaches based on the "scientific" method, as opposed to more traditional approaches based on the "engineering" method.

It is generally recognized that there is uncertainty in almost every component of hydrologic models, and much research has been devoted to their particular impact on hydrologic forecasts. Most noticeably, research can be found focusing on model inputs (Carpenter and Georgakakos 2004; Gourley and Vieux 2005; Hong et al. 2006; Wu et al. 2011; Kirstetter et al. 2010; Kirstetter et al. 2012; Kirstetter et al. 2015), model parameters (Beven and Binley 1992; Duan et al. 2003; Vrugt et al. 2005; Vrugt et al. 2008), model structure (Georgakakos et al. 2004; Carpenter and Georgakakos 2006; Clark et al. 2011; Gupta et al. 2012), and streamflow observations (Petersen-Overleir 2004; Di Baldassarre and Montanari 2009; McMillan et al. 2010). Uncertainty resulting from its multiple sources manifests in errors of the model output with high complexity

in their structure. As a result, it is sometimes difficult to discern random uncertainty from epistemic uncertainty. Through uncertainty characterization, it is possible to explain limitations in the modeling system that can be reduced, and separate randomness so it can be modeled through a probabilistic approach (Beven 2013).

1.3 Floods, flash floods, and operational hydrologic forecasting system in the United States

Flooding has been the subject of intensive research efforts because these hydrometeorological phenomena are considered to be the ones that produce the most devastating effects on lives and infrastructure on a global scale. On average, floods cause more than 20,000 deaths and adversely affect about 140 million people yearly over the globe (Adhikari et al. 2010). Forecasting this natural hazard is a challenging task especially in regions where in-situ observations are sparse or non-existent. A particular type of flooding that represents a challenge for prediction in the hydrometeorological field is flash flooding. A flash flood is a rapid flooding of water over land or in a stream that results from heavy rainfall or a sudden release of impounded water from a logjam or dam (Hong et al. 2010). Flash floods differ from fluvial floods due to their quick response to rainfall (within a few hours) and to the fact that they generally occur in small watersheds, such as headwater basins. The biggest challenge of flash floods is their short window for warning.

In the United States, flash floods cause about 100 deaths per year and are considered a top weather-related hazard (Ashley and Ashley 2008; Gourley et al. 2012; Clark et al. 2014). Although the prediction of these hydrometeorological hazards is a

subject of great interest, the forecasting tools used in operational systems in the United States have not been significantly upgraded in over forty years (Clark et al. 2014). Much development in terms of the knowledge of the geophysical hazard, availability of remote-sensing platforms, such as satellites and radars, geographical information systems, and computational resources has occurred in the past two or three decades and, thus, new forecasting methodologies are warranted for operational implementation (Gourley et al. 2014).

1.4 Scientific problem and research objectives

1.4.1 Hypotheses

Hypothesis 1: If simulated streamflow errors are correlated with at least one of the properties of basins (namely geomorphometry, hydro-climatic regime, soil and land cover/use), then at least one of those properties could be used to explain the spatial variability of simulated streamflow errors.

Hypothesis 2: If *Hypothesis 1* is true and if the association between simulated streamflow errors and the property (or properties) of the basin can be modeled, then simulated streamflow errors could be predicted in basins whose property (or properties) is (or are) encompassed in the training dataset.

1.4.2 Objectives

This work seeks to establish the foundation for the development of a framework based on a synergy between physical and stochastic dynamics for improved operational

flood forecasting over the Conterminous United States. Specific objectives in this research are listed as follows:

- To configure a hydrologic modeling system based on geomorphological parameters, hydro-climatic regime, soil properties, and land cover/use for flood and flash flood forecasting over the Conterminous United States;
- To develop an evaluation approach that describes uncertainty in the hydrologic modeling system through simulated streamflow errors;
- To investigate descriptors of the spatial variability of simulated streamflow errors for uncertainty modeling over the Conterminous United States.

1.5 Dissertation structure

The research work included in this dissertation consists of four studies that have been part of the development of the hydrologic modeling approach in FLASH and that establish fundamental elements of the design of an innovative probabilistic methodology for flood forecasting. Specifically, this work presents research to devise a framework for explicit descriptions of uncertainty in hydrologic forecasts and its macro scale patterns of variability. The progressive development of such a framework is presented in this dissertation as described below.

Chapter 2 presents a study that illustrates the concept of hydrologic uncertainty characterization, entitled "*Modeling and Disentangling Uncertainty in Streamflow Forecasts.*" Moreover, the work therein demonstrates the usefulness and importance of characterizing uncertainty for improved hydrologic forecasts. The case study establishes important elements of the skill diagnostic strategy that specifically respond to the

second objective of this dissertation. Additionally, it demonstrates the implementation of data assimilation and its natural linkage with uncertainty description. Data assimilation has been acknowledged as an important methodology for operational forecasting systems, whose application in such settings still requires research efforts.

Chapter 3 is entitled "*Configuration of a Distributed Hydrologic Model for Streamflow Simulation over the Conterminous United States.*" This chapter describes the hydrologic processes believed to be relevant for the modeling of streamflow. It also presents the conceptualization of these processes in a mathematical model. Moreover, elements of the configuration of the hydrologic model over the Conterminous United States are discussed. Consideration of these aspects is important for the characterization of uncertainty because of the inevitable limitations in the perceptual and conceptual descriptions of the physical system. Likewise, choices in the overall modeling strategy can significantly contribute to uncertainty and impact the skill of hydrologic forecasts.

Chapter 4 is entitled "*Estimating a-priori Flow Routing Parameters for Streamflow Simulation over the Conterminous United States.*" It elaborates on the parameter estimation aspect of the work presented in Chapter 3. Specifically, an innovative approach to *a-priori* estimation of flow routing model parameters based on physical theory and multi-dimensional statistical modeling is presented. A key aspect of the approach is the regionalization of information available only at gauged locations. Furthermore, the regionalization represents one of the core concepts of the framework devised herein for the description of the macro scale variability of hydrologic forecasting uncertainty. The work on Chapters 3 and 4 responds to the first objective of this dissertation.

Lastly, Chapter 5 presents the fourth study entitled "*Modeling the Macro Scale Characteristics of Peak Flow Error for Probabilistic Flood Forecasting.*" The work therein responds to the third objective and specifically addresses the hypotheses of this dissertation. It consists of an evaluation of the simulation skill of the hydrologic modeling system described in Chapters 3 and 4, using methodological elements of the diagnostic approach described in Chapter 2. More importantly, a series of experiments is presented that explore the association between simulation errors and geophysical factors. The work concludes with a statistical modeling exercise that demonstrates the explanatory power of the geophysical factors for uncertainty characterization.

Chapter 2. Modeling and Disentangling Uncertainty in Streamflow Forecasts

2.1 Introduction

In Chapter 1, it was discussed that a change from the deterministic paradigm to a probabilistic one requires addressing numerous challenges. Among them, the description of uncertainty in hydrologic forecasts is arguably one of the most important ones. This is because characterizing uncertainty involves numerous difficulties associated with its nature, the non-linearity of the hydrologic system, and the current methodologies to quantify inaccuracies and their variability. A further difficulty is associated with the assessment of the uncertainty characterization scheme. Intuitively, direct evaluation of an uncertainty characterization cannot be performed because there is no reference (i.e., truth). Assessment of the correctness of the characterization of uncertainty is arguably not necessary. Instead, assessing uncertainty characterization's usefulness is key for its implementation and purpose. A modeling technique that capitalizes on the characterization of uncertainty is data assimilation. Data assimilation uses statistical information of the error to optimally combine observations and model forecasts to provide the best estimate of the system's states and thus improve subsequent forecasts (Evensen 1992; Clark et al. 2008). Various studies have shown that the success of data assimilation is contingent on the description of the predictive uncertainty. For example, Reichle et al. (2008) indicated that poor specifications of error inputs can have detrimental effects in the assimilation of soil moisture. Maggioni et al. (2012) found that an elaborate description of the error in satellite quantitative

precipitation estimates can improve the hydrologic estimates from assimilation. Moreover, Liu et al. (2012) point out that although data assimilation is becoming a standard tool in operational forecast systems, it is not technically ready partially because of the lack of methodologies to properly quantify uncertainty. The fact that data assimilation represents an essential component in hydrologic forecasting systems and that it relies on how uncertainty is represented, postulates it as an appropriate objective diagnostic tool for assessing the usefulness of a given characterization scheme.

This chapter is included herein as a demonstration exercise of the usefulness in characterizing uncertainty for improved hydrologic forecasts. Particularly, the aim of this work was to address the challenge of assessing uncertainty characterization with an approach able to disentangle the individual impact that uncertainty in rainfall inputs and in hydrologic model parameters has on simulations of streamflow for flood forecasting. The approach is based on a heuristic methodology involving sequential data assimilation and error metrics that are commensurate with the hydrologic phenomena of interest. Results from this study are intended to provide insights on how to describe uncertainty that results from multiple sources in the hydrologic modeling system, and that is entangled in state variables used to diagnose the skill of the forecasts. Likewise, it might be possible to learn how the usefulness of the characterization of uncertainty can be assessed and how multiple sources of uncertainty can be modeled. Additionally, this work can offer general guidance on how uncertainty needs to be defined for data assimilation systems, which has been noted to lack in the scientific literature.

2.2 Characterizing uncertainty in the hydrologic modeling system

An ensemble forecasting approach was employed in this work for the characterization of uncertainty in hydrologic model forecasts. A major goal of ensemble forecasting is to estimate the flow-dependent uncertainties of the forecast (Hamill 2006) and the true probability distribution of future states of the system (Gourley and Vieux 2006). To quantify the uncertainty, ensembles are built by taking samples from the assumed error Probability Density Function (PDF) and running the model forward in time (Liu and Gupta 2007). Statistical information can then be computed from the ensemble to characterize uncertainty in the forecast.

Creating ensembles involves perturbation of the modeling system's components to account for the variability induced by uncertainty from multiple sources. In principle, each individual component in the modeling system can be perturbed in an attempt to represent the multivariate nature of uncertainty. However, difficulties arise when trying to explicitly characterize uncertainty from sources that have no reference information to compare against (e.g. there is no "true model structure" that can be used to characterize uncertainty in models whose structure deviates from it; Gupta et al. 2012). Additionally, characterizing the entanglement of uncertainty caused by interdependencies among modeling components may represent a challenge. These aspects need to be considered when designing ensemble-based systems for a particular modeling problem. In this study, the design of ensembles is simplified by only targeting sources of uncertainty known to be significant for the specific modeling objective. The rationale of this approach is elaborated in the following sections.

2.2.1 Study area

The Tar River basin in coastal North Carolina was selected as the area of study in this work. The basin is periodically affected by heavy rainfall from tropical storms and hurricanes, at which time major flood events occur. Specifically, the hydrologic modeling focused on the catchment of the US Geological Survey (USGS) stream gauge station located at Tarboro (USGS 02083500; Figure 2.1), which includes the upper Tar River and Fishing Creek sub-basins. Streamflow observations at this location are available at sub-hourly time intervals. The catchment has a drainage area of 5,653 km^2 and is located on the coastal plain. The Tar is a perennial river with a mean daily flow value of about 62 m^3/s at Tarboro. The minimum recorded daily flow is 0.79 m^3/s with a maximum of 1,996 m^3/s, which occurred after the landfall of hurricanes Dennis and Floyd in September of 1999. Flooding and subsequent catastrophic events caused by these two storms have been the subject of several studies and prompted research efforts on this area (e.g. Colby et al. 2000; Bales 2003; Van Cooten et al. 2011; Dresback et al. 2013). The analysis in this study focused on a time period of 3 years (11/21/2003 through 11/21/2006). Catchment response was studied at the hourly scale, which was deemed appropriate, given the size of the basin and the observed overall duration of flow events (on the order of few days).

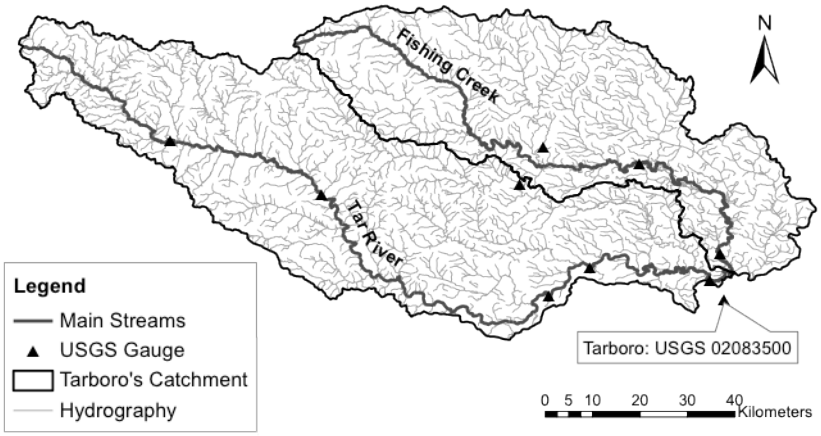

Figure 2.1: Study area showing the hydrography and USGS streamflow gauge stations.

2.2.2 Uncertainty in satellite-based quantitative precipitation estimates

Satellite-based estimates of rainfall have found widespread applications in hydrology because of its availability over the globe. However, there are still many challenges involved in their use because of the inherent uncertainties due to the indirect nature of the measuring technique (Gourley et al. 2010). The satellite-based QPE product used in this work was the Tropical Rainfall Measuring Mission (TRMM)'s near real-time infrared and microwave merged precipitation product 3B42RT (version 7). The 3B42RT estimates are available at the resolution of 25-km and 3-hour time intervals (Huffman et al. 2007). The 3B42RT combines passive microwave low earth orbit (LEO) precipitation estimates and infrared precipitation products from geostationary satellites used to fill the LEO gaps in the 3-hourly times intervals. The QPE product used herein as ground reference was the U.S. National Weather Service (NWS) Multisensor Precipitation Estimation (MPE) product. MPE combines information from satellite, radar, and rain gauges at a resolution of 4-km and 1-hour (Briedenbach and Bradberry 2001; Fulton 2002; Seo et al. 2010). Both 3B42RT and MPE estimates were available for the 3-year period of study. To account for the effects

of the differences in resolution (Vergara et al. 2014), MPE estimates were resampled to accommodate 3B42RT's 25-km/3-hour resolution.

The uncertainties associated to satellite estimates suffer from the variable detection performances and include systematic errors, as well as random effects propagating from merging infrared and microwave passive estimates (Maggioni et al. 2014). In this study the focus was on the impact of random errors on hydrological modeling. It assumes (conditionally) unbiased precipitation estimates. Similarly to Kirstetter et al. (2012; 2015), the transfer between 3B42RT and MPE is analyzed through conditional distribution functions. The sets of MPE distributions given a 3B42RT estimate are studied using the generalized additive models for location, scale, and shape (GAMLSS) technique. Only pairs for which MPE and 3B42RT are both nonzero are considered in the calculations. The errors associated to detection limitations of the satellite product were not considered.

GAMLSS aim at modeling the parameters of a response variable's distribution. Two main assumptions were made: 1) the response variable MPE is a random variable following a known parametric distribution with density f conditional on the parameters (μ, σ), and 2) the observations MPE are mutually independent given the parameter vectors (μ, σ). Each parameter is modeled as a function of 3B42RT (the explanatory variable) using monotonic (linear/nonlinear or smooth) link functions. More details are provided by Rigby and Stasinopoulos (2001; 2005), Akantziliotou et al. (2002) and Stasinopoulos and Rigby (2007). A wide variety of distributional forms are available within GAMLSS. A number of conditional two-parameter density functions (lognormal, normal, reverse gumbel, logistic, gamma, etc.) were tested to fit the data. The

distributions of MPE (not shown here) were generally found to be unimodal and asymmetric. The goodness of fit on the whole dataset was checked with the Akaike Information Criteria (AIC) for each of the semiparametric density fits. The lognormal distribution was found to be the most appropriate:

$$f(R_{ref} \mid \mu,\sigma) = \frac{1}{R_{ref}\sqrt{2\pi}\sigma} e^{-\frac{\left(\ln R_{ref}-\mu\right)^2}{2\sigma^2}} \tag{2.1}$$

The function above was used to model the conditional MPE distributions, where the location μ is linked to the expected MPE value, and the scale σ is representative of random errors. The rainfall trends for each parameter are fitted using locally weighted scatterplot smoothing (LOESS), which are more flexible than polynomials or fractional polynomials for modeling complex nonlinear relationships. It is a polynomial curve determined by 3B42RT fitted locally by weighted polynomial regression (see Cleveland et al. 1992). This ensures to convert 3B42RT estimates into (conditionally) unbiased precipitation estimates. For a given conditional distribution of the response variable MPE, the conditional quantiles can be expressed as a function of 3B42RT. Figure 2.2 shows the quantiles of the fitted MPE distribution models as a function of 3B42RT. It can be observed that the conditional PDFs of MPE present a high conditional spread due to the high bias in 3B42RT estimates.

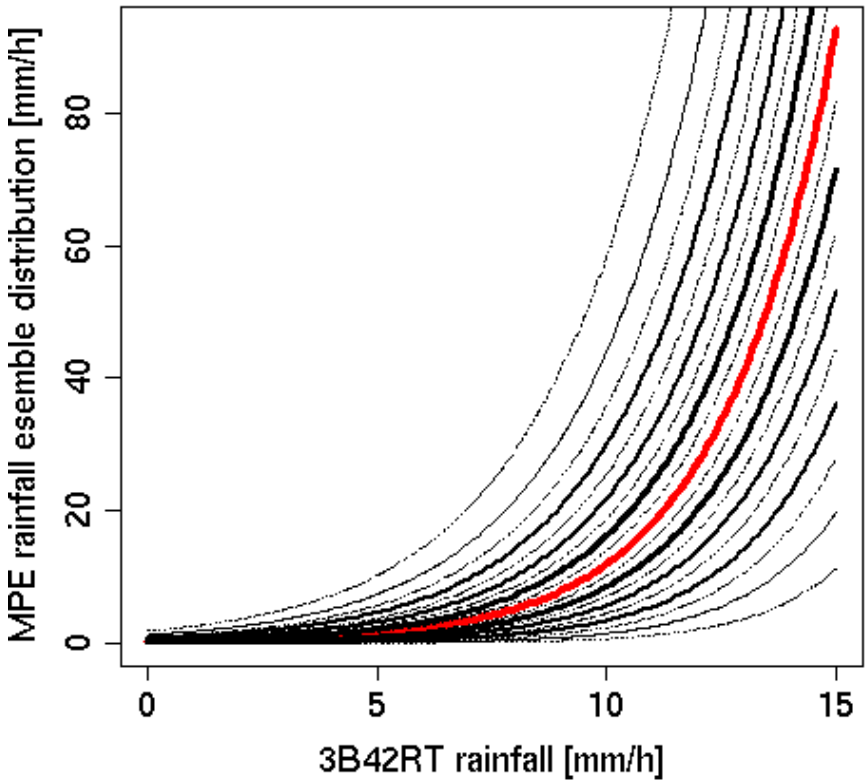

Figure 2.2: Probabilistic Quantitative Precipitation Estimates model. The different curves represent different quantiles of the conditional distributions. The median is highlighted with the red line.

2.2.3 Uncertainty in hydrologic model parameters

The hydrologic model employed in this study is a parsimonious rainfall-runoff model called HyMOD, a subject of several hydrologic model calibration studies (e.g. Wagener et al. 2001; Vrugt and Bouten 2003; Vrugt et al. 2008) and research on hydrological data assimilation (e.g. Moradkhani et al. 2005; Vrugt et al. 2005; Vrugt and Robinson 2007; Smith et al. 2008). HyMOD is based on the Probability Distribution Model (PDM) developed by Moore (1985) and the Nash cascade of linear reservoirs (Nash 1957). Figure 2.3a presents schematics of these two modeling components. The PDM model describes the generation of excess rainfall (*ER*) depending on the current moisture content *C* at time *t* in the catchment, the precipitation *u* and the evapotranspiration *ET*. The model assumes that *C*(*t*) varies across the

catchment. This spatial variability of soil moisture capacity is described by a storage capacity distribution function (Wagener et al. 2001; Smith et al. 2008). ER is then partitioned into subsurface and surface runoff and routed through two series of Nash linear tanks denoting slow and quick flow, respectively. Streamflow at the outlet is obtained by arithmetic aggregation of the two flow components.

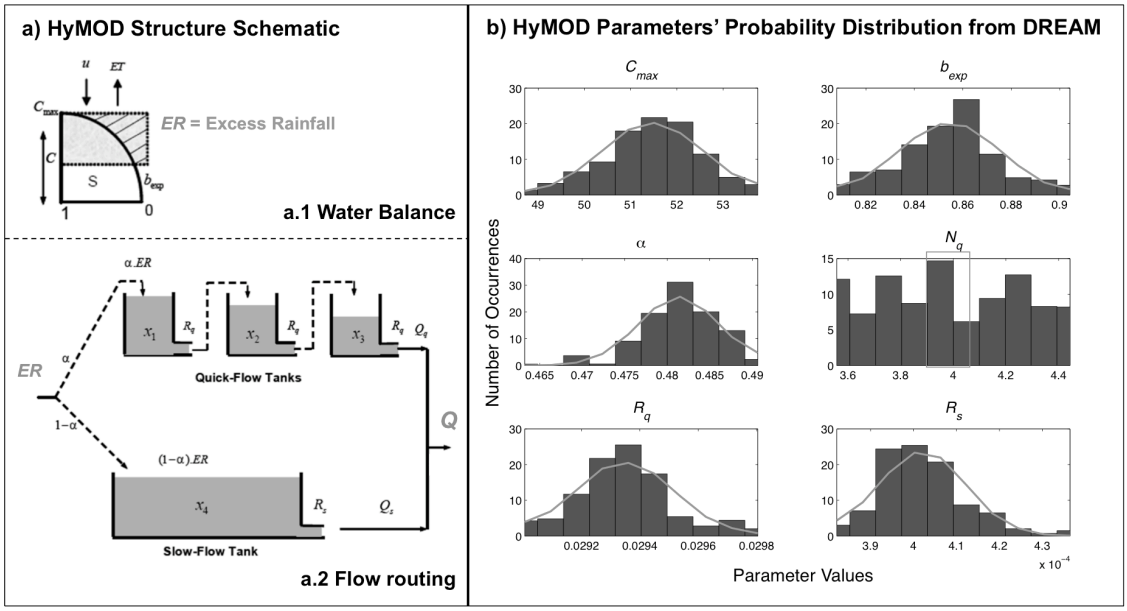

Figure 2.3: HyMOD structure and estimated parameter distributions. a) HyMOD structure schematic presenting the water balance component (a.1) and flow routing (a.2) (adapted from Moradkhani et al. 2005); and b) Parameter distributions generated from a 3-year calibration using DREAM. Lines over the bar graphs represent a fit to a Gaussian distribution.

The abstract conceptualization of hydrologic processes in HyMOD makes it difficult to derive parameters based on physical characteristics of the hydrologic system (e.g. soil properties, which are heterogeneous across a basin). Therefore, model calibration is usually required to estimate HyMOD's parameters. A consequence of this approach is that uncertainty in the estimates, which primarily stems from characteristics of the fitting data and the compensation of inaccuracies in other components of the modeling system (e.g. model structure and input data), is of multivariate nature due to

interdependencies artificially created in the optimization process. Accordingly, the method for uncertainty characterization employed herein was based on a joint probability distribution of HyMOD's parameters derived from a model calibration. HyMOD was calibrated using the DiffeRential Evolution Adaptive Metropolis (DREAM) automatic optimization algorithm presented by Vrugt et al. (2009), which uses a formal Bayesian approach to estimate the posterior probability distribution of model parameters. Figure 2.3b shows the univariate conditional probability distributions of a HyMOD's parameters subset corresponding to the 30^{th} percentile of the optimized objective function. Notice that the distribution of the number of quick flow tanks parameter (N_q) appears to be uniform and has values in the interval $3.5 < N_q < 4.5$. Intuitively, this particular parameter can only take values of non-zero cardinal numbers. However, N_q was calibrated in DREAM considering a continuous range of real numbers between 1 and 10. Other studies featuring HyMOD have fixed N_q to a value of 3. In this study, N_q is allowed to acquire different values in order to represent the variability of flow timing with an extra degree of freedom (i.e. additional to k_o, which is also intended to represent flow delay to the outlet).

The first two moments of the joint probability distribution described above were used to produce ensembles of HyMOD's parameters employing a multi-dimensional random number generator, as described by Wang and Bishop (2005). The first moment of the joint probability distribution corresponds to the parameter set found to be optimal from the calibration process, while the second moment corresponds to the covariance of the parameter sets included in the subset defining the univariate distributions. To model

the variability around the first moment, an n X n covariance matrix Σ_{par} of the deviations of all considered parameter sets from the optimal set was constructed:

$$\Sigma_{par} = \mathrm{cov}\left(\begin{bmatrix} \chi' - \chi_1 \\ \vdots \\ \vdots \\ \chi' - \chi_m \end{bmatrix}_{m \times n} \right) \tag{2.2}$$

where n is the number of model parameters, m is the number of set of model parameters, χ' is the optimal parameter set, and χ_i is the i^{th} parameter set of the considered subset of m parameter sets. The covariance matrix Σ_{par} needs to be decomposed into its eigenvalues and eigenvectors:

$$\Sigma_{par} = \left\langle \varepsilon \varepsilon^T \right\rangle = E \Omega E^T \tag{2.3}$$

where ε are the perturbations for ensemble generation, the columns of E contain the eigenvectors and the diagonal of Ω the corresponding eigenvalues. The perturbations ε are obtained through:

$$\varepsilon = x_1 e_1 + x_2 e_2 + ... + x_n e_n \tag{2.4}$$

where e_i, is the i^{th} eigenvector corresponding to the i^{th} eigenvalue ω_i, and x_i is the i^{th} univariate random value. These random values are generated using parameterized normal distributions with mean equal to zero and variance equal to the i^{th} eigenvalue of Σ_{par}:

$$x_i \sim N\left(0, \sqrt{\omega_i}\right) \tag{2.5}$$

The normality assumption of random noise in (2.5) reasonably fits the univariate distributions of HyMOD's parameters shown in Fig. 1b, with the exception of

parameter N_q. Examination of the marginal probability distribution of parameter N_q, however, reveals a shape that approximates a Gaussian PDF (not shown here). Therefore, parameter N_q univariate variability is also modeled using normal random noise.

Wang and Bishop (2005) employed the method described above in a methodology to improve reliability of under dispersive ensembles. In their methodology, only the directions where ensembles show under dispersion (i.e. in the direction of eigenvectors whose corresponding eigenvalues are positive) are perturbed. Eigenvectors in the direction of over dispersion (i.e. negative eigenvalues) are not perturbed because perturbations would make the ensemble more over dispersive. In the case worked herein, HyMOD's parameters deviations covariance matrix displayed under dispersion in all directions (i.e., all eigenvalues are positive) and, thus, perturbations were applied to all eigenvectors.

2.2.4 Streamflow data assimilation

Assimilation algorithm considerations

Because streamflow results from the space and time integration of the different hydrological processes occurring over the drainage upstream a particular location (e.g., basin outlet), a technique based on smoothing should arguably be the most adequate method when assimilating flow observations. Smoothing methods adjust model states for a window of space and time in order to improve model integration trajectory. However, this requires iterative integration of the model, which represents a very expensive computational process. Filtering on the other hand, which is also known as

"updating", is used for modifying model variables (e.g., input or parameters) at each assimilation cycle (Neal et al. 2007), which is computationally efficient and, thus, more suitable for operational systems. The Ensemble Kalman Filter (EnKF; Evensen 1994, 2003), a widely-used data assimilation algorithm in hydrology (e.g. Moradkhani et al. 2005; Vrugt and Robinson 2007; Clark et al. 2008; Komma et al. 2008; Reichle et al. 2008; Xie and Zhang 2010; Nie et al. 2011; Li et al. 2013), was the technique used in this study. In a series of experiments where the EnKF performance was compared to that of its smoothing-based counterpart (i.e., the Ensemble Kalman Smoother; EnKS), Li et al. (2013) showed that filtering can yield similarly accurate results as smoothing does when both water balance and routing states are simultaneously corrected. In this work, this approach of simultaneous model states updating is followed, and, thus, filtering is deemed an appropriate choice. Lastly, although other filtering algorithms, such as the particle filter (PF), has been shown to be more robust and effective in the presence of non-Gaussian distributions of model residuals (DeChant and Moradkhani 2012), the EnKF can still provide the best linear unbiased estimates (i.e., first order accuracy). Furthermore, the EnKF is a better choice for operational systems because it generally requires a less number of ensemble members than the number of particles required by the PF.

The EnKF is a Monte Carlo simplification of the Extended Kalman Filter (EKF; Jazwinski 1970), a non-linear version of the Kalman Filter (Kalman 1960). The mathematics of the algorithm have been extensively described in many publications (see for example Evensen 1994, 2003; Evensen 2009), and therefore it is not done herein. The most important advantage of the EnKF over the EKF is that background

error statistical information is computed from the ensembles, thus linearization of the model and observation operator are not necessary. This particular approach to the retrieval of error statistics in the EnKF's algorithm establishes a natural linkage between ensemble forecasting and data assimilation (Hamill 2006), and represents a fundamental aspect of the methodology in this study.

Observation error specification

The deterministic form of the EnKF, entitled Ensemble Square Root Filter (EnSRF; Whitaker and Hamill 2002; Hamill 2006), was employed herein to avoid the need of perturbations on the observations. Although perturbations of the observations are necessary because otherwise the error covariance of the analysis is systematically underestimated (Hamill 2006), this can have a detrimental effect in the analysis itself (Clark et al. 2008). The EnSRF uses a reduced Kalman gain (i.e., the weight of the innovations) to update the perturbations and only requires specification of the observational error term. To accomplish this, uncertainty in streamflow observations was estimated with a simple model that considers the heteroscedasticity of measurement residuals (i.e. the variability of error deviations). The method is based on standard error values reported in the literature (e.g. Sauer and Meyer 1992; Di Baldassarre and Montanari 2009) and a log-linear direct relation between streamflow errors and recurrence intervals, which is consistent with claims of Sorooshian and Dracup (1980) and Vrugt et al. (2005). The regression fit and data used to construct the model was described in Vergara (2011) and is not included herein. The equation of observation residuals takes the following form:

$$Q_{error}(\%) = \begin{cases} 5 & for \quad RP < 1.0\,yr \\ 14.4 \ln RP + 5 & for \quad RP > 1.0\,yr \end{cases} \tag{2.6}$$

The parameters of the model in equation (2.6) correspond to particular characteristics of the measurements at the study site (e.g., frequency of flow values and reported level of accuracy of the USGS gage station).

2.3 Experimental design

The experiments in this study are based on a heuristic optimization of the uncertainty characterization scheme used to generate streamflow ensembles. The scheme consists of one component for the satellite-based QPEs and another one for the hydrologic model parameters, as described in Sections 2.2.3 and 2.2.4, respectively. The focus of this exercise was not to find an optimal configuration of the scheme, but rather to examine the sensitivity of the forecasting skill to the interactions between the uncertainty characterizations in the two components. Moreover, the overall objective was to extract information relevant to the usefulness of the uncertainty characterization of the hydrologic modeling system in the context of flood forecasting.

2.3.1 Uncertainty characterization optimization

The optimization was focused on two parameters of the uncertainty characterization scheme, each controlling the spread of the error variability modeled in each component. The parameter for the component on the hydrologic model controls the number (or fraction) of the standard deviations of the distribution from which the univariate random noise is drawn, and is referred to hereafter as P_{sf}. The values of P_{sf} considered were in the range 0.0 to 1.6, with increments of 0.2. A value of P_{sf} equal to

0.0 indicates that the first moment of the joint probability distribution (i.e., the parameter set found to be optimal from the hydrologic model calibration process) is employed. The maximum value of P_{sf} was constrained by the boundaries of the hydrologic model parameters' feasible ranges. Likewise, the parameter of the satellite-based QPE component controls the quantile distance above and below the conditional median of the fitted distribution (Figure 2.2) and is referred to hereafter as R_{dm}. The values of R_{dm} considered were in the range 0 to 45%, with increments of 5%. A value of R_{dm} equal to 0% indicates that the conditional median of the QPE reference estimates is employed. A value of 45% indicates that values within the 5% and 95% quantiles of the fitted distribution are used to build the ensembles.

Ensembles were generated from all possible combinations of P_{sf} and R_{dm} values from the ranges described above. In order to neglect sampling effects, all ensembles were built with 400 members, a size deemed adequate for statistical data retrieval. Each 400-member ensemble was used to simulate streamflow at a time step of 1-hour for the 3-year period of study. The skill of these simulations was quantified using the metrics of performance listed and discussed in Section 2.3.2. A qualitative and quantitative analysis was then performed to describe the optimality of the uncertainty characterization in a multi-objective fashion.

2.3.2 Skill assessment approach

Initial tests of this study examined the use of metrics of goodness-of-fit, such as the root mean squared error or the Nash-Sutcliffe Coefficient of Efficiency (NSCE; Nash and Sutcliffe 1970), commonly used in hydrology to test the performance of simulations (not included in the results presented in this work). The aggregation of the

errors and the scale at which these metrics are commonly integrated (i.e., over the entire period of simulation) overshadows the impact of the uncertainty in the hydrologic model parameters (i.e., no sensitivity to changes in P_{sf}). A property of metrics needed for the analysis herein is the ability to diagnose biases resulting from the dynamics operating at the scale of "quick" runoff generation. Therefore, the approach to skill evaluation employed herein was based on a selection of metrics that are commensurate with the scale at which floods are observed, and that are consistent with operational forecasting applications.

In operational hydrologic modeling systems, a common approach to the detection and prediction of flooding events is based on the use of flood thresholds (e.g., Reed et al. 2007). This is a minimum value of streamflow (or water depth) above which flooding is believed to occur (e.g., bank-full discharge; Williams 1978). A required prediction skill in such an approach is to simulate a basin's peak response, as depicted in a storm hydrograph. Consequently, the skill analysis herein was performed at the scale of individual flooding events. Figure 2.4 presents the selected events within the 3-year period of study and includes the simulation from the deterministic model (i.e., using the optimal parameter set and forcing the model with satellite-based QPE). The events were selected based on their recurrence interval and the ability of the deterministic model to simulate a noticeable response. With the exception of the event occurring in May of 2005, all events exceeded the 1.01-year recurrence interval.

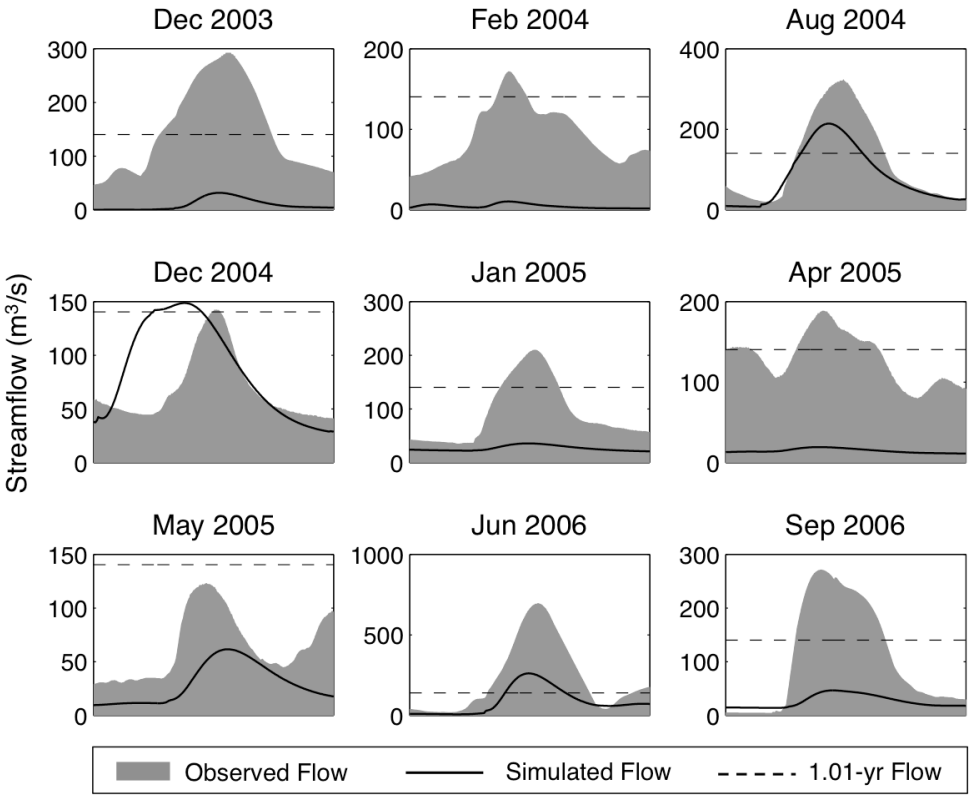

Figure 2.4: Selected streamflow events and baseline simulations using the deterministic model.

Based on the aforementioned considerations, three metrics to assess the skill of the ensembles' location (i.e., mean) were used in these experiments: Relative Peak Error (in units of %), Peak Time Error (in units of hours), and Relative Volume Error (in units of %). The Relative Peak Error describes the ability of the modeling system to simulate the magnitude of a flooding event. It is computed according to the following:

$$Peak\_Error(\%) = \left(\frac{Q_{sim}^{peak} - Q_{obs}^{peak}}{Q_{obs}^{peak}} \right) \times 100\% \qquad (2.7)$$

where Q_{obs}^{peak} is the event's observed peak flow in m^3/s and Q_{sim}^{peak} is the event's simulated mean peak flow in m^3/s. A negative value of the Relative Peak Error indicates underestimation of the event's peak flow, while a positive value indicates overestimation of the event's peak flow. The Peak Time Error describes the skill in

simulating flood wave routing in an aggregated manner. It was computed using serial date numbers, which represent the fractional number of hours from a reference date and time (e.g. 01-Jan-2000 00h):

$$Peak\_Time\_Error(hours) = Dt_{sim}^{peak} - Dt_{obs}^{peak} \tag{2.8}$$

where Dt_{obs}^{peak} is the serial date number of the observed peak flow in hours and Dt_{sim}^{peak} is the serial date number of the simulated mean peak flow in hours. A negative value of the Peak Time Error indicates peak flow is simulated early, while a positive value indicates peak flow is simulated late. Lastly, the Relative Volume Error describes the ability of the modeling system to simulate total runoff generation during a flooding event. It is computed as:

$$Volume\_Error(\%) = \left(\frac{\sum_{i=1}^{N} \left(Q_{sim}^{i} - Q_{obs}^{i} \right)}{\sum_{i=1}^{N} Q_{obs}^{i}} \right) \times 100\% \tag{2.9}$$

where Q_{obs}^{i} is the i_{th} observed streamflow value and Q_{sim}^{i} is the i_{th} simulated mean streamflow value of the event with a length of N hours. The series of Q in 2.9 are assumed to have the same Δt. When this is not the case, both observed and simulated hydrographs need to be integrated before computing the relative differences. A negative value of the Relative Volume Error indicates underestimation of the event's total runoff, while a positive value indicates overestimation of the event's total runoff. A schematic of the information of the hydrograph that these metrics aim to characterize is presented in Figure 2.5 below.

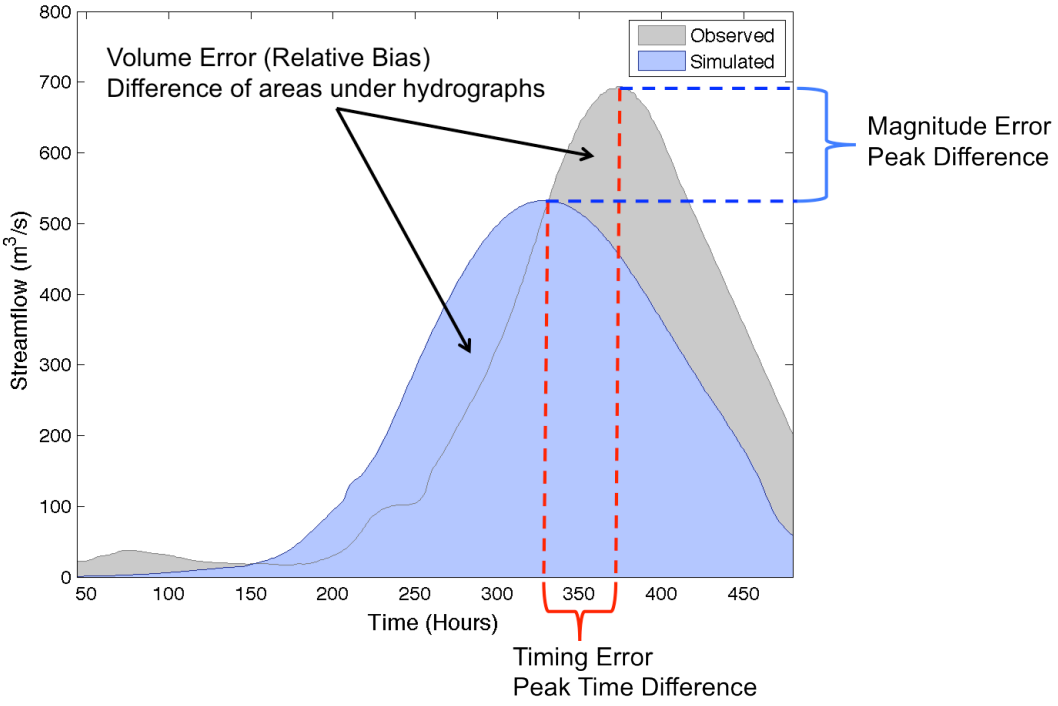

Figure 2.5: Schematic of hydrograph elements of interest and corresponding error metrics.

Additional to the aforementioned metrics of deterministic skill (i.e., skill of the ensembles' location), the probabilistic skill of the ensembles was evaluated (i.e., skill of the ensembles' spread). The ensemble characteristics most commonly evaluated are reliability, also known as "calibration" or "empirical validity", and sharpness (Hamill 2001; Carney and Cunningham 2006; Gneiting et al. 2007). Reliability refers to the ability of the ensemble to make good probabilistic predictions, which means that if the ensemble indicates that there is a P probability of a given flow value, the long-run proportion that actually occur turns out to be P (Carney and Cunningham 2006). Sharpness refers to the concentration of the predictive distributions around the observation (Gneiting et al. 2007). In other words, sharpness evaluates how spread out or how sharp the forecasts are (Carney and Cunningham 2006). Reliability is commonly measured using the rank histogram (Hamill 2001). Reliability was determined by two

metrics based on the rank histogram: 1) the skewness of the histogram, which identifies systematic problems with the ensemble, and 2) an error metric formulated herein and referred to hereafter as the Rank Histogram Relative Error (RHRE), which indicates how non-uniform the ensemble's rank histogram is. The metric is based on a fit to a beta distribution:

$$RHRE = \left(|1-\beta_1| + |1-\beta_2|\right) \times \frac{\beta_{\min} - 1}{|\beta_{\min} - 1|} \qquad (2.10)$$

where β_1 and β_2 are the maximum likelihood estimates of the beta distribution parameters, and $\beta_{\min}$ is the estimate with lower value. A perfectly uniform histogram has both parameters β_1 and β_2 equal to 1. If both parameters are greater than 1, the histogram will tend to have a dome shape indicating over dispersion. If both parameters are less than 1, the histogram will tend to have a "U" shape indicating under dispersion. The sign of the RHRE value reflects on the aforementioned: positive values indicate tendency to a dome shaped histogram, while negative values indicate tendency to the "U" shaped histogram. The skewness of the histogram corresponds to the difference between β_1 and β_2, where the sign indicates whether the ensembles densities are biased to the right (i.e., positive skew) or to the left (i.e., negative skew).

The Continuous Ranked Probability Score (CRPS; Carney and Cunningham 2006) was used here to assess sharpness. The CRPS is defined as the difference between the predicted and observed Cumulative Distribution Functions (CDF), and it is suitable for probabilistic forecasts of continuous variables, such as streamflow. For the different sample points, the CRPS is computed and then its mean value is reported. A

CRPS value of zero is perfect, while there is no upper bound to be defined. This is a measure commonly used to compare two or more ensembles.

2.3.3 Open loop vs. data assimilation runs

As part of the multi-objective strategy for the optimization, a comparison was performed between simulations without data assimilation, referred to hereafter as *open loop runs*, and those with data assimilation, referred to hereafter as *EnSRF runs*. This particular aspect of the evaluation targeted the skill of the uncertainty characterization (represented in the ensembles) in providing useful statistical information to the EnSRF for state updating and forecast improvement. The intent in comparing the performance of the simulations in the open loop and EnSRF runs was to test consistency between the ability of the ensembles in producing good forecasts and in activating the improvement capabilities of data assimilation.

EnSRF runs consisted of simulations during the 3-year period where observations of streamflow were assimilated at every time step. Assimilation of streamflow was employed to update all state variables of the hydrologic model (i.e., both water balance and routing model states). The skill metrics of the EnSRF runs were computed for the first guess (i.e., background), which in this case corresponds to the 1-hour forecast (forecast herein refers to the integration of the model forward in time based on updated states with the EnSRF). Additionally, the performance of the filter was evaluated at different lead-times (6-hr, 12-hr, 24-hr, 48-hr and 120-hr) for the June-July event of 2006, which was caused by a tropical storm.

Initial run tests showed that filter divergence occurred for runs with perturbations on model parameters alone (i.e., $P_{sf} = 0\%$). This is most likely due to over

constraint (i.e., low spread) in the specification of these ensembles, which causes underestimation of error covariance. This particular issue was not further investigated because it was only observed for these ensembles. Moreover, this set of ensembles (i.e., those with P_{sf} = 0%) was removed from the analysis due to its negative impact on the contouring of forecast skill.

2.4 Discussion of results

2.4.1 Physical significance of ensemble characteristics

Figure 2.6 illustrates the impact that different combinations of the uncertainty model parameters, R_{dm} and P_{sf}, have on the streamflow ensembles and associated forecast skill for a single streamflow event with the open loop model. The 3x5 multi-panel figure consists of 3 color-scaled contour plots located in the center (i.e., row 2, columns 2 – 4) featuring the event peak, peak time and volume errors, where the *x*- and *y*-axis of the contour plots correspond to the values of P_{sf} and R_{dm}, respectively. The contours are constructed from the computation of each error metric for all the ensembles resulting from all possible combinations of P_{sf} and R_{dm}. The 12 time-series plots on the outer panels (Figure 2.6 a – l) present hydrographs of the streamflow ensembles sampled from different locations over the contour plots.

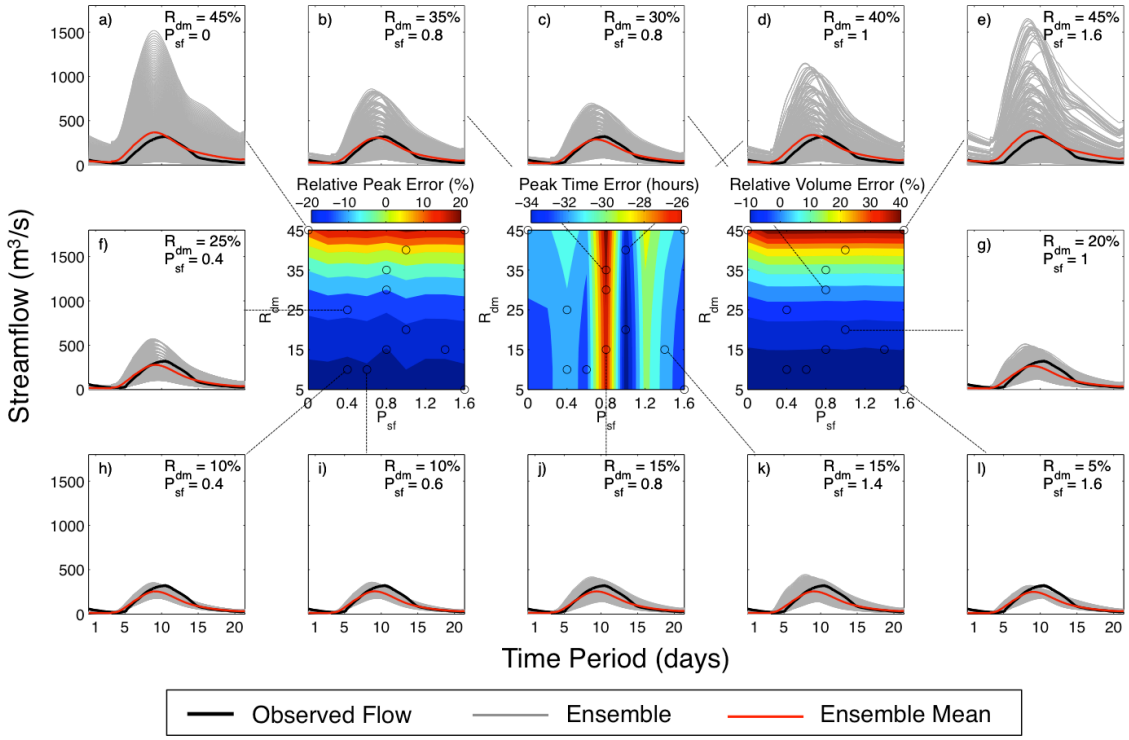

Figure 2.6: Sample 2-D surface plots of error metrics associated with open loop runs for the August 2004 event. The three panels in the center of the figure correspond to colored contour surfaces of the relative peak error (left), peak time error (middle), and relative volume error (right) of ensemble means. The outer panels (a – l) around the center surface plots correspond to ensembles featuring different combinations of P_{sf} and R_{dm}.

A feature that needs to be highlighted from the contour plots is the orientation of the error gradients with respect to the error model parameters. Both the relative peak error and the relative volume error display a vertical gradient along R_{dm}, while the peak time error displays a horizontal gradient along P_{sf}. This indicates that these metrics are able to effectively disentangle the independent impact of the two sources of uncertainty.

The correlation between both peak and volume errors and rainfall error are arguably intuitive. During a storm, the magnitude of streamflow is mostly defined by overland runoff, which results from the excess of rainfall after infiltration demands are satisfied. Moreover, at some point during the event (and for some events occurring under saturated soils conditions), the infiltration rate reaches a constant, which results in overland runoff being a linear function of rainfall. The three vertical panels on the far

left of Figure 2.6 illustrate the aforementioned situation, where ensembles are mainly based on rainfall perturbation, and they display spreads in the direction of the streamflow magnitude (i.e., vertical direction). Clearly, higher values of R_{dm} yield larger ensemble spreads and, as a consequence, higher mean streamflow values and forecasts of flooding events of higher magnitudes.

The correlation between peak time error and hydrologic model parameters error might not be as intuitive. Besides controlling aspects of the infiltration in the water balance, the hydrologic model parameters control the space-time integration of the total runoff over the basin. Particularly, the flow routing component in the hydrologic model represents processes over the land surface that delay streamflow (e.g., surface friction due to roughness, depression storage, and the topographic controls on runoff velocity). Therefore, the timing of the flow is directly dependent on model routing parameters. A perhaps counterintuitive feature observed in the peak error contour plot is the fact that perturbations on hydrologic model parameters do not seem to influence it, despite that there is a correlation between peak magnitude and timing in hydrologic models: that is, given a constant runoff volume, if the flow routing is modeled with higher velocities, the timing (as defined by the time-to-peak length) will be shorter, and the peak magnitude higher, than a flow routing modeled with lower velocities. The modeling of this behavior can be observed from the ensembles on the bottom row, where a spread of the peak flow can be evidenced to increase in an up-left to down-right direction, as the magnitude of P_{sf} increases. The mean of the peak errors, however, is not affected by this increments of spread. The peak time error contour plot does not display a defined direction of increasing values in contrast to those of the peak and volume errors. It

displays several local maxima and minima (e.g., along the R_{dm} = 5%, maxima at P_{sf} = 0.8 and P_{sf} = 1.2, and minima at P_{sf} = 0.6 and P_{sf} = 1.0). This particular feature indicates that optimal ensemble skill may require more than just high variance, as opposed to what the peak and volume error surface plots suggest.

2.4.2 Objective assessment of the uncertainty characterization with the EnSRF

Background skill

Figure 2.7 has the same configuration of Figure 2.6 but for model runs with the EnSRF. Although similar observations can be made in terms of the main driving error parameter (i.e., R_{dm} driving most of the errors in peak and volume, while P_{sf} driving most of the peak time error), the contour plots show that both error parameters influence the orientation of error gradients. The latter is an indication of the sensitivity of the error metrics to the interaction between P_{sf} and R_{dm} values in generating error statistics that are balanced by the EnSRF. Moreover, these interactions seem to be strongest at P_{sf} = 0.8 along the entire R_{dm} range, a feature consistent in all contour plots for the EnSRF runs. These error parameters values in fact correspond to where the peak time error contour plot of the open loop run shows a steep peak in its gradient, and where the peak magnitude error contour plot of the open loop shows a noticeable perturbation of its gradient (Figure 2.6). Additionally, the hydrograph plots (Figure 2.7 a – l) show the reduction of spread in the ensembles, which is indicative of success in the assimilation process. Significant variability in the resulting ensemble spread can be observed, which also points to the impact of the choice of P_{sf} and R_{dm} values on the performance of the EnSRF.

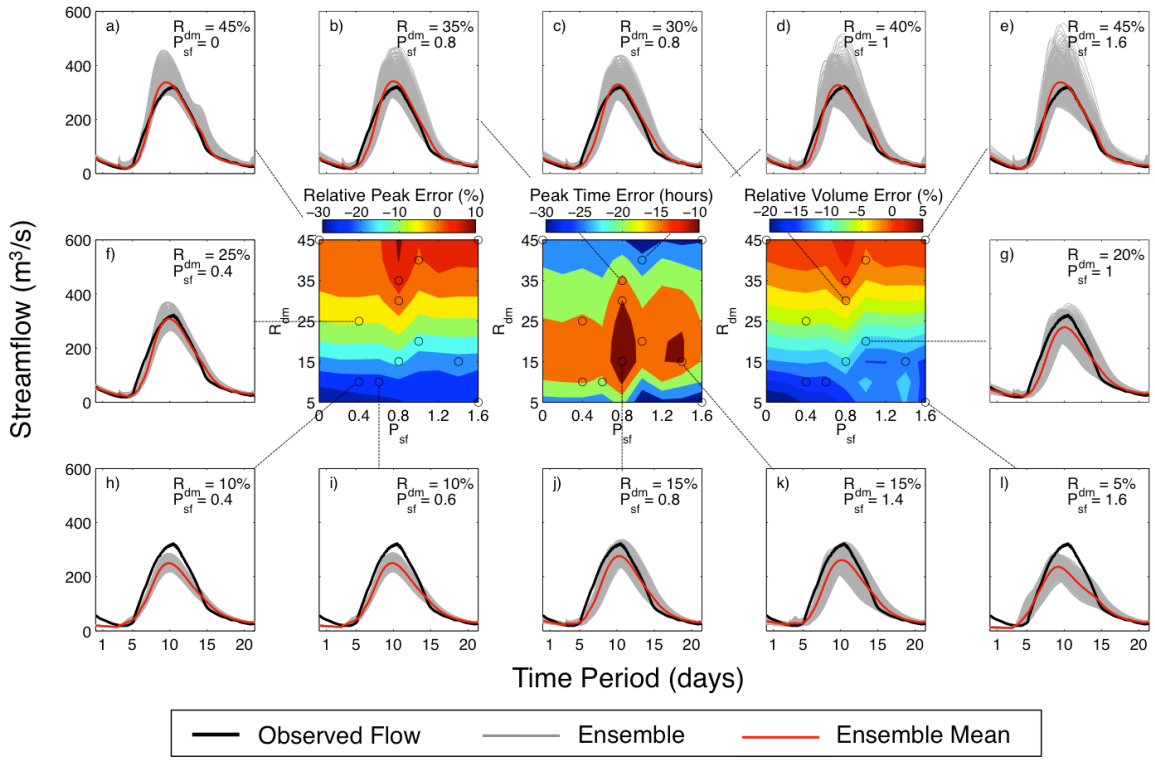

Figure 2.7: Same as Figure 2.6 but for the EnSRF runs.

An overall evaluation of all events considered in this study is summarized in Figure 2.8. Consistent features discussed for the single event in Figure 2.6 and Figure 2.7 can be observed for the average errors. The plots for the open loop runs show that, overall, the ensembles tended to underestimate the streamflow magnitude. Only at the upper extreme of the R_{dm} range do ensembles approach the events' peak and volume magnitudes. In terms of timing, peak flows tended to be early, with the exception of the ensembles along the 0.8 P_{sf} value. The EnSRF runs plots show that data assimilation improved the performance for the majority of ensembles. The exception is for ensembles with R_{dm} values less than 10%, for which no noticeable change occurred in terms of peak and volume errors, and deterioration of peak timing skill is observed. This once again highlights the importance of the interactions of error parameters in providing useful statistical information to EnSRF for effective adjustments. For the peak and

volume errors, a strong convergence toward the 0.8 P_{sf} value for values of R_{dm} larger than 35% can be observed. Lastly, there are strong interaction between P_{sf} and R_{dm} for the peak time error, drawing a clear boundary between early and late peak flows at around 30% R_{dm}.

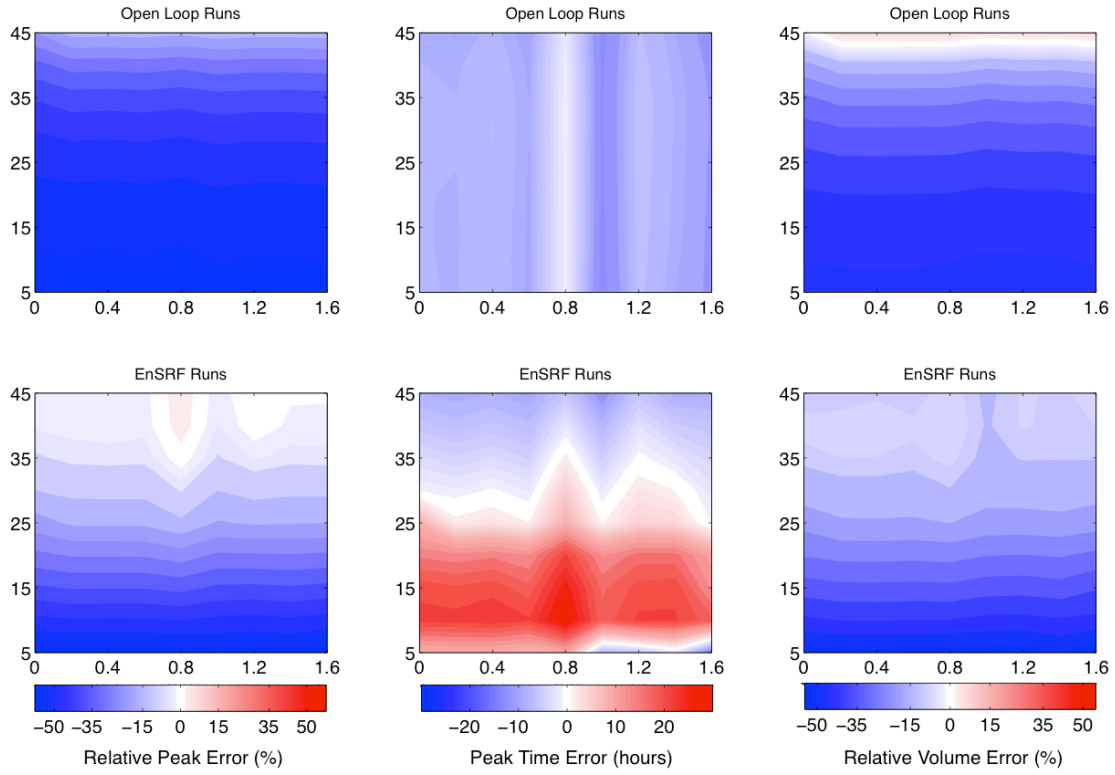

Figure 2.8: Overall deterministic performance of ensembles over the events. Upper panels present the open loop runs, and lower panels present the EnSRF runs. The values of P_{sf} are presented in the x-axis and the values of R_{dm} are presented in the y-axis.

Short to medium range forecast skill

Surface contour plots of the skill metrics for different lead-times are presented in Figure 2.9. Naturally, the overall skill diminishes as the lead-time increases. The timing skill seems to be the most impacted from this inherent reduction in the time-propagation of state adjustments. This is due to the fact that the peak time error is sensitive to the model trajectory, which is not explicitly adjusted during the filtering process (i.e., as

opposed to a smoothing process). A consequence of the latter is that peak timing is subject to higher uncertainty. At the 120 hours lead-time, the peak time error surface shows a rather different set of features than those with shorter forecast length. The relative peak error shows the most consistent features for all lead-times. Only at the 120 hours lead-time does the relative volume error show a clear optimum. It can be observed, however, that this area of optimal performance gradually appears as forecast lead-time increases.

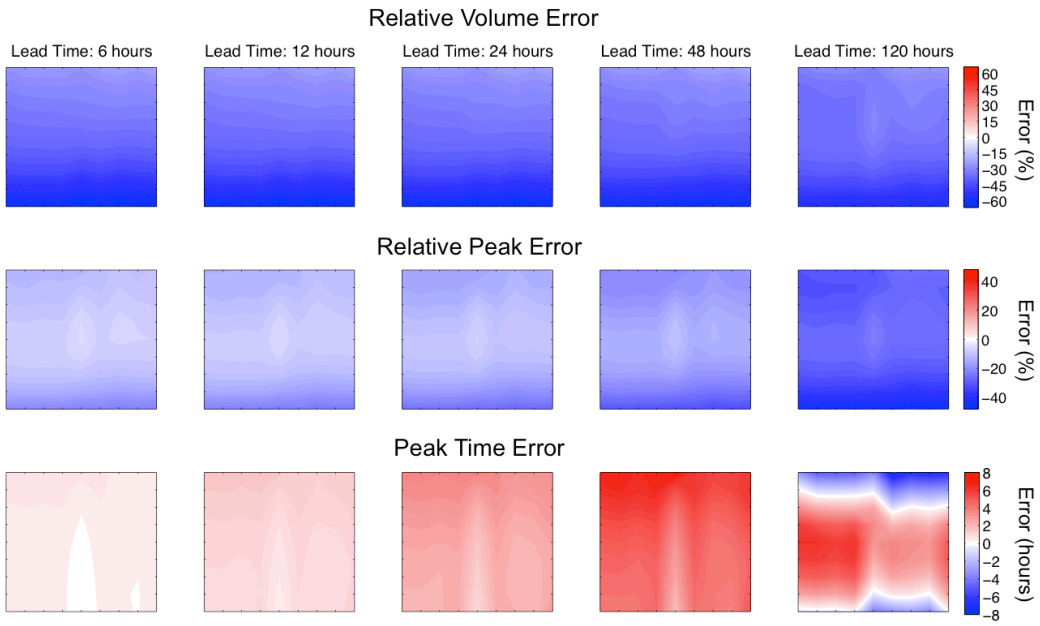

Figure 2.9: Evolution of deterministic model skill with data assimilation for various forecast lead times for the June of 2006 event. The values of P_{sf} are presented in the x-axis and the values of R_{dm} are presented in the y-axis.

2.4.3 Probabilistic skill evaluation

An evaluation of the reliability and sharpness of the ensembles was performed on the open loop runs. Flows exceeding the 1-year recurrence interval during the 3-year period of study were employed to compute cumulative probability functions and rank histograms to evaluate ensembles' sharpness and reliability, respectively. Figure 2.10

shows contour plots of the RHRE, rank histogram skewness and CRPS. All ensembles show under dispersion and negative bias (i.e. densities toward lower values of streamflow). This can be explained by the restrictions of the rainfall uncertainty characterization strategy in terms of modeling the detection limitations of the satellite-based estimates (see Section 2.2.3). There were several rainfall storms during the 3-year period of study that were essentially missed by 3B42RT, which resulted in significant underestimation of the simulated hydrologic response.

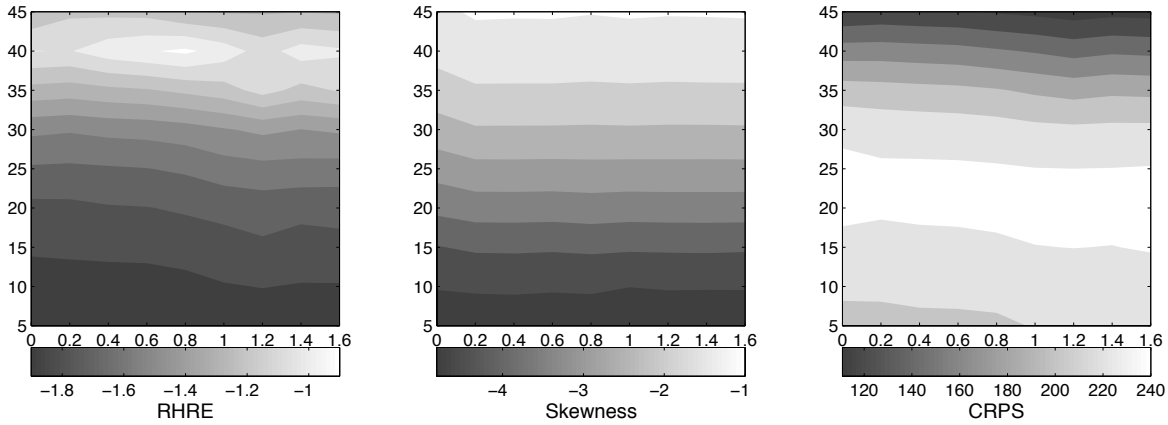

Figure 2.10: Probabilistic forecast skill of ensembles from the open loop runs. Flows exceeding the 1-year recurrence interval during the period of study were considered for the computation of these metrics. The values of P_{sf} are presented in the x-axis and the values of R_{dm} are presented in the y-axis.

An accurate evaluation of the probabilistic skill needs longer periods of data (> 10 years) to define the long-term characteristics that makes it statistically meaningful. Nevertheless, the analysis herein focused on the relative differences among the different ensembles in the context of the heuristic optimization. Overall, the gradients of these metrics show a clear vertical orientation along R_{dm}. Both CRPS and the skewness of the rank histogram lack an optimum and indicate that better skill is obtained toward the upper limits of R_{dm}. On the other hand, RHRE shows sensitivity to P_{sf}, although the overall trend is consistent with what other two metrics display. This sensitivity is

particularly significant for values of R_{dm} greater than 30% with an optimum at R_{dm} = 40% and P_{sf} = 0.8. In general, the behavior of these probabilistic skill metrics is consistent with the observations made on the deterministic ones.

2.4.4 Optimal parameters of uncertainty model

The contour plots in Figure 2.8 through 2-9 allow for the identification of different combinations of P_{sf} and R_{dm} values that yield optimal skill based on all considered metrics. Both peak and volume error plots for the open loop runs suggest values of R_{dm} exceeding 40%, independent of the P_{sf} value, while the peak time error plot points to a value of 0.8 of P_{sf}, independent of the R_{dm} value. As discussed in Section 2.4.3, a rather consistent behavior is seen in the reliability and sharpness plots. An absolute optimal on the reliability error metric was observed at R_{dm} = 40% and P_{sf} = 0.8.

The strong interactions between P_{sf} and R_{dm} in the EnSRF runs significantly constrain the options for optimal performance. In terms of the peak magnitude error, the absolute optimum is found at P_{sf} = 1.2 and R_{dm} = 40%, with a value of 0.54%. However, a secondary area of low peak magnitude error (i.e., local minimum) can be observed around P_{sf} = 0.8 and R_{dm} = 35%, with a value of 0.76%. For the peak time error, interactions between the full range of values P_{sf} and R_{dm} values between 25% and 35% yield similarly low values (approximately -1.6 to 1.6 hours). The absolute optimum, however, is found at R_{dm} = 35% and P_{sf} = 1.2, with a value of -0.37 hours. The volume error also display several local minima along R_{dm} = 40%, but the absolute optimum is found at the intersection with P_{sf} = 0.8, with a value of -6.77%. Lastly, the analysis on the single event in 2006 revealed that the region of optimal values of P_{sf} and R_{dm}

persisted for the different lead times and were more evident as the forecast length increased.

If the objective of this exercise were the selection of optimal values for the uncertainty model parameters P_{sf} and R_{dm}, the choice would be 0.8 and 40% respectively. An interpretation of these values would be that the majority of the uncertainty is in the satellite-based rainfall estimates, but that uncertainty in the estimation of hydrologic model parameters is not negligible. The high underestimation of rainfall is not surprising given the well-known issues in satellite-based estimates, as described in Section 2.2.3 and evidenced in the hydrographs presented in Figure 2.4. Also, significant impact from the uncertainty in hydrologic model parameters is expected because of the limitations in the estimation process and the simplifications in the representation of the dynamics of the real physical system.

2.5 Summary and conclusions

Quantification of uncertainty is becoming a key aspect for operational hydrologic forecasting frameworks due to its usefulness in decision-making situations and growing awareness of the limitations in deterministic systems. However, describing uncertainty is not an easy task due in part to the many challenges associated with its inherent multidimensional nature. In this study, an exercise to explicitly characterize variability in streamflow simulations originating from uncertainty in satellite-based quantitative precipitation estimates and from hydrologic model parameters was presented in the context of operational flood forecasting. The independent impacts of the two sources of uncertainty on the system's ability to simulate different aspects of the rainfall-runoff process leading to flooding were analyzed. A heuristic approach to

optimization was employed to test a diversity of characterizations featuring various levels of interaction between uncertainty coming from rainfall estimates and hydrologic model parameters. Sequential data assimilation was included in the evaluation strategy to study the statistical information content provided by the different characterizations.

The multi-objective approach to model performance evaluation proved useful in the disentanglement of the individual impact of the two sources of uncertainty characterized herein. The choice of metrics of skill (or error) and the scale at which they are integrated are critical to properly describe uncertainty and identify its sources. Simplifying the modeling problem by focusing on the dominant processes helps in the selection of these metrics and facilitates the diagnostic process. There were clear signatures in the gradients of the contour plots of the relative peak and relative volume errors that responded to the characterization of uncertainty in quantitative rainfall estimates. Likewise, the gradients in the peak time contour plots exhibited a clear response to the characterization of uncertainty in hydrologic model parameters. This association between skill metrics and their corresponding source of uncertainty, whose physical significance was explained in Section 2.4.1, demonstrates the diagnostic power of the evaluation approach.

Finding a balance in the characterization of the uncertainty in rainfall estimates and hydrologic model parameters was also effectively done through the multi-objective approach to heuristic optimization. It was possible to identify a region where the model performance was best from the intersection of the optima in the different contour plots (i.e., around P_{sf} = 0.8 and R_{dm} = 40%). Furthermore, consistency in the optimal region between open loop and EnSRF runs was observed. First, this validates the notion of the

existence of a natural linkage between ensemble forecasting and the data assimilation (i.e., ensemble design considerations, such as showing reliability and sharpness, might very well lead to better performance of an ensemble-based assimilation technique). Secondly, this illustrates that high values of error covariance are not the only required input to data assimilation. More importantly, the uniqueness of a balanced characterization motivates more investigation on the development of methodologies to accurately describe uncertainty.

In general, the results of this work present quantitative evidence of the importance in correctly characterizing uncertainty in a hydrologic modeling system. The experiments herein clearly demonstrate the sensitivity of the simulations' skill to the specifications of uncertainty characteristics. This agrees with observations that other studies have made, and responds to the necessity in providing guidance on methodologies to describe uncertainty in hydrologic forecasts. Particularly, this work highlights the importance of disentangling the individual contribution of uncertainty from different modeling components. Moreover, there are interactions among the different sources of uncertainty that need to be taken into account in order to yield balanced characterizations of the variability of hydrologic forecasts.

Further investigation needs to pursue the addition of other sources of uncertainty that were not considered in this study (e.g., model structure). Including other sources of uncertainty might prove more difficult in finding a unique and balanced characterization. It would most likely require an increase in the dimensionality of the evaluation approach. One way to achieve this could be through the application of distributed hydrologic models and the use of multi-site observational datasets (e.g.,

streamflow observations and/or soil moisture). The latter type of approach can also be useful to address the challenges involved in the regionalization of uncertainty characteristics, a subject much needed of development for the improvement of forecasting capabilities at ungauged locations.

Chapter 3. Configuration of a Distributed Hydrologic Model for Streamflow Simulation over the Conterminous United States

3.1 Introduction

Defining the specific behavior of interest of the hydrological system and its corresponding dominating processes is part of the initial steps in any modeling exercise (Reusser et al. 2009). This can in turn reduce the dimensionality of the problem and lower the complexity requirement of physics representation. Moreover, simplifying the modeling objective can facilitate performance diagnostics. The objective of the hydrologic modeling in this research is the simulation of streamflow for flood forecasting. Therefore, the representation of the hydrologic physical system is focused on processes relevant to events displaying catchment response to significant rainfall storms.

Estimation of model parameters is an important step in the configuration of hydrologic models. Model parameters are essential components of the equations and approximations in hydrologic models, whose values are allowed to change in order to represent a variety of watershed physical structures. This facilitates the application of the model over different regions. Some parameters are easily related to observable characteristics of the hydrologic system, while others are conceptually derived and cannot be directly measured in the field (Boyle et al. 2000). The latter leads to the use of indirect procedures for their specification, which can result in significant uncertainty. The standard procedure for the estimation of model parameters in hydrology is known as model calibration, an *ad hoc* solution commonly rooted in the "engineering"

approach to modeling. Modeling applications that involve prediction at ungauged locations should not estimate model parameters through calibration, since the data fitting process modifies the spatial variability structure of parameters that exist over a basin. This modification due to the introduction of artifacts from the calibration process increases the uncertainty in the modeling of the hydrology across a basin, thus compromising the reliability of forecasts at ungauged locations.

This chapter presents the configuration of the distributed hydrologic modeling strategy and other methodological aspects that are the basis of the experiments presented in subsequent chapters. A discussion on the geospatial datasets used to characterize watershed attributes and the particular hydrologic processes of interest over the Conterminous United States (CONUS hereafter) is included. Likewise, the hydrologic modeling framework and its representation of the physical system is explained in detail. The chapter closes with a synthetic experiment that illustrates the impact of model calibration on the spatial consistency of model skill.

3.2 Hydrologic geospatial datasets in the Conterminous United States

Streamflow is the primary hydrologic variable analyzed in this work. It results from the natural integration in space and time of the different hydrologic processes occurring in a watershed (or basin), the main physical unit subject to measurements and modeling in hydrology (Bedient et al. 2008). This inherent dimensionality reduction represents an advantage because it offers a mean to simplify the analysis of watershed response to rainfall. Moreover, streamflow observations in the United States have high consistency in time and space. The United States Geological Survey (USGS) manages a stream gauge network of over 10,000 stations (Gourley et al. 2013), of which

approximately 9,000 gauge stations are located over CONUS (Figure 3.1). Over 5,000 gauge stations correspond to streams that have no regulation or diversion. The identification of regulated stations was done by examination of the annual peak flow historical record of each USGS gauge station, where flags indicating the level of impact by regulation or diversion are specified. The research herein is limited to natural streams because the modeling approach employed for forecasting does not consider man-made structures (e.g., dams), which alter the hydrologic response of watersheds. The spatial distribution of these unregulated gauged locations spans the majority of CONUS' surface, which ensures adequate representativeness of its watershed diversity.

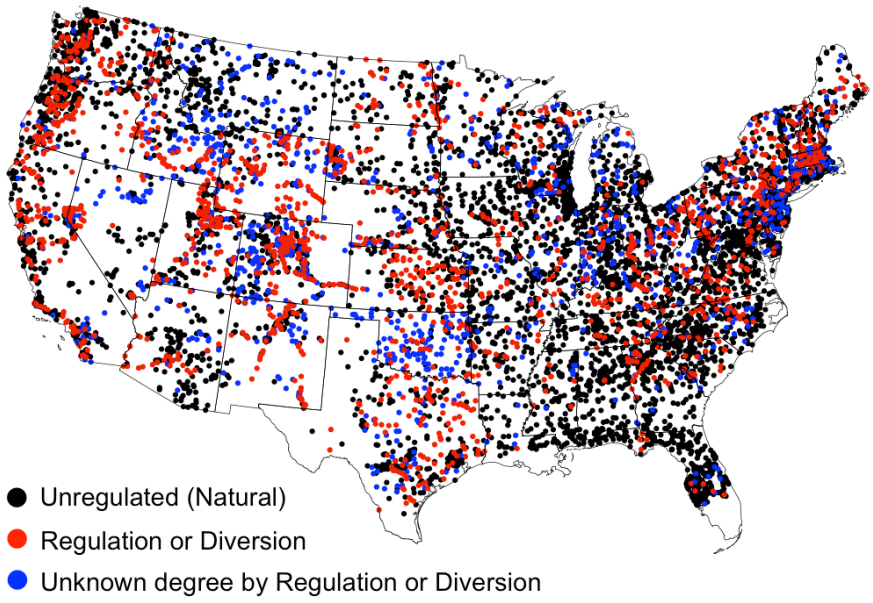

Figure 3.1: USGS stream gauge stations over CONUS. A classification for the identification of regulated catchments is included.

Several geospatial datasets for CONUS have been gathered for the development of this work. Besides being relevant to the FLASH project, studying the uncertainty of a hydrologic modeling system at the CONUS scale enables a sound analysis based on a diversity of basin characteristics and response, which is required to resolve macro scale

patterns of hydrologic response and for extending forecasting capabilities to ungauged areas.

Characteristics of the physical structure of watersheds

For effects of analysis and the hydrologic model implementation, the pixel of a rectangular grid is defined herein as the elementary unit representing a stream reach and the immediately adjacent overland area (i.e., hillslope). The particular characteristics of each stream reach, assumed to be uniform within the pixel, are uniquely determined by the flow contributed by its drainage basin, its current and past geology, topography, pedology and climate, and are part of a spatial continuum that includes the entire watershed (Dingman 2009). Therefore, several of these geophysical characteristics of watersheds were explored for the research herein. All geospatial datasets employed in this study were rendered on a rectangular grid with a 1-km pixel resolution. The grid was specifically chosen to match the radar forcing data employed for the flash flood forecasting system over the CONUS.

Watershed boundaries (i.e., watershed divide) and several basin geomorphological characteristics are defined by topography. Using Digital Elevation Model (DEM) data, it is possible to derive geomorphological parameters of any given watershed or catchment. A DEM is virtually available everywhere over the globe at high resolution (e.g., 30 meters), which enables the ability to generate geomorphological information at all gauged and ungauged locations. The DEM data used herein were based on the USGS' National Elevation Dataset (NED; Gesch et al. 2009). Figure 3.2 presents the topography over CONUS generated with a 1-km Digital Elevation Model (DEM) raster dataset. The highest elevation is found in the North

America Cordillera on the western half of CONUS at approximately 4,400 meters. Complex terrain features are defined by the arrangement of the different mountain belts over this region of CONUS (i.e., Pacific Coast Ranges, the Cascade Range and Sierra Nevada), which result in numerous independent drainages of relatively small size. The Appalachians Mountains on the far eastern region of CONUS is a secondary significant mountain range.

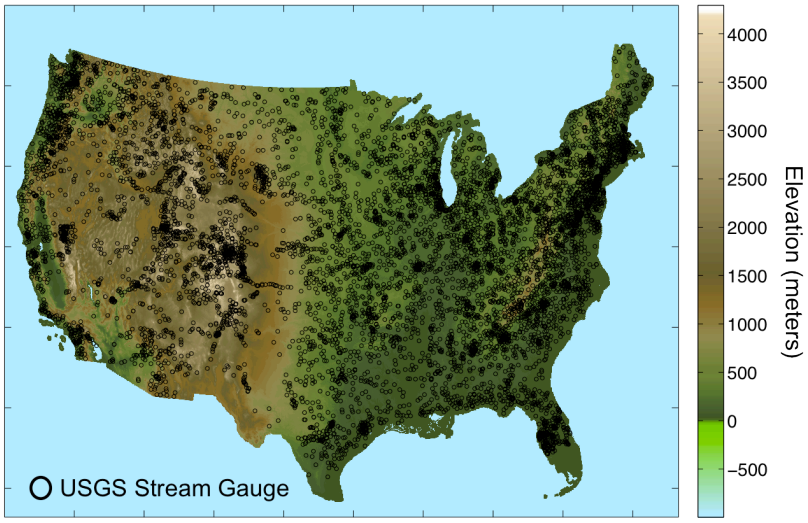

Figure 3.2: Digital Elevation Model (DEM) data over CONUS. The pixel resolution is 1-km.

A DEM is also the most basic dataset for the configuration of distributed hydrologic models. Flow direction and accumulation defined at each pixel of a computational grid are based on the DEM. These two parameters are essential for flow routing and the delineation of the basins. Figure 3.3 shows all independent watersheds (i.e., these basins drain either to the ocean or an inland water body such as a lake) over CONUS derived from the 1-km DEM dataset. Following the same approach employed on the derivation of these independent watersheds, basin delineation was performed for the catchment of each of the USGS stream gauge over CONUS. In this way, it was

possible to associate particular basin geophysical parameters to the response observed at the gauge.

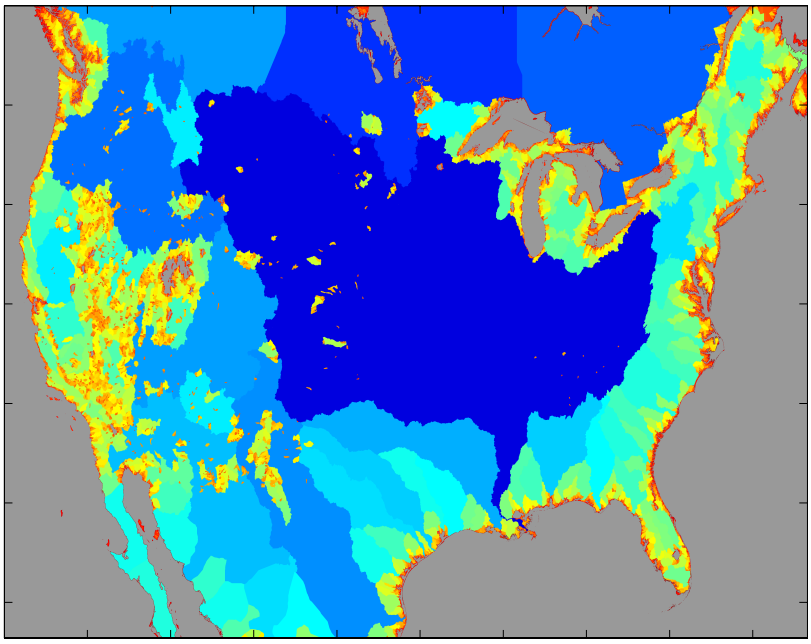

Figure 3.3: All independent catchments draining to either the ocean or an inland water body, derived from the 1-km DEM dataset.

Geomorphological variables considered herein were selected based on the studies by Schumm (1956) and, in particular, Costa (1987) who analyzed relationships between characteristics of watersheds and flash floods over the CONUS. The variables include drainage basin, elongation ratio, relief ratio, slope index, slope at the outlet, and river length. These geomorphological variables were derived automatically using an algorithm based on the DEM for the approximately 9,000 USGS stream gauges. The procedure first delineates a basin and computation of the geomorphological variables follow using raster-based equivalents of the measures needed to define each of them. Therefore, the products are subject to uncertainty due to the elevation estimates and resolution of the grid. However, results show sufficient skill for the analysis presented in subsequent chapters.

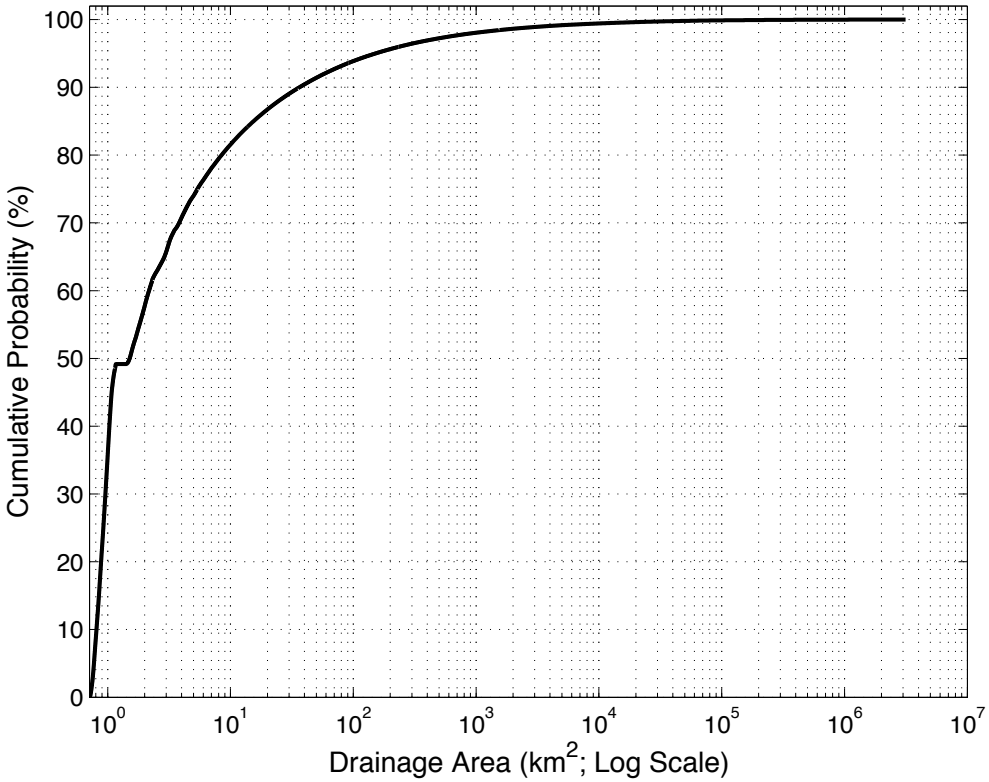

Figure 3.4: Empirical cumulative distribution of drainage areas over CONUS, computed from the 1-km drainage area grid.

Figure 3.4 above shows the empirical cumulative distribution of drainage areas derived from all pixels of the 1-km grid. It can be seen that the great majority of drainages over CONUS are small. More than 90% of the pixels represent drainages below 100 km^2, and over 95% represent drainages below 1,000 km^2. This is an important piece of information for the analysis done in this research in relation to the hydrologic modeling development over CONUS.

The hydro-climatology of basins was considered by examining mean annual precipitation and average temperature. The data correspond to the 30-year datasets prepared by the PRISM Climate Group[2] covering the period 1981 - 2010. Figure 3.5

[2] PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, created September 2013

presents the 30-yr annual precipitation and average temperature over CONUS, which shows the substantial spatial variability of hydro-climatic regimes included in the analysis of this work.

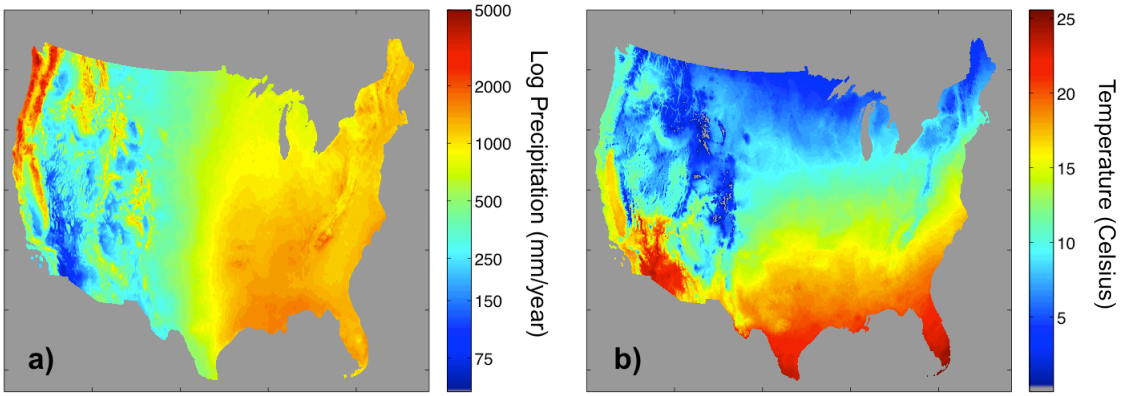

Figure 3.5: Hydro-climatic regimes over the Conterminous United States defined by a) the 30-year Annual Precipitation and b) 30-year Mean Temperature (1981 – 2010; PRISM Climate Group, http://www.prism.oregonstate.edu/normals/, created in 2013).

Soil datasets from the STATSGO database (Soil Survey Staff 1994; Miller and White 1998) were examined herein. Variables explored from this dataset include hydrologic soil group and mean depth-to-bedrock (Figure 3.6), soil class, mean rock volume percent, and erodability factor (K factor). Some of these variables can be used to directly define *a-priori* values for some of the hydrologic model (see Table 3.1).

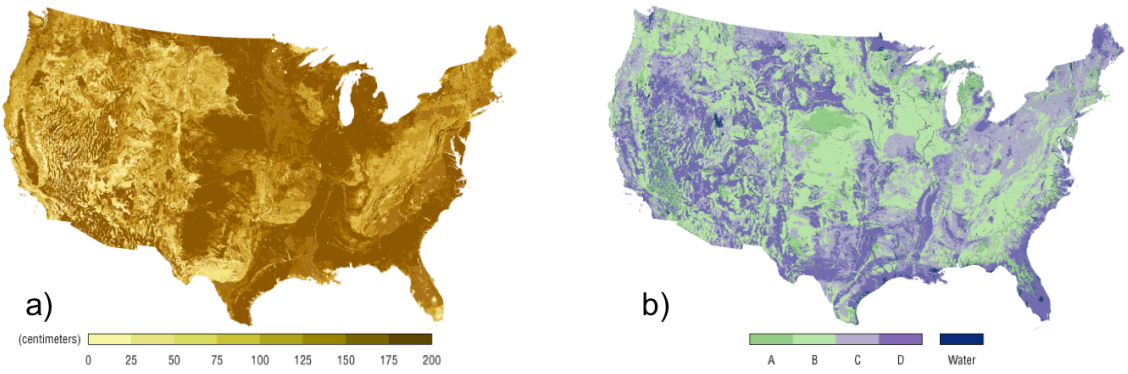

Figure 3.6: STATSGO soil products (Miller and White 1998): a) Depth to bedrock, and b) Hydrologic Soil Group (images taken from http://www.soilinfo.psu.edu).

Lastly, land cover and land use data from the National Land Cover Dataset (NLCD 2006; Fry et al. 2011) were utilized to estimate *a-priori* parameters for the water balance component of the hydrologic model, and explore the runoff (i.e., USDA NRCS) curve number in the analysis performed in subsequent chapters.

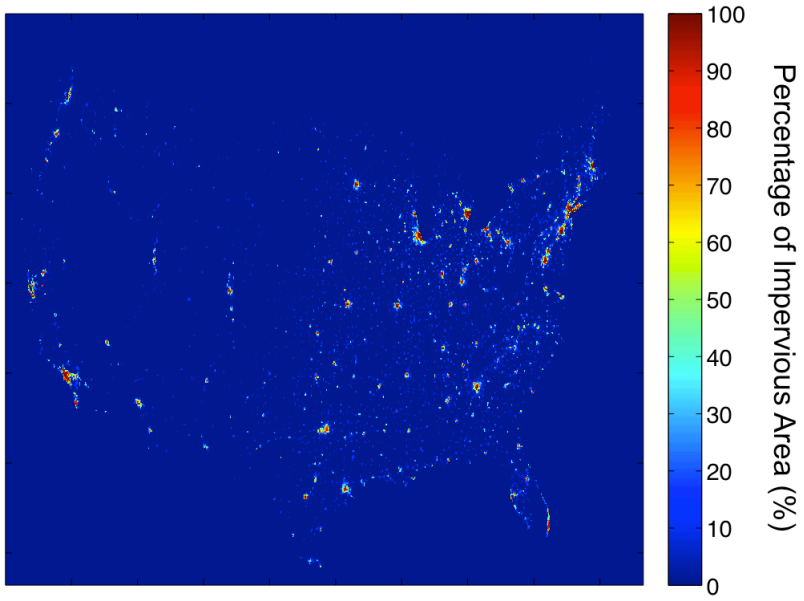

Figure 3.7: Percentage of built-up land (Fischer et al. 2008) over the Conterminous United States.

3.3 Representing the physical dynamics of the hydrologic system

3.3.1 Hydrologic modeling framework

The studies in this research primarily utilized the hydrologic modeling framework employed in FLASH to generate hydrologic forecasts. It consists of the Ensemble Framework For Flash Flood Forecasting (EF5[3]). The framework features several hydrologic model physics and capabilities to work with multiple forcing inputs to enable ensemble forecasting (Figure 3.8). It employs a rectangular grid customizable to any pixel scale and georeferenced system. Moreover, model physics in EF5 can be

[3] Ensemble Framework For Flash Flood Forecasting (EF5): http://ef5.ou.edu

integrated across different spatial scales: from the scale of an individual pixel to the continental or global scale. EF5 can also be configured for any temporal integration step size, although it is predominantly set to time steps ranging from sub-hourly (e.g., 5-min) to daily for its application in floods and flash floods forecasting.

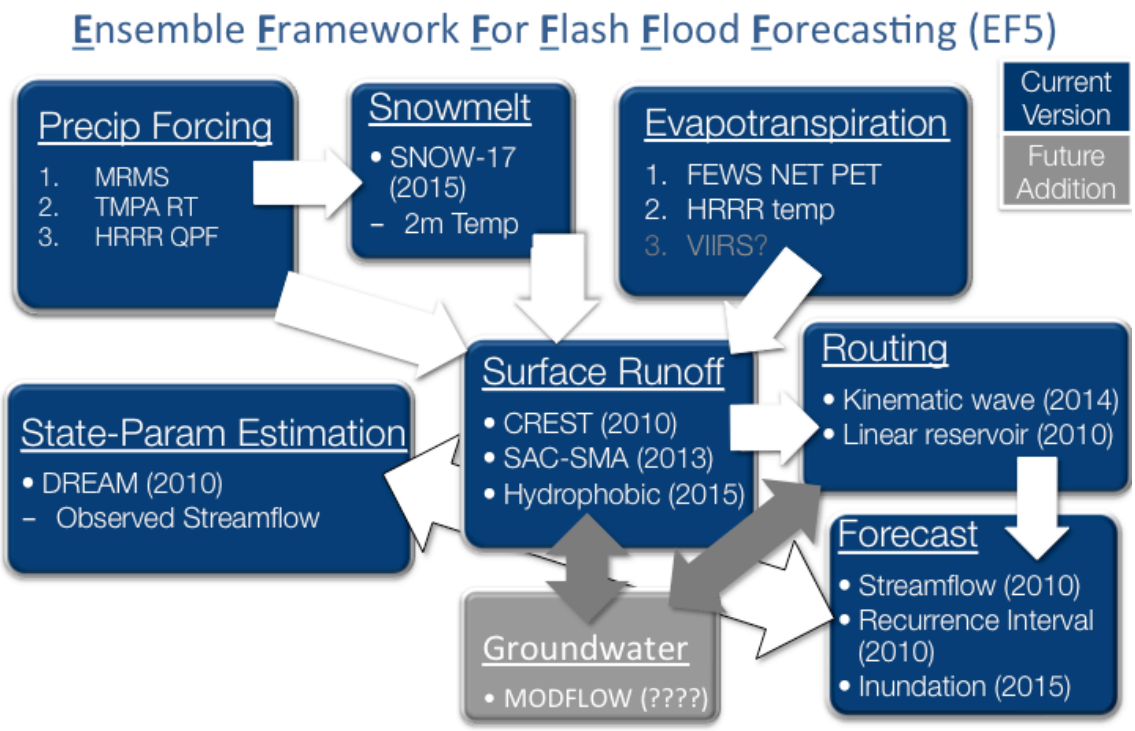

Figure 3.8: Schematic of modeling components in EF5 (http://ef5.ou.edu).

One of the core model physics in EF5 is based on the Coupled Routing and Excess Storage (CREST) distributed rainfall-runoff model recently developed by Wang et al. (2011) in a collaborative effort between The University of Oklahoma and the National Aeronautics and Space Administration (NASA). CREST is the rainfall-runoff model employed in the experiments presented in Chapters 4 and 5. Therefore, a detailed description of its structure is presented in the following section. EF5 also utilizes the physics of the Sacramento Soil Moisture Accounting Model (SAC-SMA; Burnash et al. 1973), a conceptual rainfall-runoff (CRR) watershed model widely used within the

NWS River Forecast System (Sorooshian et al. 1993; Boyle et al. 2000; Koren et al. 2004). A brief description of SAC-SMA is additionally included herein for the purpose of the modeling experiment of this chapter.

Water balance conceptualization

CREST is a conceptual hydrologic model that transforms rainfall into runoff according to the schematic in Figure 3.9. Compared to so-called physics-based hydrologic models and other conceptual models (e.g., SAC-SMA), CREST is relatively simple, which makes it attractive for operational systems. Moreover, simple structures have the property of being more *identifiable* (Gourley and Vieux 2006) than those with (sometimes excessively) higher complexity. The latter is a key consideration for the overall objective of this research, in relation to the ability to understand uncertainty in streamflow forecasts.

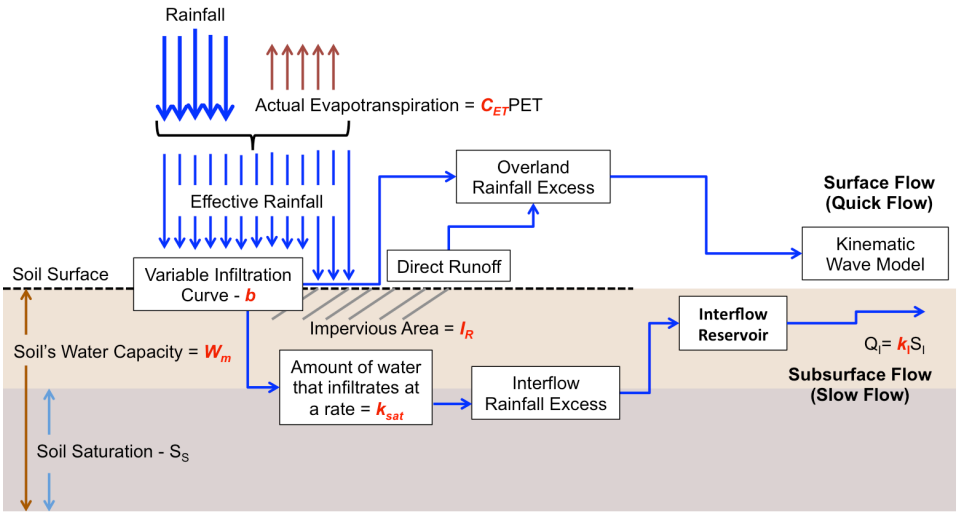

Figure 3.9: Schematic of the Coupled Routing and Excess Storage (CREST) rainfall-runoff model. CREST parameters are presented in red font.

The water balance model is based on the variable infiltration curve (Zhao et al. 1980; 1995) for the computation of excess rainfall, which is partitioned into its surface

and subsurface components through a conceptual mechanism based on hydraulic conductivity (Wang et al. 2011). The mathematical representation of the rainfall-runoff process in CREST is detailed as follows.

Rainfall P_t is first "transformed" into soil precipitation $P_{S,t}$ (effective rainfall in the schematic in Figure 3.9). The soil precipitation is defined as the amount of precipitation that makes it to the soil. $P_{S,t}$ is computed with the following piecewise function:

$$P_{S,t} = \begin{cases} 0, & for \quad P_t \le aET_t \\ \left(P_t - aET_t\right) \times \left(1 - I_R\right), & for \quad P_t > aET_t \end{cases} \tag{3.1}$$

where I_R is a parameter representing the portion of the pixel covered by impervious surface and aET_t is the actual evapotranspiration, which is computed as a linear function of the potential evapotranspiration PET_t:

$$aET_t = C_{ET} \times PET_t \tag{3.2}$$

where C_{ET} is a scalar that needs to be defined. PET_t and P_t are both model input data. Once on the soil, some of the rainfall is infiltrated according to:

$$I_t = \begin{cases} 0, & for \quad P_t \le aET_t \vee SM_t \ge W_m \\ W_m - SM_t, & for \quad \left(i_t + P_{S,t}\right) \ge i_m \\ W_m - SM_t - W_m \times \left[1 - \dfrac{i_t + P_{S,t}}{i_m}\right]^{1+b}, & for \quad \left(i_t + P_{S,t}\right) < i_m \end{cases} \tag{3.3}$$

where W_m is a parameter representing the maximum water capacity of the soil, b is the exponent of the variable infiltration curve, i_m is a parameter representing the maximum infiltration capacity defined as:

$$i_m = W_m \times \left(1 + b\right) \tag{3.4}$$

SM_t in (3.3) is a model state variable representing soil moisture, and i_t is the infiltration capacity given as:

$$i_t = i_m \times \left(1 - \left(1 - \frac{SM_t}{W_m} \right)^{1/(1+b)} \right) \tag{3.5}$$

The subtraction of infiltration to soil precipitation is referred to as excess rainfall ER_t and is given as:

$$ER_t = \left\{ \begin{array}{l} 0, \ for \quad P_{S,t} = 0 \vee P_{S,t} \le I_t \\ P_{S,t} - I_t, \ for \quad 0 < P_{S,t} > I_t \end{array} \right\} \tag{3.6}$$

Excess rainfall is partitioned into its surface and subsurface components, $ER_{O,t}$ and $ER_{I,t}$ respectively:

$$ER_{I,t} = \left\{ \begin{array}{ccc} 0, & for & P_t \le aET_t \\ temX_t, & for & ER_t > temX_t \\ ER_t, & for & ER_t \le temX_t \end{array} \right\} \tag{3.7}$$

where $temX_t$ is given by:

$$temX_t = \left\{ \begin{array}{ll} \frac{SM_t + W_{A,t}}{2W_m} \times k_{sat}, & for \quad P_t > aET_t \\ \left(aET_t - P_t \right) \times \frac{SM_t}{W_m}, & for \quad P_t \le aET_t \end{array} \right\} \tag{3.8}$$

where k_{sat} is a parameter presenting hydraulic conductivity, and $W_{A,t}$ is given as:

$$W_{A,t} = \left\{ \begin{array}{ccc} 0, & for & P_t \le aET_t \\ W_m, & for & SM_t + I_t \ge W_m \\ SM_t + I_t, & for & SM + I_t < W_m \end{array} \right\} \tag{3.9}$$

$ER_{O,t}$ is then computed as the difference between ER_t and $ER_{I,t}$ plus the water that directly runs off over impervious surface:

$$ER_{O,t} = \left\{ \begin{array}{c} 0, \ for \quad P_t \le aET_t \\ ER_t - ER_{I,t} + \left(P_t - aET_t\right) \times I_R, \ for \quad P_t > aET_t \end{array} \right\} \tag{3.10}$$

The last step in CREST computations is the update of its state variable SM_t:

$$SM_{t+1} = \left\{ \begin{array}{ll} W_{A,t}, & for \quad P_t > aET_t \\ SM_t - temX_t, & for \quad P_t \le aET_t \end{array} \right\} \tag{3.11}$$

$ER_{O,t}$ and $ER_{I,t}$ are routed with the surface and interflow routing models respectively. For the flow routing, EF5 currently features two models: a distributed version of the linear reservoir (Nash 1957), a lumped routing model commonly used in hydrology (Moore 1985; Chow et al. 1988; Vrugt et al. 2003), and the kinematic wave approximation of the Saint-Venant equations for one-dimensional unsteady open channel flow (Chow et al. 1988). The configuration of EF5 used herein couples CREST with the kinematic wave model for surface flow routing and with the linear reservoir technique for subsurface flow routing (Figure 3.9).

Distributed flow routing model

In general, there are two types of flow routing models: lumped routing models and distributed routing models, sometimes referred to as *hydrologic routing* and *hydraulic routing* respectively (Chow et al. 1988; Bedient et al. 2008). Lumped routing models usually employ empirical or conceptual ideas to describe the true mechanisms of water flow process in a hydrologic system (e.g., linear reservoir). Distributed routing models, on the other hand, consider both space and time. Furthermore, and because

water flow is a continuous variable, these models solve partial differential equations related to the physical laws governing the water movement mechanisms in a hydrologic system. Depending on the assumptions and approximations applicable to a particular hydrologic system, different distributed routing models can result.

The model selected herein for the surface flow routing was the kinematic wave approximation to the one-dimensional unsteady open channel flow equations developed by Barré de Saint-Venant in the 1800s (Beven 2011). The full implementation of the Saint-Venant equations represents the closest description of the 1-D water movement in a watershed. However, the use of alternative models by simplification of the governing equations is motivated by simpler and computationally less expensive methods for distributed flow routing. Additionally, these simpler models can capture the dominant physical processes, depending on specific flow conditions. Kinematic wave model is arguably the most widely-used distributed flow routing method in hydrologic modeling, given its simplicity, as compared to the diffusion or dynamic wave models. A general criterion to support the use of the kinematic wave approximation is based on the slope: in watersheds with predominantly steep slopes with free-flowing streams (no "backwater" effects), the flow conditions are such that the kinematic wave concept reasonably approximates the unsteady flow phenomena (Ponce 1986). Moreover, Ponce (1991) claimed that for most overland flow situations, kinematic wave approximation requirements are satisfied. Kazezyilmaz-Alhan and Medina (2007) define a minimum slope of 0.002 as a general guidance value required for kinematic wave applicability. Figure 3.10 presents a map of the applicability of the kinematic wave approximation over the Conterminous United States (CONUS) based on the aforementioned criterion.

It can be observed that the kinematic wave approximation applies for the majority of CONUS.

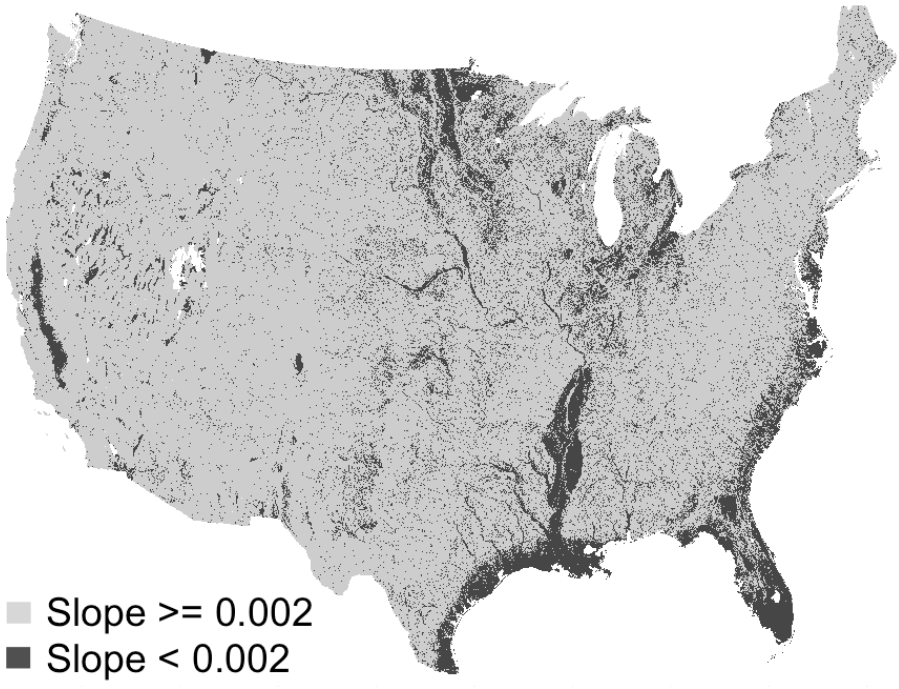

Figure 3.10: Applicability of the kinematic wave approximation over the Conterminous United States based on slope. The slope grid is based on a 1-km Digital Elevation Model (DEM) grid.

Several well-known models or modeling frameworks implement kinematic wave for the flow routing component, such as the Hydrologic Engineering Center (HEC)'s Hydrologic Modeling System (HEC-HMS; Feldman 2000) and Flood Hydrograph Package (HEC-1; Feldman 1995), the Environmental Protection Agency (EPA)'s Storm Water Management Model (SWMM; Huber and Singh 1995), the National Weather Service (NWS)'s Research Distributed Hydrologic Model (HL-RDHM; Koren et al. 2004), the Distributed Hydrology Soil Vegetation Model (DHSVM; Wigmosta et al. 1994; Wigmosta et al. 2002), and the KINEmatic Runoff and EROSion (KINEROS; Woolhiser et al. 1990) model, among many others.

The form of the kinematic wave model for channel flow routing used in EF5 (Chow et al. 1988):

$$\frac{\partial Q}{\partial x} + \alpha\beta Q^{\beta-1}\frac{\partial Q}{\partial t} = q \tag{3.12}$$

where q is the lateral inflow to the channel from overland flow routing. A more detailed description of the derivation of Equation 3.12 is presented in Chapter 4. The surface excess rainfall component obtained from water balance computations is routed as overland flow with an implementation of the kinematic wave model for a wide shallow (sheet) flow as:

$$\frac{\partial q}{\partial x} + \alpha_0 \frac{3}{5} q^{3/5-1}\frac{\partial q}{\partial t} = i - f \tag{3.13}$$

where q is the overland flow in $m^3/s.m^2$ and the lateral inflow term of equation (3.12), $i - f$, is the surface excess rainfall from the water balance in m/s (i.e., $ER_{O,t} = i - f$), and α_0 is an overland conveyance parameter defined as a function of Manning's roughness coefficient and overland slope alone. The kinematic wave model in EF5 is numerically solved using an implicit non-linear scheme, as explained in (Chow et al. 1988). The state variables for the surface routing are the overland flow q and streamflow Q.

Lastly, subsurface flow routes $ER_{I,t}$ with the following function:

$$S_{I,t+1} = S_{I,t} - k_I S_{I,t} + ER_{I,t} \tag{3.14}$$

where $S_{I,t}$ is a state variable representing interflow storage and k_I is a parameter representing the proportion of interflow storage leakage. The product term $k_I S_{I,t}$ denotes the subsurface flow routed downstream.

3.3.2 A calibration-free modeling strategy

One of the targets of this study is to enable regional forecasting of floods over CONUS. A key consideration of such a framework is the ability to extend forecasting capabilities to ungauged locations. Prediction at ungauged locations requires the configuration of hydrologic models, which includes the estimation of their parameters, over the entire area of interest (e.g., a watershed) regardless of the availability of observations of model states. An important design aspect of the hydrologic modeling strategy of this work is that model parameters are estimated using measurements related to a basin's physical structure. In other words, parameter estimation is performed using *a-priori* knowledge and information, and therefore *no calibration of the model is performed*. Model calibration consists of a process where model outputs are fitted to observations generally available at limited number of points within a watershed. It is theorized that in order for the model to match observed response at gauged locations, this procedure modifies the spatial structure of the variability of physical characteristics and, thus, the dynamics of hydrologic response across the basin. Consequently, this approach to parameter estimation is deemed not appropriate for applications that involve prediction at ungauged locations. Particularly, this method can negatively affect the analysis in this research in relation to the spatial characterization of uncertainty in streamflow forecasts over CONUS (see Chapter 5).

The majority of model parameters in CREST can be estimated directly from available geospatial datasets. Therefore, the hydrologic model was configured with *a-priori* estimates for all of its parameters. This includes seven parameters for the water balance and the excess rainfall routing (subsurface and surface), and three parameters

for the flow routing model (see Table 3.1). The kinematic wave model parameters α and β were estimated following an innovative approach, which is explained in detail in Chapter 4.

Table 3.1: CREST – Kinematic Wave model parameters and *a-priori* estimates

Parameter	Description	Source
W_m	Soil Water Capacity	STATSGO dataset (Miller and White 1998)
b	Infiltration Curve Exponent	STATSGO dataset (Miller and White 1998) and look-up table in Cosby et al. (1984)
k_{sat}	Hydraulic Conductivity	STATSGO dataset (Miller and White 1998)
k_I	Speed of subsurface flow	STATSGO dataset (Miller and White 1998) and empirical relationship based on CN number (Pokhrel et al. 2008)
coem	Manning's Coefficient for Overland routing	UMD vegetation category from 2007 MODIS. At: http://webmap.ornl.gov/wcsdown/wcsdown.jsp?dg_id=10004_32
I_R	Imperviousness Area Ratio	URB_2000 - built-up land (residential and infrastructure)" From Harmonic World Soil Database (HWSD; Fischer et al. 2008) http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/LandUseShares.html?sb=9
C_{ET}	Linear adjustment factor on Potential Evapotranspiration	Set subjectively to 1.0
Th	A threshold drainage area value above which a pixel is defined as a stream	Set subjectively to 5.0 km^2
α	Kinematic wave coefficient of the momentum equation	As defined in Chapter 4
β	Kinematic wave exponent of the momentum equation	As defined in Chapter 4
α_0	Overland kinematic wave conveyance parameter	Manning's equation using slope derived from DEM and parameter *coem*.

3.4 Impact of model calibration on spatial skill consistency

To illustrate the impact that model calibration has on the simulation skill at interior upstream points, an Observing-Systems Simulation Experiment (OSSE; Arnold Jr and Dey 1986) was devised using EF5. The experiment consisted in a calibration of a

distributed hydrologic model using simulated time-series of streamflow (i.e., synthetic observations) generated by a second distributed hydrologic model. Because the synthetic reference streamflow comes from a model, it is possible to evaluate the "skill" of the calibrated model at interior points. Additionally, it is possible to evaluate the simulation of other states of the system that are generally not available for actual physical systems.

A relatively small catchment with available observations of streamflow was employed for this exercise. The basin corresponds to the catchment of USGS stream gauge 07325800 over central-western Oklahoma on Cobb Creek, with a drainage area of about 342 km^2 (Figure 3.11). A 1-year period from January to December 2007 was used for the experiment. A significant flooding event triggered by the remnants of Tropical Storm Erin in August 2007[4] is included in this period.

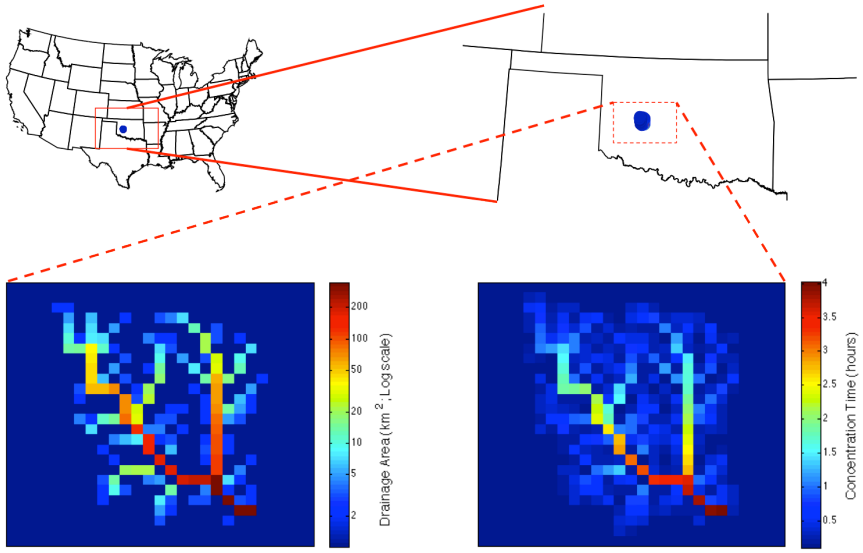

Figure 3.11: Study basin location in CONUS. Basin's drainage area grid (bottom left) and analytical concentration time grid (bottom right) are also included.

4 http://www.ok.gov/OEM/Emergencies_&_Disasters/2007/Severe_Weather_Event_20070819_Master/, Retrieved on April 2015.

EF5 simulations were generated using CREST and SAC-SMA water balance models, both coupled to the kinematic wave flow routing model. Figure 3.12 presents the schematic of SAC-SMA model. The model structure and other details of SAC-SMA are well described in Koren et al. (2000; 2003; 2004) and Yilmaz et al. (2008), and so the focus herein is on the presentation of the model parameters and their estimation.

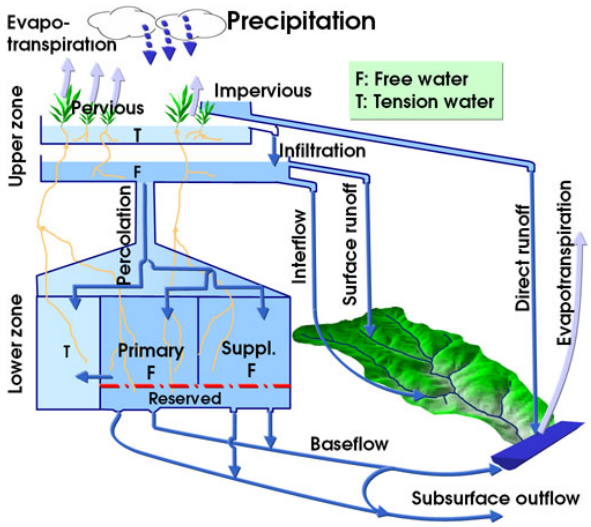

Figure 3.12: Schematic of the SAC-SMA rainfall-runoff model (http://chrs.web.uci.edu).

The model simulates runoff generation using 17 conceptual parameters (Table 3.2). Koren et al. (2000) developed an approach to estimate *a-priori* values for 11 of the SAC-SMA parameters based on the State Soil Geographic soil data (STATSGO) (Soil Survey Staff 1994, 1996). The remaining six parameters use lumped values established by the NWS from previous experience on different basins (Pokhrel et al. 2008; Yilmaz et al. 2008).

Table 3.2: List and description of Sacramento Soli Moisture Accounting Model (SAC-SMA)

	Parameter	Description
SPATIALLY DISTRIBUTED	UZTWM	The upper layer tension water capacity, mm
	UZFWM	The upper layer free water capacity, mm
	UZK	Interflow depletion rate from the upper layer free water storage, day^{-1}
	ZPERC	Ratio of maximum and minimum percolation rates
	REXP	Shape parameter of the percolation curve
	LZTWM	The lower layer tension water capacity, mm
	LZFSM	The lower layer supplemental free water capacity, mm
	LZFPM	The lower layer primary free water capacity, mm
	LZSK	Depletion rate of the lower layer supplemental free water storage, day^{-1}
	LZPK	Depletion rate of the lower layer primary free water storage, day^{-1}
	PFREE	Percolation fraction that goes directly to the lower layer free water storages
LUMPED	PCTIM	Permanent impervious area fraction
	ADIMP	Maximum fraction of an additional impervious area due to saturation
	RIVA	Riparian vegetation area fraction
	SIDE	Ratio of deep percolation from lower layer free water storages
	RSERV	Fraction of lower layer free water not transferable to lower layer tension water
	EFC	Effective forest fraction

Similarly to CREST, input data for SAC-SMA consist mainly of precipitation and potential evapotranspiration data. The models used radar precipitation estimates integrated at a 1-hr time step and climatological estimates of PET (Koren et al. 1998). A comparison of both simulations to actual observations of streamflow is presented in Figure 3.13. Several goodness-of-fit metrics were employed in the analysis presented

herein. The relative bias of a streamflow time-series is a measure of systematic under or over estimation. It is defined as:

$$Bias(\%) = \frac{\sum_{i=1}^{N}\left(Q_{sim}^{i} - Q_{obs}^{i}\right)}{\sum_{i=1}^{N} Q_{obs}^{i}} \times 100 \tag{3.15}$$

where Q_{sim}^{i} is the simulated streamflow value and Q_{obs}^{i} is the observed streamflow value at time *i,* for *N* time steps in the time-series. The relative bias can take values between minus infinity and infinity, where the perfect value is zero. A negative relative bias indicates under estimation, while a positive bias indicates over estimation. Pearson's correlation (also known as correlation coefficient) is a measure of linear association between two data series. It is defined as:

$$r = \frac{N\sum_{i=1}^{N}\left(Q_{sim}^{i}Q_{obs}^{i}\right) - \sum_{i=1}^{N}Q_{sim}^{i}\sum_{i=1}^{N}Q_{obs}^{i}}{\sqrt{N\sum_{i=1}^{N}\left(Q_{sim}^{i}\right)^{2} - \left(\sum_{i=1}^{N}Q_{sim}^{i}\right)^{2}}\sqrt{N\sum_{i=1}^{N}\left(Q_{obs}^{i}\right)^{2} - \left(\sum_{i=1}^{N}Q_{obs}^{i}\right)^{2}}} \tag{3.16}$$

Values of Pearson's correlation range between -1 and 1. A value of zero indicates no correlation. A positive value indicates direct correlation with 1 being the perfect value. Likewise, a negative value indicates indirect correlation with a perfect value of -1. A rank correlation can also be computed using 3.16 replacing the values of streamflow by their corresponding rank position in the series. The rank correlation is a better metric for non-linear associations. Likewise, it is useful in determining simulation skill to rank values regardless of their magnitudes.

It can be observed that both models consistently underestimate the magnitude of streamflow. This is partially due to the coarse temporal resolution of the precipitation

input data and at which the models were integrated. Nevertheless, the model based on CREST physics performed better overall with a relative bias of -50% and a Pearson's correlation of 0.67, while SAC-SMA had a relative bias of -85% and a Pearson's correlation of 0.63. Therefore, CREST was used to produce the synthetic observations of streamflow for the OSSE.

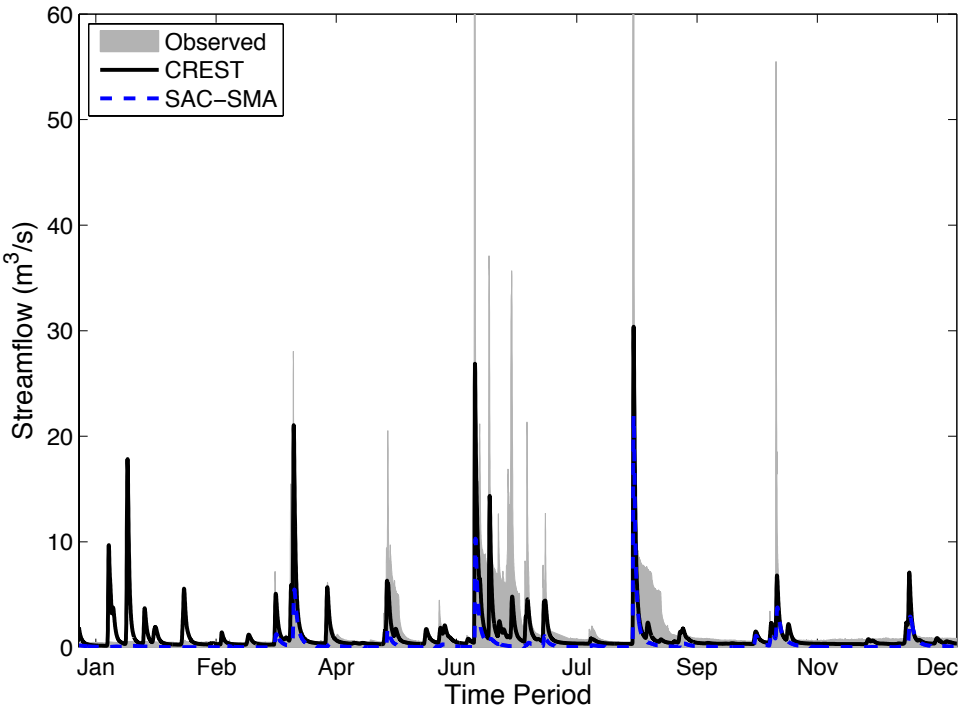

Figure 3.13: Streamflow simulations from EF5 as compared to observations from USGS stream gauge during 2007 in the study basin. Y-axis was truncated to 60 m^3/s for better visualization of simulated series. Maximum observed peak value is 200 m^3/s in August.

Figure 3.14 presents the baseline simulation of the OSSE. CREST simulations will be referred to hereafter as "truth", and are used to calibrate and evaluate the SAC-SMA model performance. The baseline simulation shows underestimation by the SAC-SMA model, with a relative bias of -72%. The correlation of the time series is significantly high, as indicated by Pearson's and rank correlation coefficients. This is expected because both models use the kinematic wave model for flow routing.

The hydrologic model was calibrated using EF5's built-in general-purpose optimization algorithm entitled Differential Evolution Adaptive Metropolis (DREAM) (Vrugt et al. 2009). DREAM is an adaptation of the widely used Shuffled Complex Evolution algorithm (SCE-UA; Duan et al. 1993). The algorithm employs Markov Chain Monte Carlo (MCMC) sampling to estimate the posterior probability density function of parameters using a formal likelihood function, which ensures a collective evolution of the model parameters (i.e., all parameters are optimized simultaneously to account for interdependencies among them). The algorithm is designed to operate in complex, high-dimensional sampling problems (Vrugt et al. 2008). The reader is encouraged to review the work by Vrugt et al. (2009) for more details regarding DREAM.

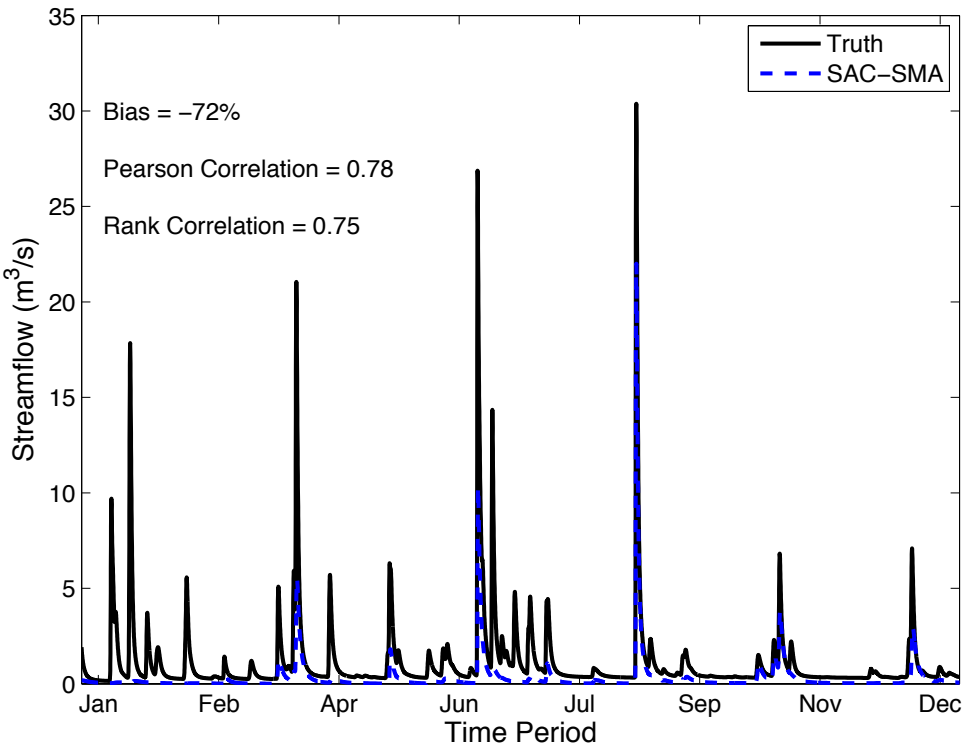

Figure 3.14: Baseline OSSE simulation: CREST simulation (black solid line) is used as reference (Truth) and SAC-SMA simulation (blue dashed line) is used as the benchmark for calibration.

Two model calibration runs are performed for the experiment. The first one assumes that no *a-priori* estimates of model parameters are available, and so spatially uniform estimates of the parameters are determined through calibration. This run is referred to as "Uniform". The second run utilizes model calibration to find scalar values for *a-priori* estimates of the parameters of the baseline model. In other words, initial spatially distributed estimates of the parameters are available, and so calibration is used to scale their values, keeping the relative differences among pixels and, thus, the underlying variability of the grids. This run is referred to as "Scaled *a-priori*". Results of both calibration runs are presented in Figure 3.15.

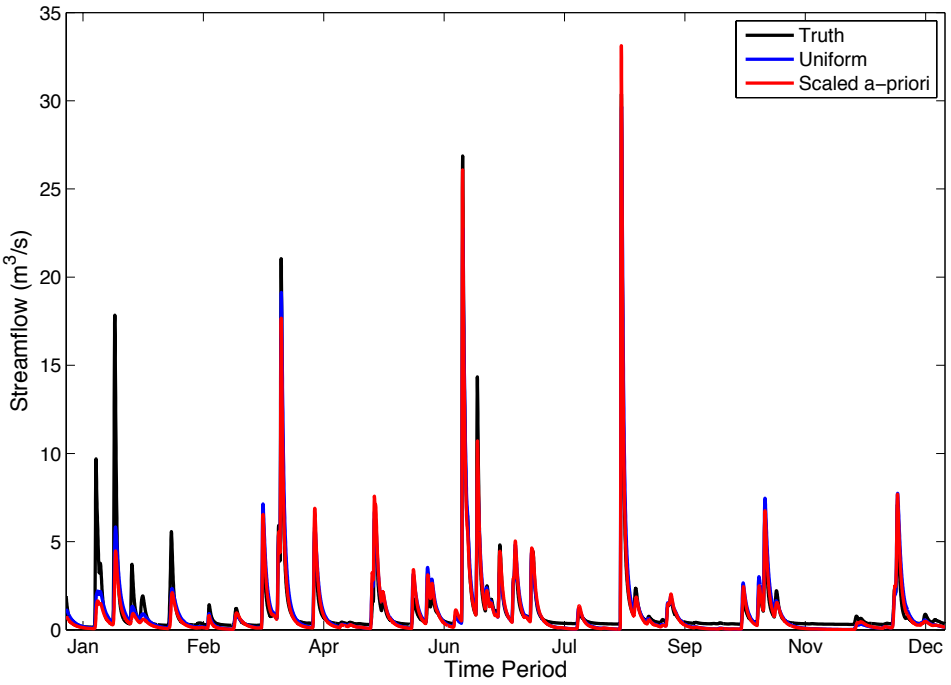

Figure 3.15: OSSE calibration results for both Uniform and Scaled *a-priori* runs.

Visual inspection of the hydrographs from the calibrated models indicates the calibrations were successful. The differences between the baseline model and the truth (Figure 3.14) have been virtually eliminated. Moreover, the differences between the two models are negligible. Additional to the relative bias and correlation metrics, other

aggregated statistics of skill commonly used to evaluate model calibrations were computed for an objective assessment. The relative root mean squared error (RMSE) is defined as:

$$RMSE(\%) = \frac{\sqrt{\frac{\sum_{i=1}^{N}\left(Q_{sim}^{i} - Q_{obs}^{i}\right)^{2}}{N}}}{\overline{Q}_{obs}} \times 100 \tag{3.17}$$

where $\overline{Q}_{obs}$ is the average of observed streamflow values. The range of the RMSE is zero to infinity. The Nash-Sutcliffe Coefficient of Efficiency (NSCE) is defined as:

$$NSCE = 1 - \frac{\sum_{i=1}^{N}\left(Q_{sim}^{i} - Q_{obs}^{i}\right)^{2}}{\sum_{i=1}^{N}\left(\overline{Q}_{obs} - Q_{obs}^{i}\right)^{2}} \tag{3.18}$$

The range of values of NSCE is between minus infinity and 1. A perfect value of NSCE is 1. A value of zero indicates that the model has the same skill as the average of observations. A negative value indicates the model has no skill. Table 3.3 presents a summary of the skill metrics for both runs. The values in the table confirm the assessment by visual inspection. The underestimation at the beginning of the period can be explained by insufficient simulation spin-up (warm-up), which is inconsequential for the analysis of interest.

Table 3.3: OSSE calibration statistics of skill for both Uniform and Scaled *a-priori* runs.

	Bias (%)	**RMSE (%)**	**NSCE**	**Pearson Corr.**	**Rank Corr.**
Baseline	-72.16	155.24	0.40	0.78	0.75
Uniform	-2.98	72.45	0.87	0.93	0.91
Scaled *a-priori*	-12.96	72.66	0.87	0.93	0.92

Both calibrated models were employed to generate gridded streamflow maps for the entire period of simulation at every hourly time-step. With these multi-location time-series, it is possible to compute statistics of skill and generate grids that show their spatial consistency. Figure 3.16 presents grids of relative bias (%) for each model over the basin of study. It can be observed that the effect of calibration varies. Some interior locations are improved, while some others are deteriorated. Overall, the improvement is mainly observed on the main streams, while deterioration occurs on smaller drainages where overland processes are represented. Moreover, it can be seen that some locations are severely deteriorated, and that in general calibration causes overestimation for the majority of the basin. Additionally, there are not noticeably differences between the two calibrated models.

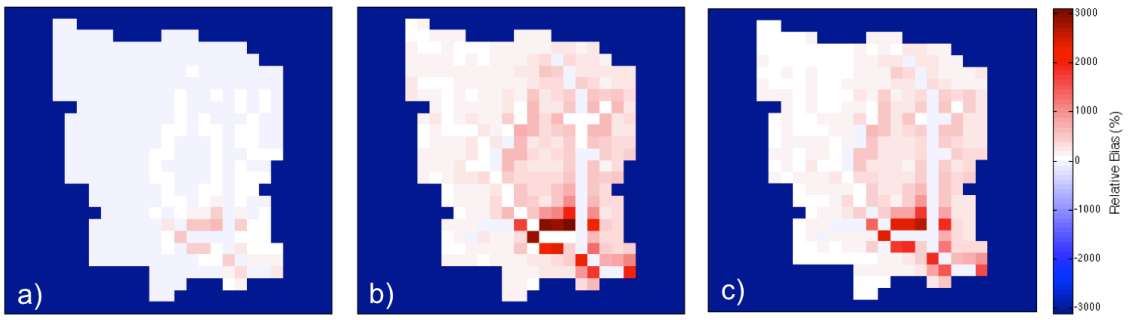

Figure 3.16: Comparison between a) baseline, b) calibrated uniform, and c) calibrated scaled *a-priori* simulations based on relative bias (%).

A pixel-to-pixel comparison of skill with respect to the baseline model and a summary of the spatial distribution of relative bias are presented in Figure 3.17. The information in panel a) can be interpreted as follows: points located above the 1:1 line indicate that calibration increased the relative bias; conversely, points located below the 1:1 line imply that calibration decreased the relative bias; points over the 1:1 line indicate no change. Ideally, points should tend to move to the horizontal line at 0.0 %

relative bias. It can be seen that for the great majority of points, the relative bias increases. For some of these points, the value is over five times larger after calibration. The distributions of relative bias values over the basin shown in panel b) demonstrate the spatial inconsistency of skill that is achieved at the outlet. Moreover, there is greater variability in this skill for the calibrated models, as indicated by the larger spread of the distributions.

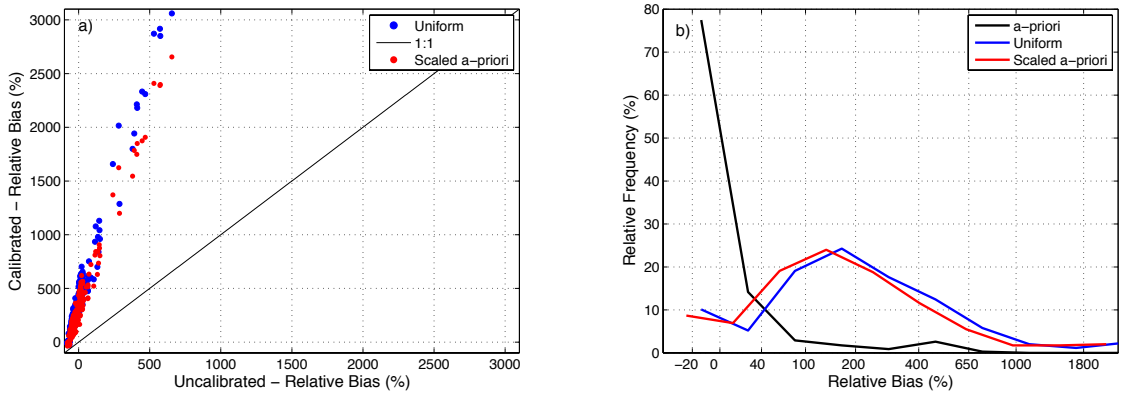

Figure 3.17: Overall comparison of skill over the basin between for the baseline and calibrated models: a) Scatter plots of relative bias (%) at every pixel; and b) distributions of relative bias (%).

The same plots were produced for the rank correlation coefficient. This particular statistic of skill is of great interest for flood forecasting because it indicates ability to rank flows. This skill is required specifically for correct detection of flooding events. Figure 3.18 shows a similar pattern of deterioration and improvement as the one discussed for relative bias. The majority of improvement is seen on and around the main streams, while deterioration is particularly evident on overland pixels. Likewise, both calibrated models result in almost equivalent maps.

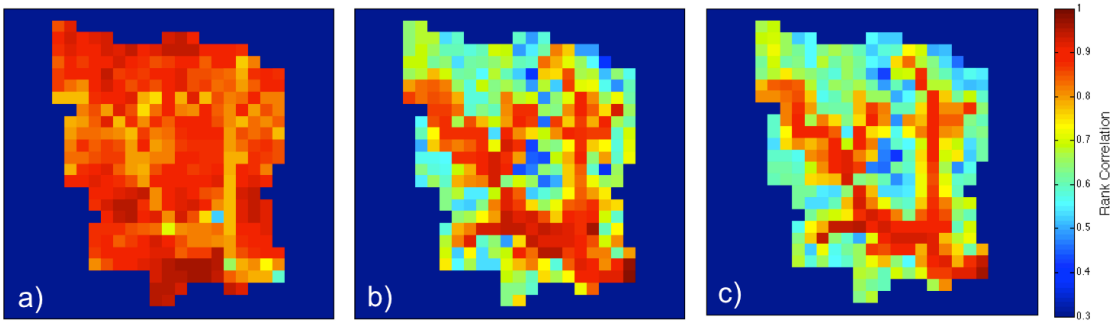

Figure 3.18: Same as Figure 3.16 but for the rank correlation coefficient.

The pixel-to-pixel comparison and summary of skill distribution over the basin in Figure 3.19 shows in more detail the significant deterioration caused by model calibration. Specifically, it can be seen the significant variability of rank correlation over the basin in both plots. This once again demonstrates the spatial inconsistency of skill.

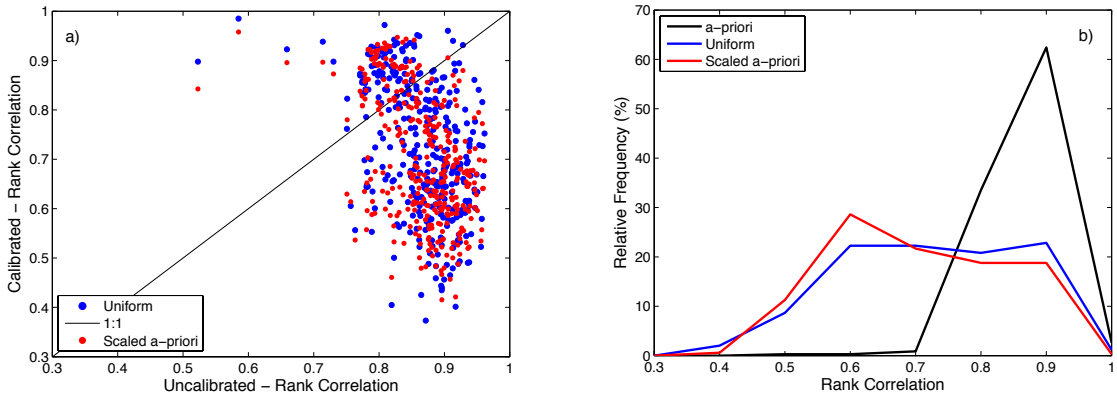

Figure 3.19: Same as Figure 3.17 but for the rank correlation coefficient.

Besides what was discussed in the aforementioned assessment, these results highlight a critical downside of calibration. Because both CREST and SAC-SMA employ the same physics and *a-priori* parameter estimates for flow routing, the baseline model results in already skillful simulations of flow timing, which is indicated by the high values of correlation in panel a). By definition, the "Uniform" calibration cannot

resolve required spatial patterns of variability of parameter estimates and, thus, it is not surprising to observe deterioration of flow timing skill. The "Scaled *a-priori*", on the other hand, is based on the baseline model and so flow timing skill should not be altered.

Lastly, the impact of calibration on internal states was assessed after generating multi-location time-series of soil moisture. The same type of plots as those employed for the relative bias of gridded streamflow time-series were analyzed for soil moisture. The gridded relative bias maps presented in Figure 3.20 show that for all cases, soil moisture tends to be underestimated with respect to the truth. The similarity in spatial patterns between the baseline and "Scaled *a-priori*" runs are due to the spatially distributed parameters of the water balance. Because, again, the "Uniform" run does not have information about the spatial variability of parameter estimates, the patterns that can be observed are due to rainfall alone. Overall, it can be seen that calibration deteriorates the estimates of soil moisture.

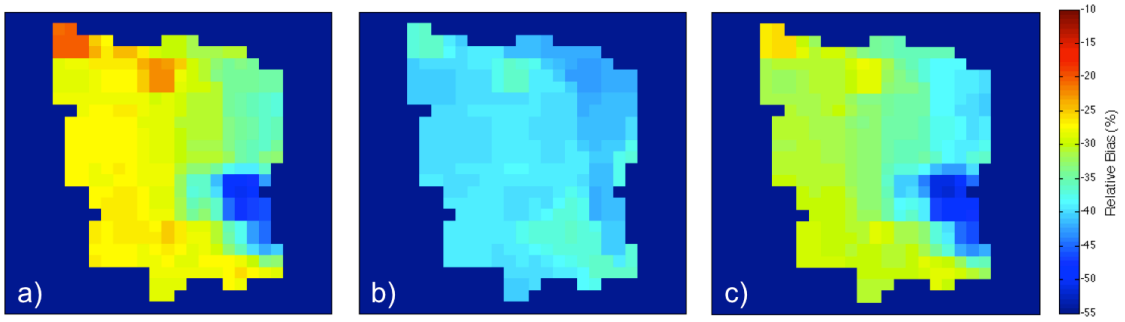

Figure 3.20: Same as Figure 3.16 but for soil moisture.

A pixel-to-pixel comparison and summary of skill distribution over the basin are presented in Figure 3.21. Both plots illustrate that for the "Uniform" run, the lack of spatial variability information has a very negative impact, and for the "Scaled *a-priori*" run, the main issue is the introduction of bias. This particular exercise with soil moisture

also highlights a limitation of model calibration to observations at a basin outlet. Because soil moisture results from the vertical re-distribution of water from rainfall-infiltration processes alone, the estimation of water balance parameters through streamflow data fitting is highly indirect and, thus, prone to introduce significant uncertainty. By the time a given amount of runoff produced at a given location reaches the outlet, the information about its origins is lost due to the natural aggregation of runoff along streams. Moreover, and depending on the size of the basin, the lag times between response observed at the basin outlet and interior locations can be high and variable across the basin.

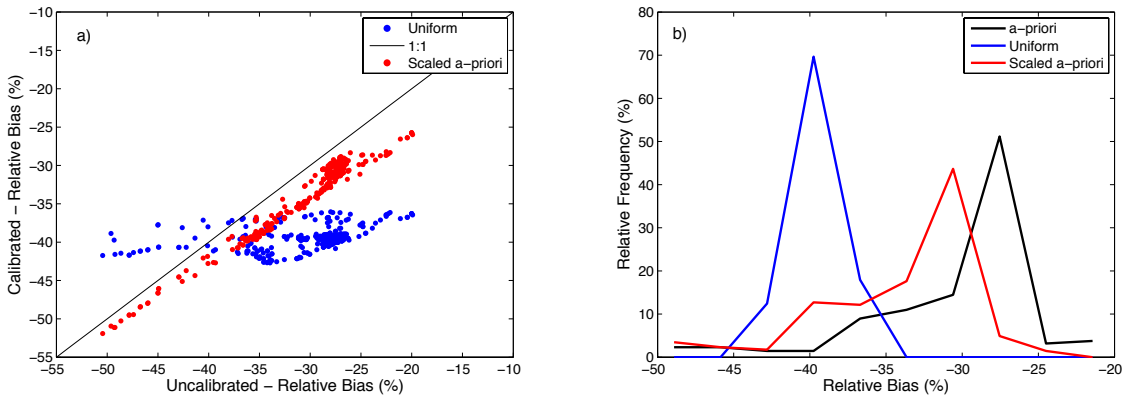

Figure 3.21: Same as Figure 3.17 but for soil moisture.

In an attempt to explain the particular locations of deterioration and improvement with respect to the basin outlet, maps of maximum linear correlation and corresponding lag time were produced for the entire basin (Figure 3.22). Linear correlation between streamflow time-series at any given location and that at the basin outlet was computed for different time lags. The maximum value of correlation and its corresponding lag time were recorded to generate the maps (panels a and b respectively). It can be seen that high correlations with low lag times are found on the main stream. It is also noticeable that as one moves upstream, the correlations decrease

while, intuitively, the lag time increases. These patterns show consistency with the overall patterns of deterioration and improvement previously discussed. However, the variability of patterns among the different skill metrics warrants a more detailed analysis in future work.

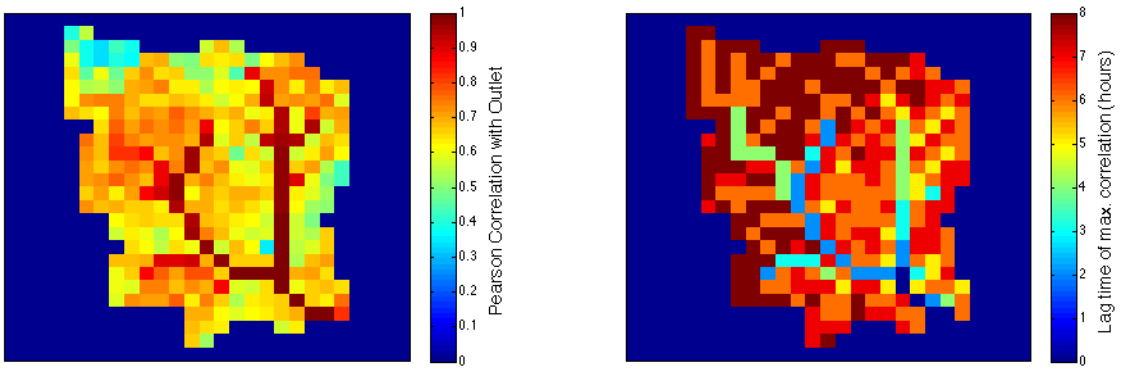

Figure 3.22: Lagged correlation between the outlet and interior points over the basin. The maximum Pearson correlation coefficient (left) is computed for a corresponding lag time (right).

3.5 Summary and conclusions

This chapter described some methodological aspects of the distributed hydrologic modeling strategy employed in this research. Different geospatial datasets describing geophysical attributes of watershed structure were presented. The modeling framework entitled Ensemble Framework For Flash Flood Forecasting (EF5) was briefly described. A detailed description of the hybrid conceptual/physics-based mathematical representation of hydrologic processes embedded in EF5 was presented. Lastly, the calibration-free approach to hydrologic modeling adopted for the main experiments of this research was discussed. Additionally, an Observing-Systems Simulation Experiment (OSSE) was presented to illustrate the impact that data fitting to limited observations used in model calibration can have on the spatial consistency of simulation skill. Main conclusions of this chapter are listed as follows:

- Geographical Information Systems and remote-sensing platforms have made it possible to obtain measurements of the geophysical characteristics of the land surface at useful scales for hydrologic analysis and modeling. Moreover, the availability of these geospatial datasets has increased to cover necessary extents (e.g., continental or even global) to support *large sample hydrologic* analysis. The datasets collected for CONUS herein show significant variability in space, which enables the ability to define explanatory macro scale patterns of hydrologic characteristics.
- The hydrologic model physics available in EF5 are deemed appropriate for the representation of runoff generation and flow routing processes leading to flooding. Particularly, the parameterization of the hydrologic model allows for direct utilization of the collected geospatial datasets that describe watershed physical structure. Moreover, the physical representation of flow routing was shown to satisfy the applicability criterion for the majority of CONUS.
- The calibration OSSE illustrated the negative impact that data fitting at a basin outlet (or at any location in a given watershed) can have on the simulation skill at interior locations, particularly overland. This is a critical result with implications for ungauged prediction applications. Furthermore, the analysis on internal states highlighted the usefulness of describing the spatial heterogeneity of physical structure parameters of the land surface. More importantly, the demonstration in the experiment supports the calibration-free modeling approach chosen for this research.

Chapter 4. Estimating *a-priori* Flow Routing Parameters for Streamflow Simulation over the Conterminous United States

4.1 Introduction

Providing useful estimates of the response of a hydrologic system (i.e., a catchment or watershed) at all locations (i.e., gauged and ungauged) is arguably *The Challenge* in rainfall-runoff modeling. This was the main subject of the past decade-long focus of the International Association of Hydrological Sciences (IAHS) through its Prediction at Ungauged Basins (PUB) initiative (Sivapalan et al. 2003), which, although it promoted scientific productivity, was largely unsuccessful in achieving its main goal (Hrachowitz et al. 2013). The underlying challenge of PUB can be phrased as *how do we generate equally skillful model estimates at all locations regardless of whether there are measurements of the model output or not?* A key aspect involved in this challenge is the regionalization problem in hydrologic modeling, which is primarily concerned with the estimation of parameters at ungauged locations (Beven 2011). The parameters' main role is to enable the versatility of the model in simulating a diverse set of hydrologic processes and responses, thus facilitating the application of the model at all locations.

The estimation of hydrologic model parameters has been the concentration of many studies for over two decades, the majority featuring model calibration techniques (e.g., Sorooshian et al. 1993; Boyle et al. 2000; Duan 2003; Gupta et al. 2003; Vrugt et al. 2006; Vrugt et al. 2008). However, model calibration is a technique primarily developed for lumped hydrologic models. This is because the spatially aggregated conceptualization of processes and parameterization in lumped models makes it difficult

to employ an approach based on characterizations of the spatial variability of the basin physical structure (e.g., topography or soil texture properties, such as hydraulic conductivity). Process-based distributed hydrologic models, on the other hand, are specifically designed to take advantage of the ever-increasing availability of geospatial datasets from geographical information systems and remote-sensing platforms to resolve the dominant spatial patterns of the hydrologic system. Consequently, distributed hydrologic models can be configured using *a-priori* methods for parameter estimation, which are naturally consistent with the PUB challenge and the regionalization problem.

While work on *a-priori* estimates for water balance model parameters based on soil properties have been reported to the literature (e.g. Koren et al. 2000; Yao et al. 2012), few efforts have been devoted to deriving spatially-distributed flow routing parameter estimates without conditioning from calibration. The primary objective of routing models is to describe the space-time evolution of water flow throughout a watershed, catchment or stream network. Moreover, flow routing is essential in the description of flood wave timing, which not only establishes when a flooding event occurs, but also the magnitude and duration of the flood. Flood wave timing is critical in forecasting approaches that rely on threshold-based methodologies for detection (e.g., Reed et al. 2007). Some studies like the ones of Montgomery and Gran (2001) and Finnegan et al. (2005) have analyzed controlling factors of the downstream variability of channel characteristics related to routing parameters. Koren et al. (2004) discuss a methodology in which rating curve data at the basin outlet can be propagated upstream to populate all grids within the watershed with estimates of the flow routing parameters.

However, no study has reported a methodology to estimate flow routing parameters at continental scales.

In this work, the spatial variability of parameter estimates of the kinematic wave model, employed for the distributed flow routing in the hydrologic model described in Section 3.3.1 of Chapter 3, was studied at the continental scale to devise an estimation approach based on regionalization. The choice of a physics-based model (i.e., models formulated from physical laws) is centered on the fact that model parameters are either based on or correspond to actual measurements of the physical system (Boyle et al. 2000), which facilitates the process of *a-priori* estimation. Moreover, the approach used herein to study the spatial characteristics of parameter estimates explores associations with several geophysical properties of the land surface. Using a model whose conceptualization of the physical system significantly departs from reality would prove difficult (if not impossible) to find aforesaid associations. Consequently, the overall goal of this study is to find *a-priori* estimates of kinematic wave routing parameters in order to enable regional forecasting of floods and flash floods at a continental scale with a distributed hydrologic modeling system.

4.2 Derivation of the Kinematic Wave approximation

The one-dimensional form of the Saint-Venant equations relate to the fact that spatial variations of velocity can be neglected both horizontally and vertically across the channel when the interest is in the main direction of water flow (i.e., along the channel). Similarly, the water surface elevation is assumed to be constant horizontally at any section of the channel. In hydrologic applications at the watershed, catchment or stream network scales (e.g., hundreds of meters to a few kilometers), the aforementioned

approximations are acceptable. The Saint-Venant equations are derived from the Eulerian view of motion, where physical laws are applied to the continuum of a fluid as it passes through a control volume. The concept is applied through the Reynolds transport theorem, which relates the time rate of change of a mass-dependent property of the fluid to the external factors causing this change (Chow et al. 1988). Applying the theorem to conservation of mass and momentum, Newton's second law of motion, and neglecting lateral inflow, wind shear and eddy losses, the one-dimensional Saint-Venant equation for continuity is given as:

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = q \tag{4.1}$$

where Q is the flow, A is the channel cross-section area, x is a horizontal distance and t is time. Likewise, the equation for momentum is given as:

$$\frac{1}{A}\frac{\partial Q}{\partial t} + \frac{1}{A}\frac{\partial}{\partial x}\left(\frac{Q^2}{A}\right) + g\frac{\partial y}{\partial x} - gS_o + gS_f = 0 \tag{4.2}$$

where g is the acceleration due to gravity, S_o is the slope of the bottom of the channel, and S_f is the friction slope. Terms in Equation (4.2) above represent the different physical processes governing flow momentum (from left to right): the local acceleration, the convective acceleration, the pressure force, the gravity force and the friction force.

Equations (4.1) and (4.2) above represent the governing equations for one-dimensional, unsteady, open channel flow. Simplifications in the Saint-Venant equations result in different distributed routing models. When equation (4.1) and (4.2) are applied in full (i.e., no simplifications), the method is called a dynamic wave model. When the acceleration (i.e., inertial) terms are neglected in (4.2), the method is called a

diffusion wave model. Finally, if the acceleration and force (i.e., pressure) terms are ignored in (4.2), the method is called a *kinematic wave model*. The kinematic wave simplifications yield $S_o = S_f$, which means that the flow is assumed uniform and, thus, a function of depth or channel's cross-section area alone. Consequently, the form of the kinematic wave equation for momentum becomes:

$$Q = \alpha A^{\beta} \tag{4.3}$$

where α and β are the *kinematic wave model parameters*. Substitution of (4.3) in (4.1) yields an expression for solving for Q as the only dependent variable (Chow et al. 1988; also shown in Equation 3.12):

$$\frac{\partial Q}{\partial x} + \alpha\beta Q^{\beta-1}\frac{\partial Q}{\partial t} = q \tag{4.4}$$

where q is the lateral inflow to the channel.

4.3 Methods for the estimation of the kinematic wave parameters

The standard method to estimate the kinematic wave parameters is based on an assumed channel cross-section shape and the application of Manning's equation, which accounts for the slope and the roughness of the channel (Bedient et al. 2008). Commonly used shapes to model natural streams' channel cross-section are rectangle, trapezoid and the parabola (Dingman 2009). Each of these has explicit functions for the estimation of α and β derived from Manning's equation. A caveat of this method is precisely the need for explicit specification of channel cross-section shape. Because of the mathematical manipulation of Manning's equation, it is difficult to use the actual cross-section shapes of natural streams, which are rather irregular. Moreover, the

assumption of regular shapes, on the other hand, consequently leads to the assumption of prismatic channels (i.e., assuming the entire channel has a constant shape).

An alternative method is based on statistical analysis of rating curve data. Field measurements at stream gauges provide a mean to estimate the parameters α and β directly. Based on the form of the momentum equation shown in (4.3), a power function relating streamflow and channel cross-sectional area can be fitted to data measured in the field (Figure 4.1). The field data needs to encompass a wide range of flows to have a representative sample able to describe the relationship. Usually, the majority of the data comes from flows of low to average magnitudes (although it can also include some significantly high flows), because of difficulties in measuring in the field under flooding conditions (Beven 2011). Nevertheless, this approach offers a way to directly estimate kinematic wave parameters, which implicitly accounts for channel cross-section shape, roughness, and slope.

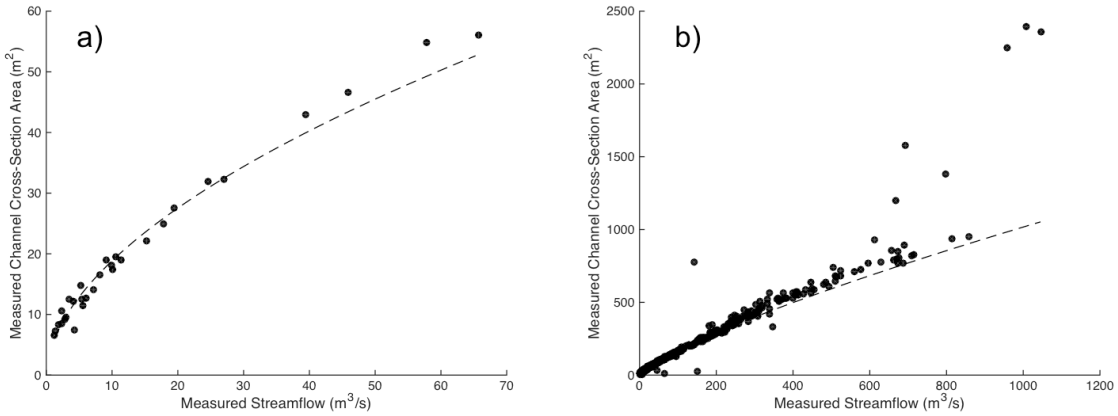

Figure 4.1: Power fit to rating curve data for streamflow (x-axis) and cross-section area (y-axis) measured in the field for USGS stations: a) 01118010 (~531 km^2) and b) 02083500 (~5654 km^2). The dots correspond to the field measurements and the dashed line to the power law regression fit.

This method has been described for the configuration of the HL-RMS distributed model in Koren et al. (2004) and in unpublished work by the Office of

Hydrologic Development (OHD). They present a methodology to propagate the estimates of the rating curve parameters obtained at gauged locations to upstream locations (i.e., ungauged) using several empirically derived geomorphological functions based on drainage area alone. While their results show reasonable skill, their methodology is aimed at estimating routing parameters at the local scale. Additionally, some aspects in their methodology, such as the use of drainage area alone to define the variability of the parameter estimates and the upstream propagation approach, are simplistic and questionable. Intuitively, flow conditions in non-regulated streams (i.e., no regulation or diversion structures) are defined by both local and upstream regional factors and, thus, a downstream approach is preferred.

4.4 Methodology of the *a-priori* estimation

The approach to estimating kinematic wave parameters presented herein is based on the rating curve method described in Section 4.3. The main aspect of the strategy was the investigation of explanatory geophysical factors of the spatial variability of rating curve parameters at a macro scale, with the aim of estimating kinematic wave parameters. This data intensive exercise represents a case of what has been called the "*fourth paradigm of science*" (Hey 2012) and the concept of "*large sample hydrology*" (Gupta et al. 2013). The ultimate goal of this study was to enable river flow routing simulation with a distributed hydrologic model for flash flood forecasting over CONUS without calibration (i.e., without model parameter fitting to a streamflow time-series).

4.4.1 Field measurements of streamflow and channel cross-section area

Using the record of stream gauge stations in the database described in Gourley et al. (2013), field measurement data from the U.S. Geological Survey (USGS) archive were obtained. A series of filtering steps were taken in order to robustly generate an appropriate sample for the statistical analysis of the spatial variability of rating curve parameters. Particularly, the subset of gauges with no evidence of regulation, as explained in Section 3.2 of Chapter 3, was employed herein. Next, an automatic processing script was employed to fit the streamflow and channel cross-sectional area data to a power-law function following Equation (4.3) for each of the selected USGS stations (see example in Figure 4.1). An evaluation of the goodness-of-fit yielded a final sample size of 4,943 stream gauges employed in the analysis of this work.

4.4.2 Watershed characteristics as explanatory variables

As discussed in Chapter 3, the particular characteristics of each stream reach are uniquely determined by several geophysical variables that operate at local and at the watershed scale. Consequently, these geophysical attributes of watersheds were considered as potential explanatory factors of the variability of rating curve parameters.

4.4.3 Multidimensional analysis of kinematic wave parameters variability over CONUS

In this work, the spatial variability of the kinematic wave parameters was analyzed through conditional distribution functions. The sets of α and β distributions were studied using the Generalized Additive Models for Location, Scale, and Shape (GAMLSS; Stasinopoulos and Rigby 2007) technique. The GAMLSS model aims to

simulate the parameters of a distribution of the response variable (i.e., α or β) according to the values assumed by some explanatory variables (i.e., the geophysical characteristics of basins). GAMLSS was chosen over other multidimensional analysis methods (e.g., principal component analysis or a canonical correlation analysis) because modeling the complete conditional distributions enables diagnostic capabilities on the resulting estimates. More importantly, this method explicitly acknowledges the inherent uncertainty of the estimates, which can be employed for probabilistic applications.

Both parameters α and β were analyzed separately following the same approach. To simplify the description of the methodology, the GAMLSS modeling procedure on α alone is explained as follows. Two main assumptions were made: 1) the response variable α is a random variable following a known parametric distribution with density f conditional on the location parameter μ and the scale parameter σ, and 2) the observed α values are mutually independent, given the parameter vectors μ and σ. Each distribution parameter was modeled as a function of the explanatory variable using monotonic (linear/nonlinear or smooth) link functions. More details are provided by Rigby and Stasinopoulos (2001; 2005), Akantziliotou et al. (2002) and Stasinopoulos and Rigby (2007), particularly on the model fitting and selection. It involves identifying a suitable distribution of α, the explanatory variables and the link functions. The estimation method is based on the maximum likelihood principle, and the model selection is carried out by checking the significance of the fitting improvement in terms of information criteria, such as the Akaike Information Criterion (AIC), the Schwarz Bayesian Criterion (SBC) and the generalized AIC (GAIC; Stasinopoulos and Rigby 2007). Forward, backward, and step-wise procedures were applied to select the

meaningful explanatory variables, supervised by diagnostic plots to check the fitting performance, as discussed in Stasinopoulos and Rigby (2007).

A wide variety of distributional forms are available within GAMLSS. A number of conditional two-parameter density functions (lognormal, normal, reverse gumbel, logistic, gamma, etc.) were tested to fit the data. The goodness-of-fit on the whole dataset was checked with the AIC for each of the semi-parametric density fits. The logistic distribution was found to be the most appropriate:

$$f_y(y \mid \mu,\sigma) = \frac{1}{\sigma} e^{-\left(\frac{y-\mu}{\sigma}\right)} \left\{1 + e^{-\left(\frac{y-\mu}{\sigma}\right)}\right\}^{-2} \qquad (4.5)$$

The function above was used to model the conditional α distributions, where the location μ is linked to the expected α value, and the scale σ is representative of prediction uncertainty. After selecting the distribution family, the structure of the model was refined through an iterative procedure by trying several combinations of explanatory variables. The trends for each parameter are fitted using penalized splines, which are more flexible than polynomials or fractional polynomials for modeling complex nonlinear relationships. Lastly, the goodness-of-fit was checked by computing the residuals, first four moments, their Filliben correlation coefficient, and quantile-quantile plots (Stasinopoulos and Rigby 2007). Values of these scores are presented in Table 4.1 below.

Table 4.1: Score values of goodness-of-fit for GAMLSS models for α and β.

Summary of the Quantile Residuals	α	β	Ideal - Gaussian
Mean	0.03	-0.01	0.00
Variance	1.00	1.00	1.00
Skewness	0.38	0.03	0.00
Kurtosis	3.36	3.41	3.00
Filliben Correlation	0.99	1.00	1.00

4.4.4 Hydrologic validation strategy

Additional to the statistical verification explained above in Section 4.4.3, a strategy based on a hydrologic evaluation was employed herein. The methodology evaluates the estimates of the kinematic wave parameters through an assessment of the hydrologic model implementation over CONUS. The hydrologic model was configured with *a-priori* estimates for all of its parameters, as discussed in Chapter 3. This includes seven parameters for the water balance and the excess rainfall routing (subsurface and surface), and the kinematic wave parameters α and β for river routing, the subject of this study (see Table 3.1). Climatological mean monthly potential evapotranspiration data (Koren et al. 1998) were used as part of the hydrologic model inputs. High resolution (1-km/5-min) quantitative precipitation estimation data from the Multi-Radar/Multi-Sensor system (MRMS[**]; Zhang et al. 2011; Zhang and Coauthors 2015) were utilized to force the hydrologic model. A period of 10 years (2002 – 2011) was used to generate simulations of streamflow at a 5-min time step.

An event-based approach to skill evaluation was followed herein. Individual streamflow events were selected with an algorithm that utilizes a threshold value and a hydrograph separation procedure. The algorithm uses a time-series of streamflow, and it returns the hydrograph components for every event in the time-series. The first step in the procedure is the identification of streamflow maxima (i.e., peak flows). The hydrological independence of these maxima needs to be checked. This is, a given event cannot set the initial conditions of a candidate subsequent event. This is common for multi-peak events. The algorithm checks that no peak flow is located within the period

[**]Multi-Radar/Multi-Sensor System (MRMS): http://www.nssl.noaa.gov/projects/mrms/

of direct runoff of another identified peak flow. Additionally, a buffer is established to further assure events are apart enough in time. The buffer here was defined using analytical estimates of the basin's time of concentration. The analytical method is based on an empirical equation for watershed lag time developed by Mockus (1961), which is based on watershed slope, stream length and the NRCS (SCS) curve number. Figure 4.2 presents the estimates of concentration time for the selected basins.

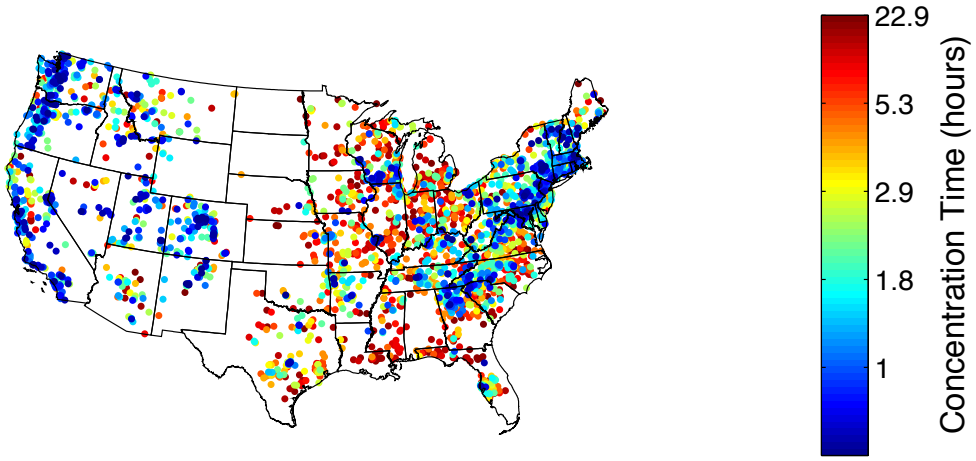

Figure 4.2: Basin analytical concentration time (hours) for selected basins. Color scale is normalized using the data's empirical cumulative distribution.

If two peak flows are determined to be too close (i.e., dependency check is not cleared), only the event with highest magnitude is kept. Direct runoff is computed after baseflow has been determined. For baseflow separation, the Local Minimum Method (Pettyjohn and Henning 1979; Arnold et al. 1995) is employed. An event was defined as that exceeding the 90^{th} percentile flow value of the historical record at each gaged location. Figure 4.3 presents an example of the automatic hydrograph separation.

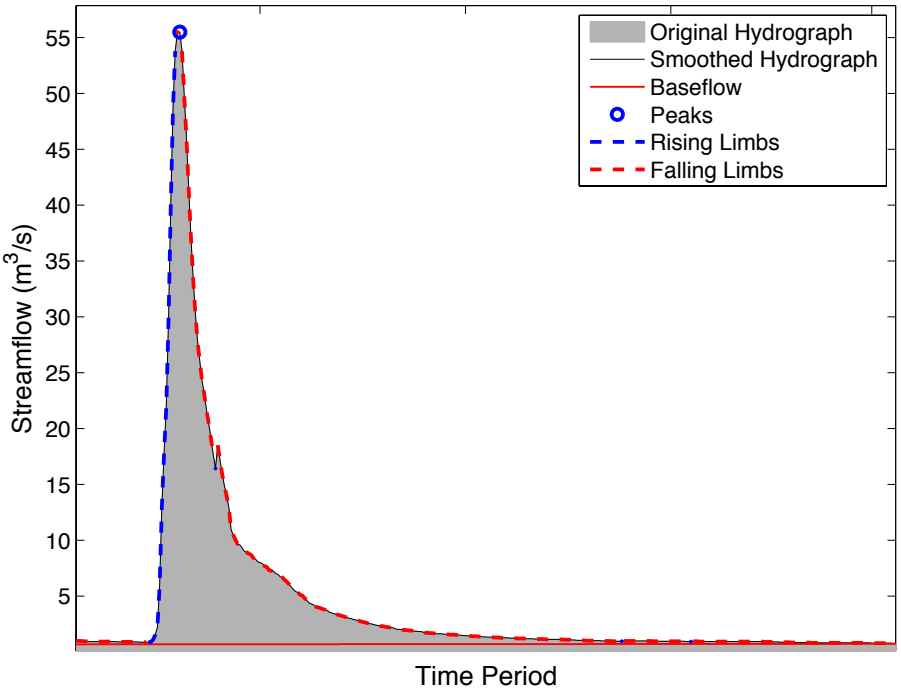

Figure 4.3: Sample output of the automatic event selection based on hydrograph analysis.

The evaluation employed stream gauge stations with no regulation and a drainage area less than 1,000 km^2, which is a representative scale for the majority of drainages over CONUS (> 95%; Figure 3.4). The aforementioned procedure for event selection resulted in an evaluation sample consisting of 75,496 events from 2,680 basins. This filtering was performed in order to reduce the impact of uncertainty from sources unrelated to the estimation of kinematic wave model parameters. Naturally, not all sources of uncertainty can be effectively neglected or accounted for. However, the quantitative approach to skill evaluation employed herein is able to target specific signatures of the modeling of flood wave routing. Two metrics to assess the skill of the simulations were used in these experiments: Relative Peak Error (Equation 2.7) and Peak Time Error (Equation 2.8). In Chapter 2, it was demonstrated the use of these two metrics allows one to disentangle the impact of rainfall and flow routing uncertainty.

4.5 Discussion of modeling results

4.5.1 Association of α and β with watershed geophysical characteristics

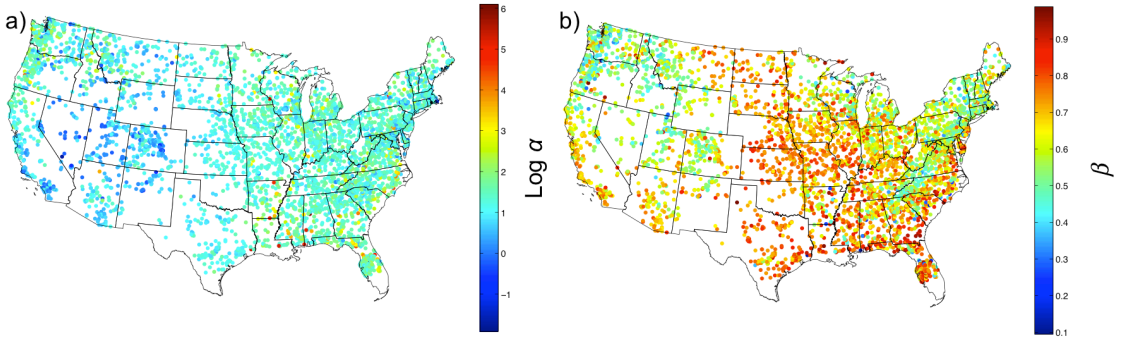

Figure 4.4: Spatial distribution of rating curve parameters for the catchments of the selected USGS stream gauges over the CONUS: a) α in log scale; and b) β.

Figure 4.4 presents the values of the rating curve parameters from all selected USGS stations over CONUS. An initial visual assessment of the spatial variability of both parameters reveals distinct patterns associated with the hydro-climatology and topography across the CONUS. Specifically, α variability appears correlated with the mean annual precipitation and β shows a strong association with relief ratio (Figure 4.5a). The β parameter also presents features corresponding to some clusters observed in the mean rock volume percent. This is consistent with findings of Finnegan et al. (2005) in relation to the scaling of channel geometrical characteristics, depending on the material in which the channel is developed.

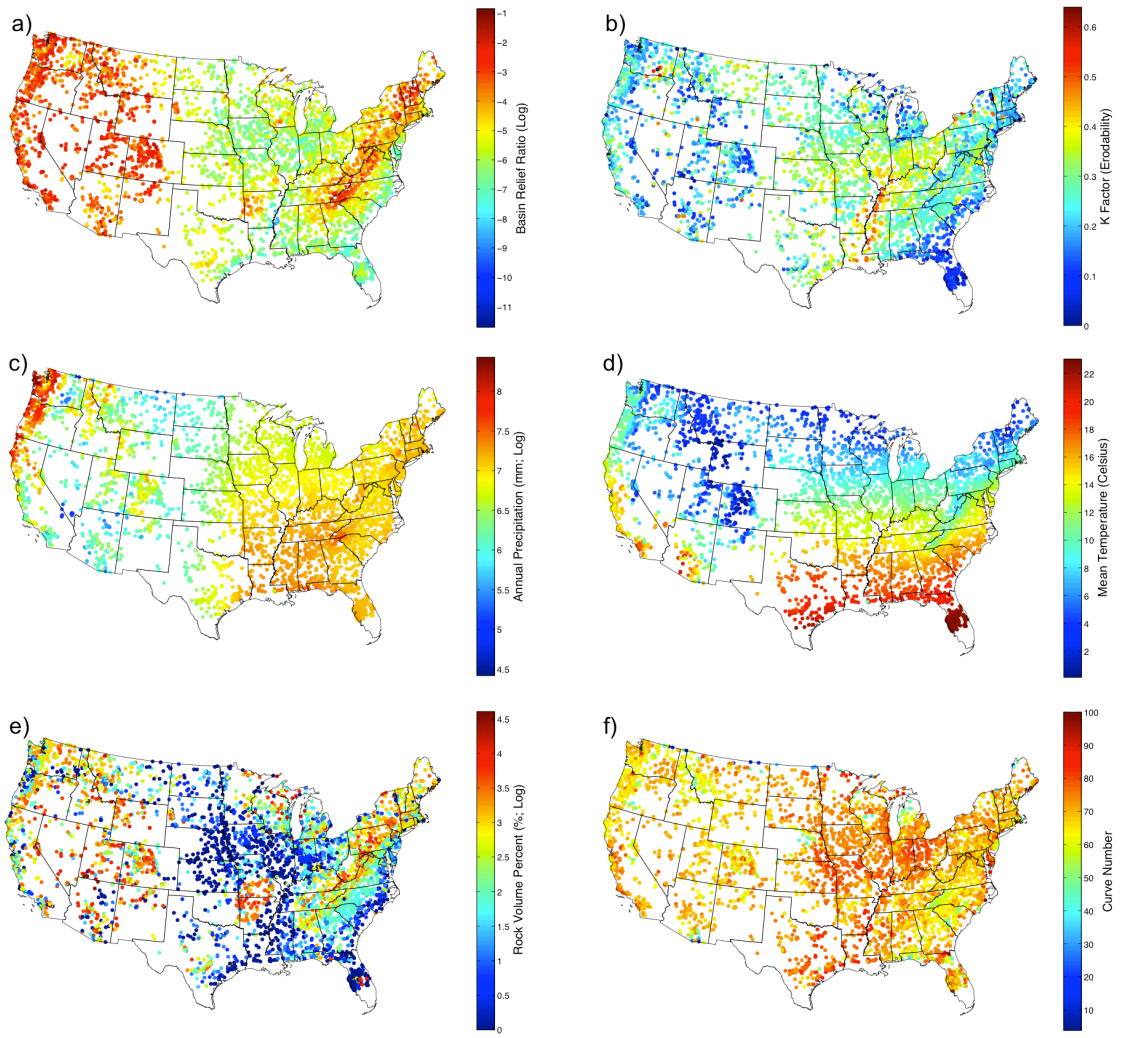

Figure 4.5: Sample of geospatial datasets used in the analysis of spatial variability of rating curve parameters: a) Relief ratio (log scale); b) K factor (Erodability); c) Mean annual precipitation (log scale; mm/year); d) Mean temperature (Celsius); e) Mean rock volume percent (log scale; %); and f) Runoff Curve Number.

Scatterplots illustrating the aforementioned associations are presented in Figure 4.6. The scaling effect of drainage area on the α parameter is arguably not surprising, given its well-known relationship with channel width used in fluvial hydraulics (Montgomery and Gran 2001; Dingman 2009). An interesting feature, however, is the conditioning of this scaling by the hydro-climatology of the basins. Likewise, the relationship between the β parameter and relief ratio shows dependency on the mean rock volume. Further analysis of associations between the rating curve parameters and

geophysical characteristics was performed through 2-D and 3-D methods, such as density-colored scatterplots. However, it was not possible to observe additional significant relationships because the conditioning of the associations, which are a consequence of the interactions of several geophysical factors considered, needs to be assessed through high-dimensional analytical methodologies, such as GAMLSS.

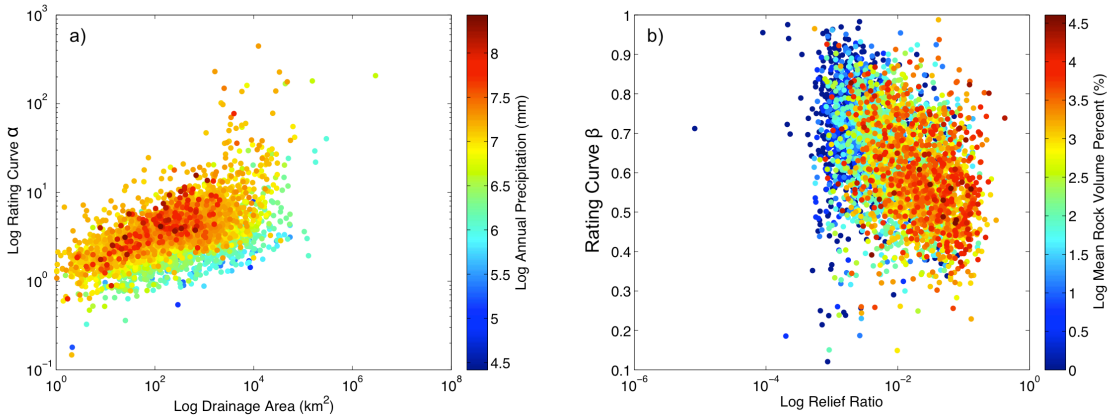

Figure 4.6: A sample of the obtained results from the analysis of associations of kinematic wave model parameters to geophysical variables. Only variables displaying strong associations are included.

4.5.2 Multi-dimensional modeling with GAMLSS

The GAMLSS model was constructed following the methodology explained in Section 4.4.3. The geophysical variables retained by GAMLSS and their corresponding statistical significance values are presented in Table 4.2. The model identified several of the important factors that were discussed in the simpler 2-D analysis discussed above in Section 4.5.1. Drainage area, relief ratio, rock volume and the hydro-climatic variables are highlighted by their significance levels (i.e. probability of rejection of zero). This can be interpreted as a sign of robustness of the GAMLSS model.

Table 4.2: Statistical significance of explanatory variables in GAMLSS model. Not retained or not considered variable are marked with '-'. Significance is expressed as a probability of rejection.

Variable	Alpha	Beta
Basin Area (km^2)	0	-
Elongation Ratio	0.001	-
Relief Ratio	0	0
Slope Index	0.001	-
Slope to Outlet	0.001	0.001
Annual Precipitation (mm/yr)	0	0
Mean Temperature (Celsius)	0	0
K Factor (Erodability)	0	0
Depth-to-Rock (cm)	0.001	-
Rock Volume (%)	0	0
Soil Texture (*b* parameter)	0.05	-
Curve Number	0.001	0
River Length (m)	-	0

The goodness-of-fit of the resulting model is shown in Figure 4.7. Overall, GAMLSS displays skill to predict the values of α and β, as indicated by their correlation coefficient values of 0.73 and 0.63, respectively. However, significant inaccuracies can be observed on the upper end of the rating curve for α and the lower end of the rating curve for β. An investigation of the rating curves associated with these estimates revealed a flow rate-dependent hysteresis at the corresponding gauged locations. The methodology followed herein for the fitting of rating curves does not account for this behavior and, thus, the estimates of the power-law regression parameters will have significant uncertainty. Moreover, the conditions that need an elaborate description of the hydraulics in an open channel (e.g., dynamic wave model) are out of the scope of the flow routing modeling subject of this work.

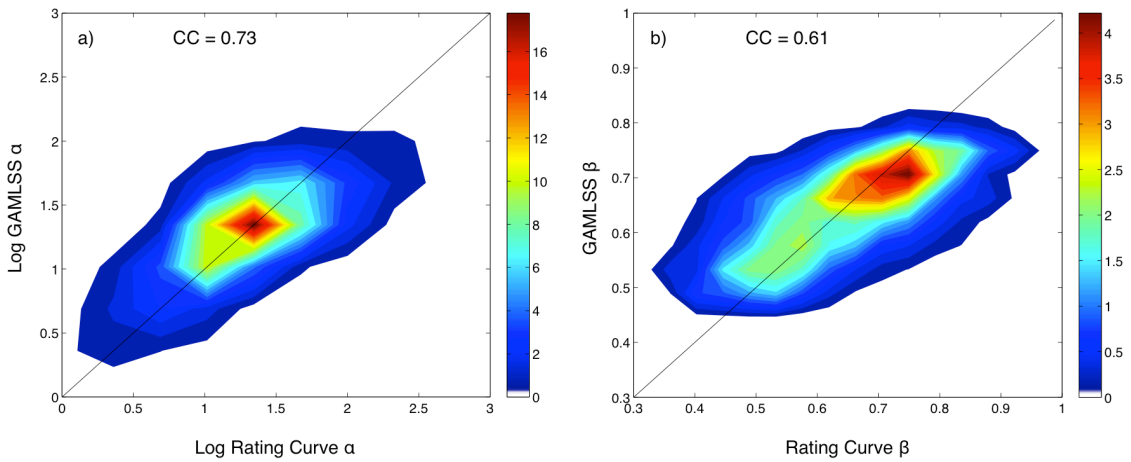

Figure 4.7: Evaluation of the *goodness-of-fit* of the GAMLSS model estimates of kinematic wave model parameters α (left) and β (right). Color-scale represents the data density.

The model fit with GAMLSS was employed to produce 1-km grids of the kinematic wave parameters over the CONUS. Each of the geophysical variables used in the analysis was available over the entire computational grid for which the hydrologic model was configured. Some of the ranges of the explanatory variables for the prediction dataset are larger than those for the training dataset (Table 4.3). The methodology, however, allows for a supervised extrapolation that was implemented herein. The association between each explanatory variable and the response variable was examined along the entire range of available values. If the trajectories show indication of stability toward the boundaries, then extrapolation is deemed acceptable. Otherwise, the boundaries of the training dataset are enforced.

Table 4.3: Explanatory variables retained by GAMLSS. The minimum, mean and maximum values of each variable are included for the training and prediction datasets.

Variable	Training Dataset			Prediction Dataset		
	Min	**Mean**	**Max**	**Min**	**Mean**	**Max**
Basin Area (km^2)	1	2,421	2,926,080	0.71	804	3,138,200
Elongation Ratio	0.262	0.819	2.718	0.197	1.104	7.899
Relief Ratio	$8x10^{-6}$	0.022	0.421	0	0.020	1.099
Slope Index	$2x10^{-5}$	0.012	0.375	0	0.032	1.417
Slope to Outlet	$2x10^{-4}$	0.023	0.208	0.000	0.037	3.005
Annual Precipitation (mm/yr)	121	1,053	4,463	$2.8x10^{-3}$	792	5,675
Mean Temperature (Celsius)	0.0	11.0	22.9	-5.5	11.0	25.5
K Factor	0.000	0.256	0.640	0.000	0.259	0.640
Depth-to-Rock (cm)	9	130	176	9	125	191
Rock Volume (%)	0	12	100	0	14	100
Soil Texture (*b* parameter)	2.79	5.29	11.55	2.79	5.49	11.5
Curve Number	8	70	92	0	70	100
River Length (m)	10,071	68,879	5,282,430	638	10,506	5,440,000

Figure 4.8 presents samples of the *a-priori* estimates of the kinematic wave parameters α and β and their corresponding grids of standard deviation. The main spatial patterns observed on the grids clearly correspond to climatology of precipitation and relief (cf, Figure 4.8a and 4.8b with Figure 4.5a and 4.5c). A closer examination of the α grid also shows the influence of catchment size, as indicated by high values at large streams. This is consistent with the analysis on geophysical characteristics discussed in Section 4.5.1 above. Additionally, it can be observed that the estimates have low standard deviations, which indicates that the GAMLSS model has good precision.

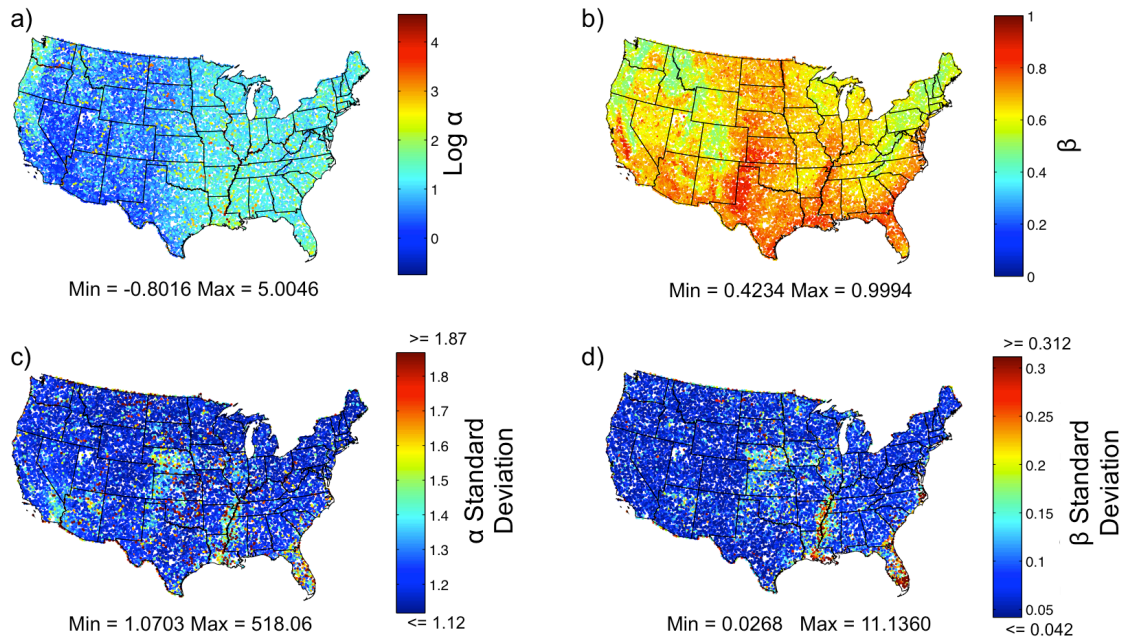

Figure 4.8: Samples of a) α *a-priori* estimates grid and b) β *a-priori* estimates grid, c) standard deviation of α *a-priori* estimates and d) standard deviation of β *a-priori* estimates. Standard deviation colormaps are stretched to 2% and 98% percentiles.

Some regions display noticeably higher standard deviation such as in Nebraska, northwestern Kansas, Iowa, Illinois, the Mississippi valley, Florida, and southern California. Locations with significantly higher deviations are generally scattered, although some clusters can be observed for the β estimates over Florida, the Mississippi valley and on the coast of North Carolina. Visual inspection of the maps of the different geophysical variables points to flat areas (where the kinematic wave model may not apply) and sandy soils as possible factors for this variability in the estimates. A rigorous and elaborate analysis of this particular aspect of the estimation should be performed in future works to understand these specific factors of uncertainty.

4.5.3 Evaluation of hydrologic simulations

Discussion on event-based evaluation and flow routing signatures

Streamflow at any given location (e.g., an outlet) results from the convolution of flood wave routing of upstream reaches. Therefore, the analysis herein on streamflow simulation is representative of the integrated impact of the estimates of the kinematic wave parameters. A sample of the simulation of streamflow events demonstrating model skill and different signatures of the simulated flood wave routing is presented in Figure 4.9. The events were selected from a historic group of floods occurring in September of 2009 in the southeast of the United States, where eleven fatalities resulted from flash floods and floods and a total of $270M USD of damage occurred (NWS 2010). In general, the hydrologic model with its *a-priori* configuration (i.e., no calibration) shows good skill in reproducing the hydrologic response to rainfall in each of the cases. The variability in the magnitude and timing of the peaks is due to uncertainty from several sources, including those in radar rainfall estimates and the hydrologic model itself.

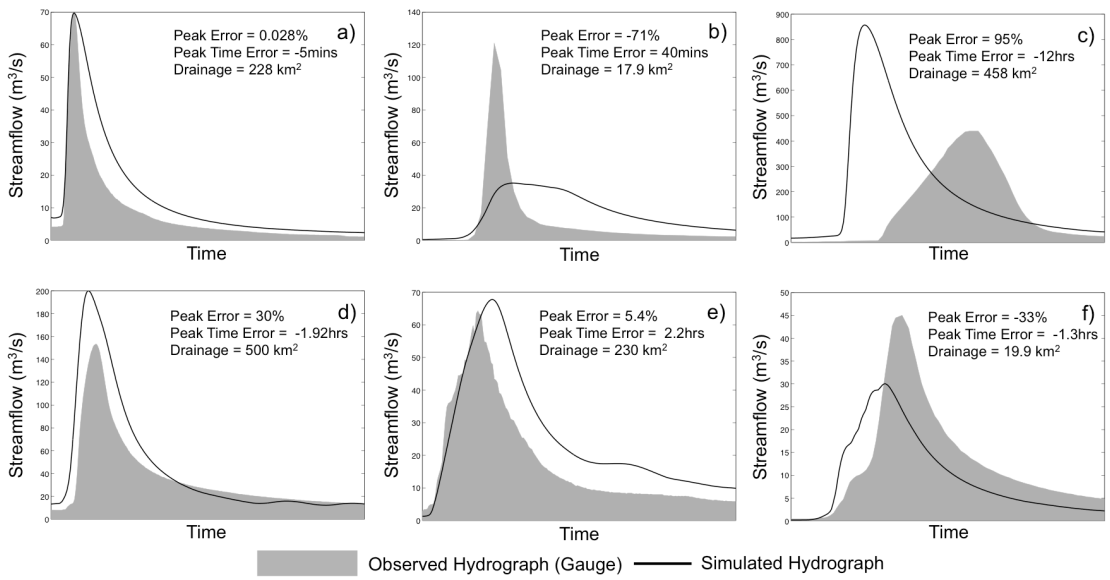

Figure 4.9: Sample hydrographs showing different simulated flow routing skill signatures. The hydrographs correspond to events occurred during September of 2009 on the Southeast of the United States: a) near perfect routing (Mississippi), b) late and low peak (Arkansas), c) early and high peak (Tennessee), d) slightly early and high peak (Tennessee), e) late and high peak (Georgia) and f) slightly early and low peak (near Atlanta, Georgia).

General signatures of flow routing modeling in streamflow hydrographs can be described with the cases shown in Figure 4.9. Early and high (overestimated) peaks indicate that, overall, the flood wave is routed too fast (panels c and d), displaying a tendency for "flashy" responses. Late and low (underestimated) peaks indicate that the flood wave is routed too slowly (panel b) and shows attenuated responses. Both types of model behavior have an impact on the detection and prediction of floods in systems that rely on flooding thresholds: too fast flow routing will tend to over predict the occurrence of floods (i.e., increased false alarm rates), while too slow flow routing will tend to under predict the occurrence of floods (i.e., increased miss rates). In addition to the aforementioned cases, there are events that display strong signatures of the interaction between uncertainty in the runoff generation component (i.e., the water balance) and in the flow routing. In panel e) of Figure 4.9, there is overestimation of the

magnitude with a late peak, which indicates overestimation of excess rainfall in combination with a slow flow routing. On the other hand, panel f) shows a case where the peak is underestimated but occurs early, which indicates underestimation of the excess rainfall and fast flow routing. Lastly, the "ideal" case is presented in panel a) with a near perfect flood wave timing, although minor overestimation of the total volume can be observed.

Event-based evaluation over CONUS

Taking into consideration the aspects discussed above, an evaluation of the 75,496 events from the selected 2,680 basins was performed. Histograms of peak time error and relative peak error are shown in Figure 4.10. The peak timing obtained from the *a-priori* estimation of routing parameters is remarkably skillful. The peaks tend to be early only 15 to 25 minutes. Moreover, the standard deviation is about 3.7 hours, which represents a skill acceptable for flash flood forecasting. The peak magnitude, on the other hand, tends to be underestimated. Furthermore, its frequency displays significant variability, indicating that high underestimation can occur. Peak magnitude errors are more likely to be related to water balance uncertainty, in which quantitative precipitation estimates from radar can play a significant role. However, routing could also explain some of the magnitude errors of peak flow, as discussed in the preceding section.

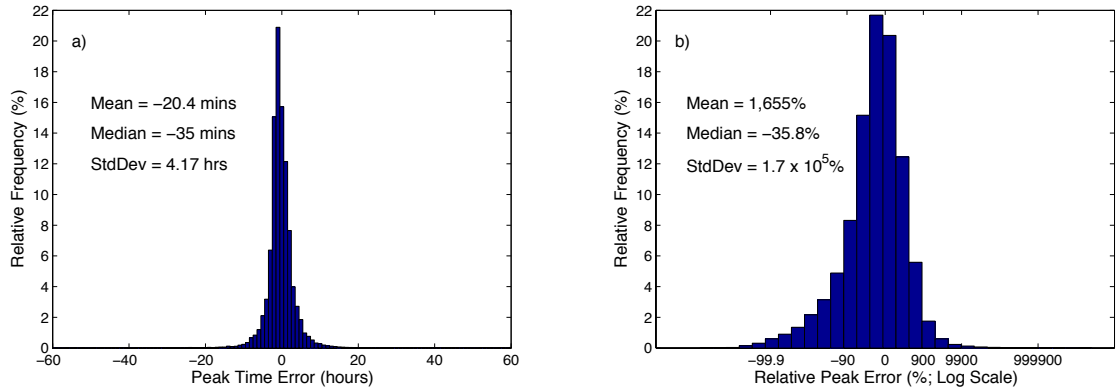

Figure 4.10: Histograms of the a) Peak Time Error (hours) and b) Relative Peak Error (%) for 75,496 events. Measures of location and scale are included for each case.

4.6 Summary and conclusions

In this work, a methodology was devised to generate *a-priori* estimates for the parameters of the widely used kinematic wave approximation to the unsteady, 1-D Saint-Venant equations for hydrologic flow routing. The approach is based on an analysis of the conditional distribution of rating curve parameters over the Conterminous United States given a set of geophysical basin characteristics, including geomorphology, hydro-climatology, pedology and land cover/land use. The main goal of this study was to enable prediction at ungauged locations through regionalization of field measurements for model parameter estimation. Key remarks of this work can be summarized as follows:

- The results of this work demonstrate the value of *a-priori* parameter estimation in a successful configuration of a hydrologic modeling system. The skill of the flow routing simulations, considering that no calibration was performed, is very good for peak flow and timing of peak flow estimation. More importantly, the skill shows consistency, as indicated by the large sample verification. Attaining

such level of skill and consistency is crucial in extending forecasting capabilities to ungauged locations.

- The resulting grids of *a-priori* estimates can be used in any hydrologic model that employs the kinematic wave model for flow routing. Moreover, the methodology presented in this study enables the estimation of the kinematic wave model parameters anywhere over the globe, thus allowing flood modeling in ungauged basins at regional to global scales.
- The approach to parameter estimation featured herein combines the power of large sample hydrology, statistical multi-dimensional analysis, and physical theory to investigate regional and local controls of the spatial variability of channel characteristics that can be parameterized using the rating curve. The results highlight the importance of regional and local geophysical factors in uniquely defining characteristics of each stream reach conforming to physical theory of fluvial hydraulics.
- An important aspect of this approach is its consistency with the scale of flood and flash flood modeling (commensurability). Furthermore, it addresses challenges in standard methodologies that rely on information whose availability might not be adequate for regional to global modeling, and whose scale is not explicitly resolved at the scale of the application.

Overall, this contribution illustrates the advantages of investigating relationships of model parameters with geophysical variables whose availability, in the form of geospatial datasets, is increasing. The particular exercise on the kinematic wave parameters leaves room for further development in terms of accuracy and adaptability to

different basin physical structures. The latter is specifically needed to extend this work to modeling applications at the global scale. Future research will tackle some of the simplifications of the implementation of the kinematic wave used herein, such as the flow-independent nature of the parameter estimates.

Chapter 5. Modeling the Macro Scale Characteristics of Peak Flow Error for Probabilistic Flood Forecasting

5.1 Introduction

The hydrologic physical system can be considered deterministic from the causality view of physics. However, because there exist limitations in our understanding of the physical system and in our ability to observe and measure it, the methodologies we employ to represent the system and predict its state can only be probabilistic. Because hydrologic models are inexact mathematical representations of the physical system, uncertainty arises in the prediction process. Uncertainty in hydrologic forecasts results from the convolution of limitations in accuracy and precision of the different modeling system components (e.g., model inputs, structure, and parameters). Disentangling and characterizing the impact of uncertainty in hydrologic simulations are necessary steps for methodologies that attempt to improve forecasts. Errors in these forecasts, defined as the level of discrepancy with respect to a reference, are the manifestation of uncertainty. However, errors in hydrologic forecasts can display complex variability in space and time. Characterization of spatial and temporal patterns of the variability of simulated watershed response can help track sources of uncertainty. Interest in the spatial patterns of hydrologic response has been of particular concern in ecological studies that try to understand the correlation of a given species' spatial distribution with environmental factors. An early effort to characterize hydrological regimes in CONUS was done by Poff (1996), where 806 streams were classified based on ecologically relevant measures of flow variability. Wagener et al. (2007) discussed

the importance of catchment classification in providing uncertainty constraints and diagnostic power for model evaluation.

In this chapter, a multi-dimensional analysis and characterization of peak flow error over CONUS is performed in order to model the spatial variability of hydrologic forecast skill. One of the goals of this characterization is to infer where significant uncertainty resides and provide guidance on efforts for improving the modeling system. The other key goal of this exercise is to establish the foundations of streamflow probabilistic forecasting framework.

5.2 Multivariate analysis of errors in peak flow simulation

Using the hydrologic model configured in Chapter 3 and the CONUS-wide 10-year simulation of streamflow completed in Chapter 4, a qualitative and quantitative diagnostic of skill of peak flow prediction is performed herein. Although the exercise includes descriptions of both the peak magnitude and peak timing, focus is placed on the former. As demonstrated in Chapter 4, the peak flow timing uncertainty has been constrained significantly through the multi-dimensional estimation of *a-priori* values for the kinematic wave model parameters (see Figure 4.10). The density of peak flow time errors approximates a Gaussian distribution well centered near 0.0 hours (median error = -35 min). The peak magnitude, on the other hand, still displays considerable uncertainty and bias, thus warranting further analysis.

5.2.1 Study basins sample for training and validation

In addition to the diagnostic evaluation, a multi-dimensional modeling of peak flow error is performed. To this end, the dataset consisting of basins in CONUS with no

regulation and a drainage area less than 1,000 km^2 was divided into two randomly selected samples for training and validation purposes, as presented in Figure 5.1 below:

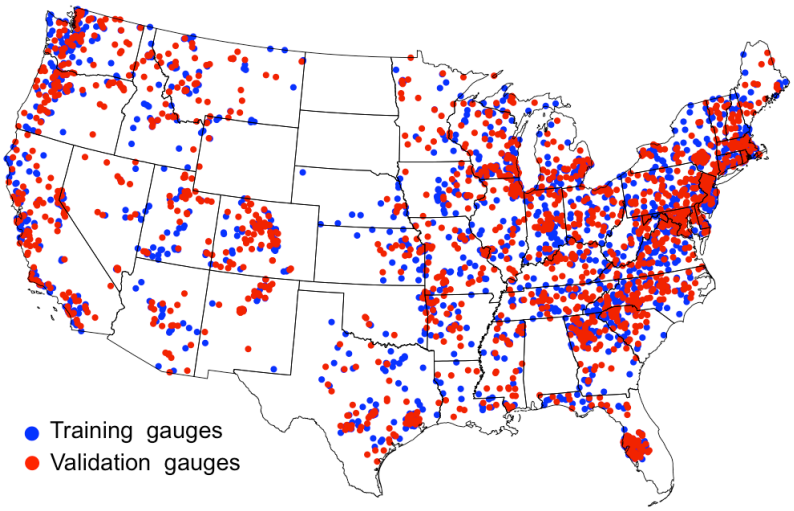

Figure 5.1: USGS stream gauges of the study sites for error analysis

The stations in both samples are well-distributed over CONUS, and appear to have a similar coverage of each region. Figure 5.2 present distributions of several basin geophysical attributes for each sample. It can be observed that the samples have significant consistency, which satisfies an important requirement of the statistical validation strategy. The simulations of streamflow covered a period of 10 years (01/01/2002 – 12/31/2011), as explained in Chapter 4.

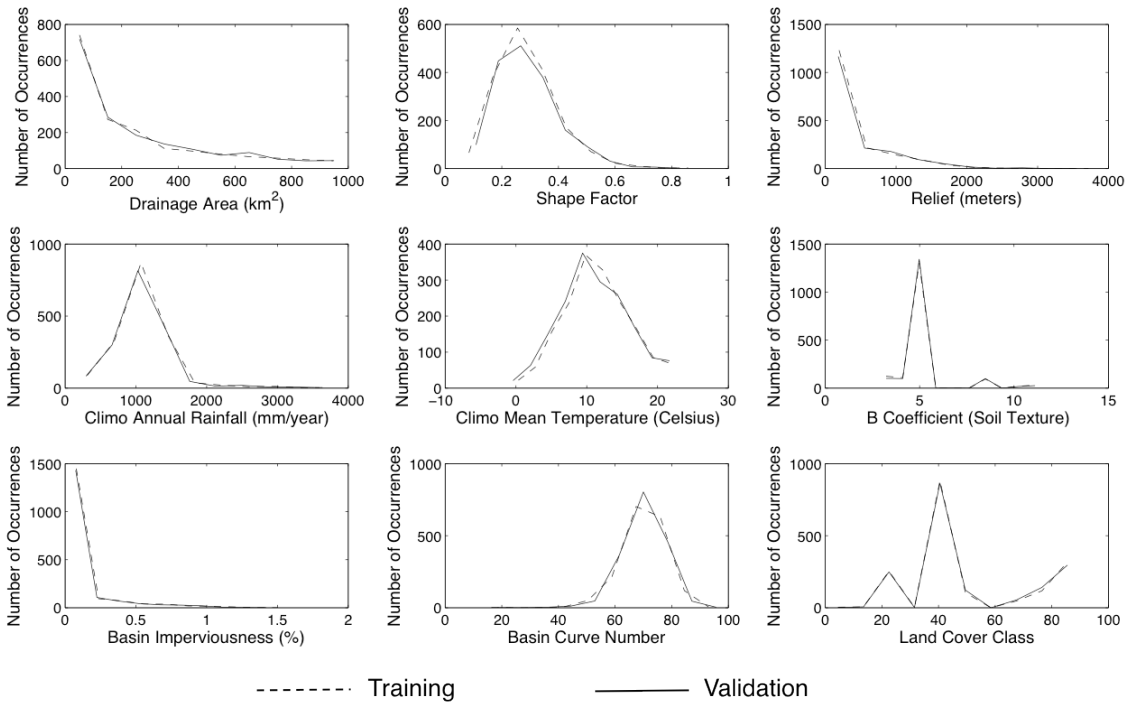

Figure 5.2: Evaluation of consistency of the geophysical factors between training and validation datasets

5.2.2 Multi-dimensional analytical strategy

For the multi-dimensional modeling of peak flow error, a built-in multi-linear regression function from the statistical package of MATLAB® (2011) was employed. The algorithm is similar to the one for GAMLSS employed in previous Chapters. It consists of a systematic method that tests the statistical significance of a given explanatory factor in the error model. The algorithm iterates, including more and less terms in the model, and compares their explanatory power. It uses hypothesis testing based on the *p*-value of each potential explanatory factor and customizable thresholds for acceptance and rejection. The algorithm was employed to objectively test the collection of geophysical factors included and described in Chapter 3.

5.3 Discussion of results

5.3.1 Initial assessment of peak flow simulation skill

A preliminary evaluation of the performance of the hydrologic model in simulating peak flow is presented in Figure 5.3. It can be observed that the model has skill to simulate peak flow overall. However, the scatter density plot (panel a) reveals significant variability in this skill. Moreover, there is evidence in the plot of the impact of issues from the radar precipitation estimates, particularly for low flows that appear "invariant" and almost "zero" in the simulations. Panel b of Figure 5.3 shows the conditional quantiles of simulated peak flow given observed peak flow. The median suggests that simulated peak flow tends to be unbiased with respect to observations, although significant deviations can be noticed for low and high flows. Specifically, low flows tend to be overestimated and high flows tend to be underestimated. It can be also observed that there is higher uncertainty for low flows, as indicated by the wider spreads of quantiles, and that this level of uncertainty decreases with increasing peak flow value.

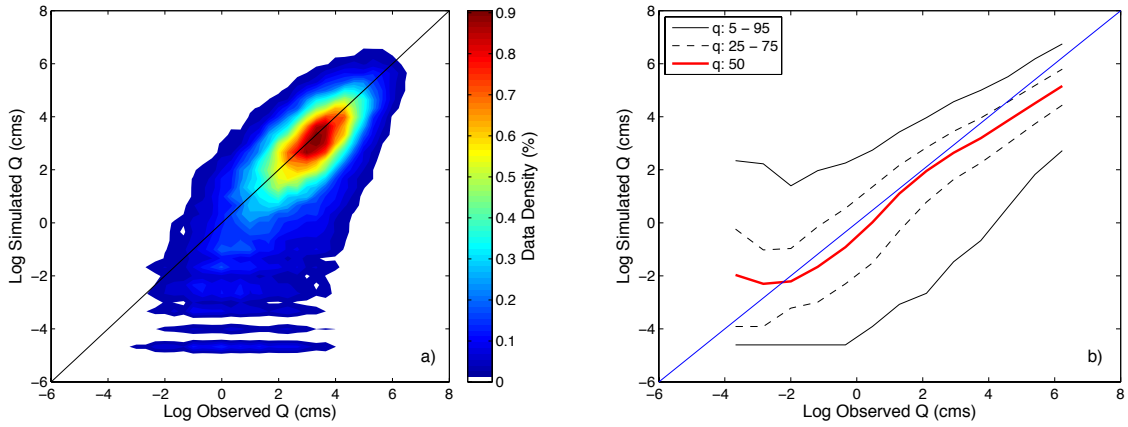

Figure 5.3: Scatter density plot of the observed and simulated peak flows for the 75,496 streamflow events. Values are compared in log scale.

Using this simple diagnostic analysis based on the large sample of peak flow events spanning a 10-year period, it is possible to construct a prognostic model for the probability of an observed peak flow given a simulated peak flow value. This conditional probabilistic model is illustrated in Figure 5.4. Up to this point, the model can only be conditioned on the simulated value of streamflow, regardless of the characteristics of physical structure and response of each basin, and external factors that control uncertainty in the different components of the forecasting system. In the following sections, an exercise to further elaborate the diagnostic analysis is discussed with the objective of finding additional conditioning factors for the probabilistic forecasting model.

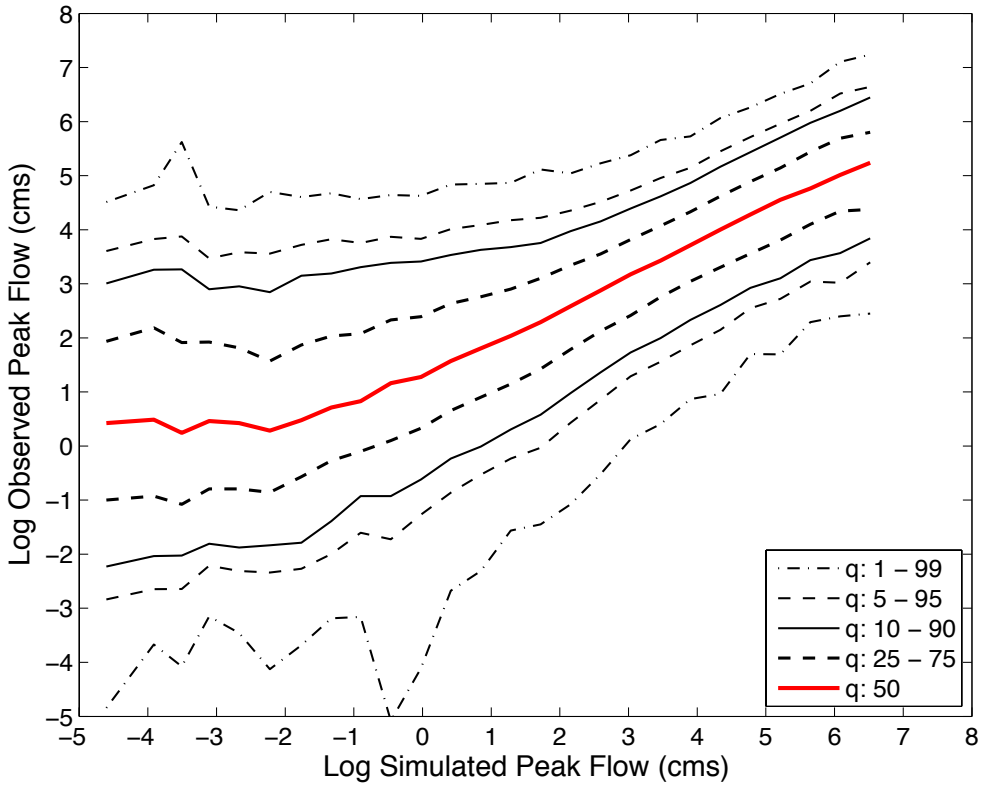

Figure 5.4: Prognostic conditional peak flow probability model.

5.3.2 Peak flow error distribution over CONUS

Distributions of peak simulation errors over CONUS are presented in Figure 5.5 and Figure 5.6. The relative peak error map highlights significant levels of underestimation over the western region of CONUS. Some noticeable underestimation is also observed over the Appalachians and, in general, on the northernmost areas of CONUS. This particular feature points to the impact of topography and temperature on the estimation of rainfall. Significant overestimation can also be seen in central CONUS, Arizona and northern states around the Great Lakes. Overestimation of peak flow can result from a combination of factors, including overestimation of rainfall, underestimation of infiltration, and the flood wave being routed too fast. Although there is not a clear clustering of these locations, as in the case of high underestimation, examination of soil datasets suggest that uncertainty in the simulation of infiltration could explain the overestimation of peak flows. Soils in the hydrologic groups A and B, with sandy texture and a relatively deep profile, seem to be associated with these locations.

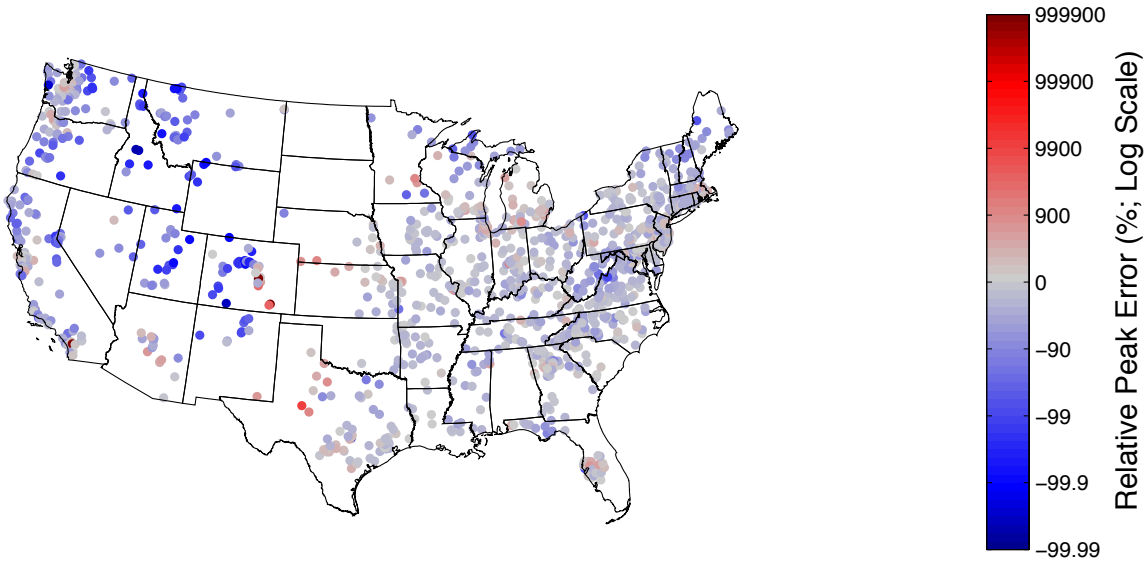

Figure 5.5: Median relative peak error (%) over CONUS in log scale. The median is computed from the total of events identified for each basin.

Some of the basins with high relative peak error are associated with negative peak time error, which indicates that the flow is routed too fast. The locations around the Great Lakes, however, show positive peak time error, indicating overestimated flows being routed too slowly. This is a signature of water balance uncertainty interacting with flow routing, as discussed in Section 4.5.3 of Chapter 4. Overall, and as discussed previously, the peak time skill is considerably higher, as shown by the low error distribution over CONUS. Some of the variability in this skill can be associated with the uncertainty in the kinematic wave parameter estimates (Figure 4.8) and the applicability of the flow routing physics (Figure 3.10). These potential explanatory factors are discussed further in the next section.

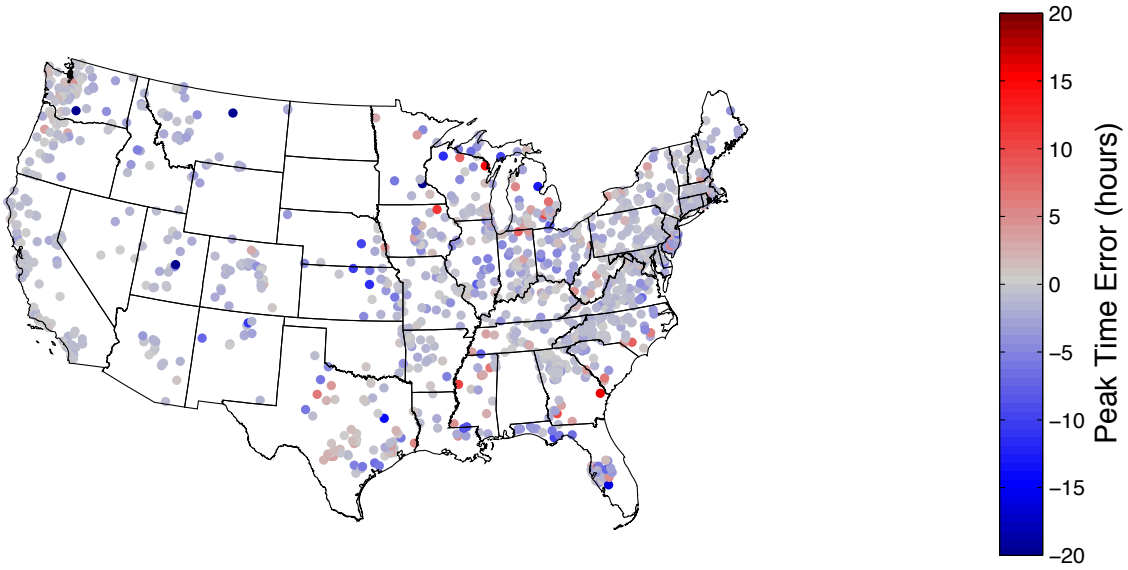

Figure 5.6: Same as Figure 5.5 but for peak time error (hours) in linear scale.

5.3.3 Explaining peak flow errors

The hydrologic model has been configured to produce simulations of streamflow at every grid over CONUS (Chapter 3). However, the physics considered in the model structure specifically represent rainfall-runoff processes near the land surface. This excludes snow storage and melting, frozen soils dynamics, and groundwater dynamics. Likewise, non-weather scatterers and limitations in radar coverage due to the complex topography of CONUS can considerably affect quantitative precipitation estimates (Hong and Gourley 2014). Because rainfall is the main forcing in the hydrologic model, uncertainty in precipitation estimates have arguably the most significant impact on streamflow simulations. Additionally, limitations in the estimation of hydrologic model parameters can also explain the variability of peak flow skill. In the following discussion, the concept of using geophysical parameters and some of their geospatial derivatives to explain uncertainty in peak flow simulations is demonstrated. The analysis is done in terms of the density distribution of peak magnitude and peak timing

errors. Figure 5.7 shows the baseline of these densities, where all 75,496 events are included.

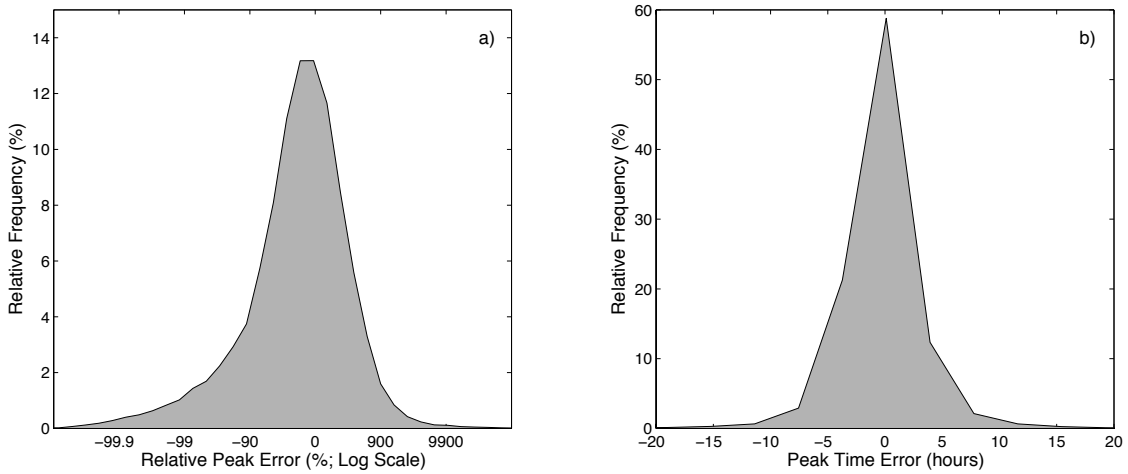

Figure 5.7: Density distributions of a) peak error and b) peak time error for all 75,496 events.

Data quality

As discussed above, the data quality of precipitation estimates can significantly impact the skill of streamflow simulation. To illustrate the impact of uncertainty in precipitation estimates due to radar coverage, the Hybrid Scan Reflectivity Height (HSRH; Figure 5.8) product from the MRMS system suite was employed to characterize peak flow errors. It is immediately evident the high variability in radar coverage over CONUS. In particular, the poor coverage in the Intermountain West is one of the most noticeable features of the plot. Hong and Gourley (2014) report that only approximately 50% of this region is included in the radar network's coverage.

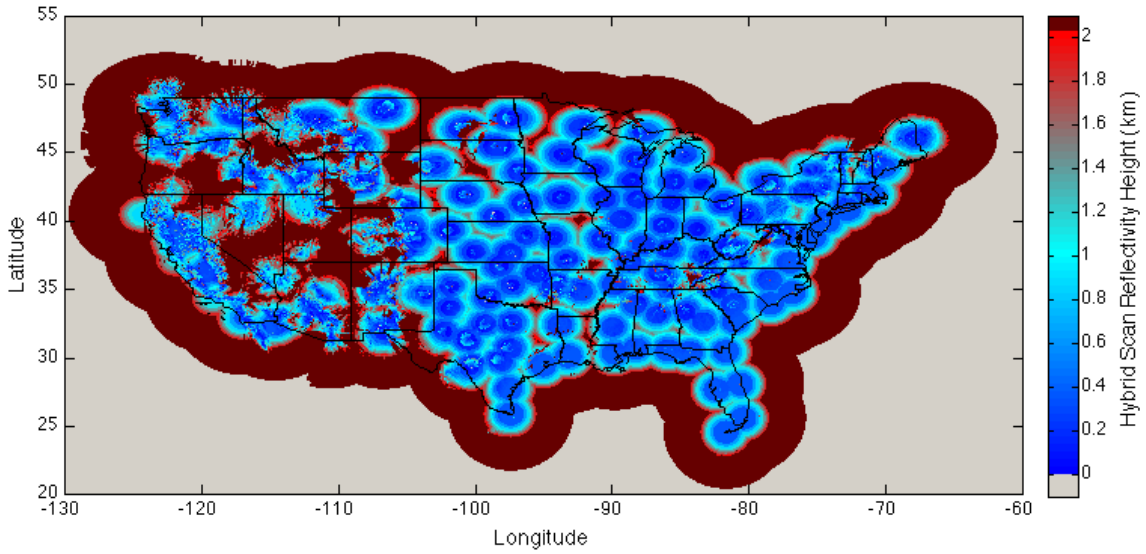

Figure 5.8: CONUS radar network's Hybrid Scan Reflectivity Height (HSRH). Color-scale is stretched to show variability below 2 km. All pixels with HSRH above 2 km are colored with dark red.

The percentage of pixels within a basin with an HSRH below 2 km was computed, and a subjectively chosen threshold of 80% was used to separate basins with adequate coverage from those with poor coverage. Figure 5.9 presents the distributions of peak flow errors for both groups of basins. It can be observed that the peak magnitude tends to be underestimated on basins with poor radar coverage. Additionally, the segregation of a poorly covered basin yields a sample with narrower peak error distribution. The peak time error density, on the other hand, is unaffected by the segregation. This is consistent with the notion that uncertainty in precipitation estimates directly affects the water balance and runoff generation, while flow timing should be virtually unaltered. Moreover, this once again highlights the disentangling attributes of these two error metrics. More importantly, the results in Figure 5.8 demonstrate that radar coverage can explain some of the variability of peak flow skill.

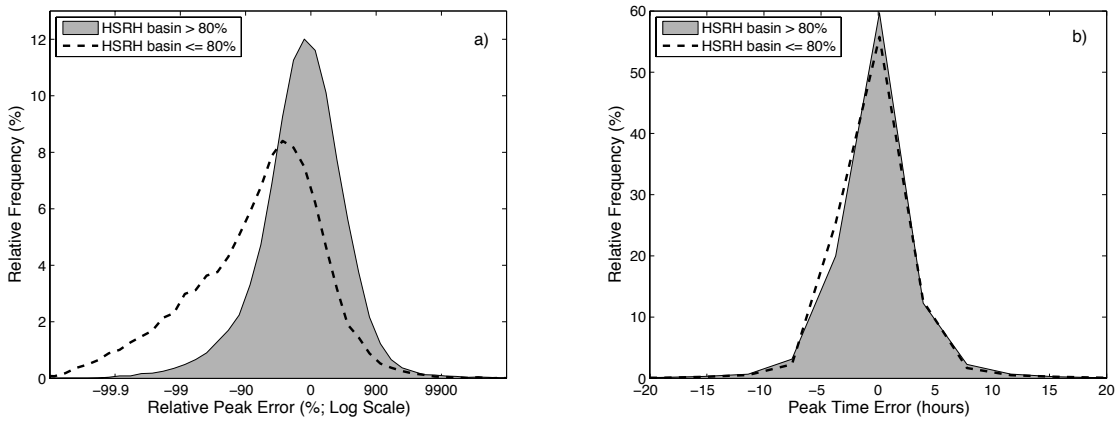

Figure 5.9: Density distributions of a) peak error and b) peak time error conditioned on the percentage of basin with HSRH < 2-km.

Physical representation

Illustration of the impact of the absence of some physical representation in the hydrologic model is done by considering snowmelt-dominated basins. The mean percentage of snow contribution to total annual precipitation was obtained from the Geospatial Attributes of Gages for Evaluating Streamflow (GAGE) dataset (Falcone et al. 2010). Figure 5.10 shows the values of this geophysical parameter for the selected basins. The mountainous regions of the west and the northernmost areas on the east have intuitively the most significant percentages of snow. Interestingly, the strongest patterns of peak flow error shown in Figure 5.5 seem to be associated with snow. It is also worth noticing that although the lack of snowmelt representation of the hydrologic model can explain these errors, there are strong spatial correlations between locations with significant snow and poor radar coverage, since both are at least partially driven by topography.

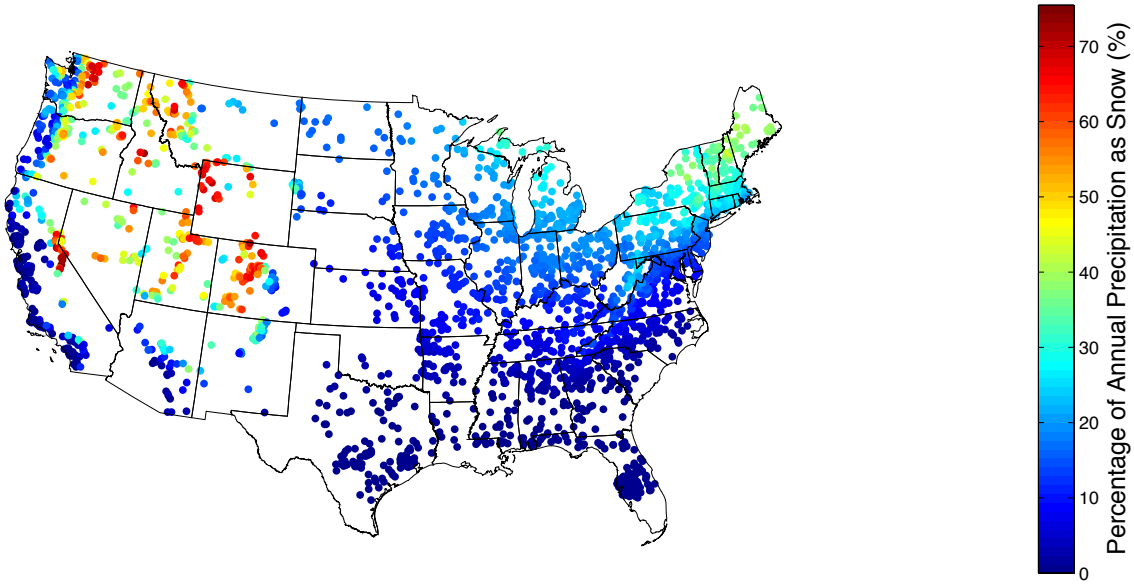

Figure 5.10: Mean percentage of annual precipitation falling as snow for the selected sample of basins.

A subjectively chosen threshold of 30% was used to separate snowmelt-dominated basins from basins with negligible snow influence. Figure 5.11 displays the distribution of peak errors for snow influence in the same fashion as the analysis on radar coverage presented in Figure 5.9. The separation of the distributions of relative peak error is very similar to that observed in the radar coverage analysis. This is once again expected due to the correlation of both parameters. The peak time error distributions are marginally different, also indicating that uncertainty due to the absence of snowmelt physics only significantly affects the water balance.

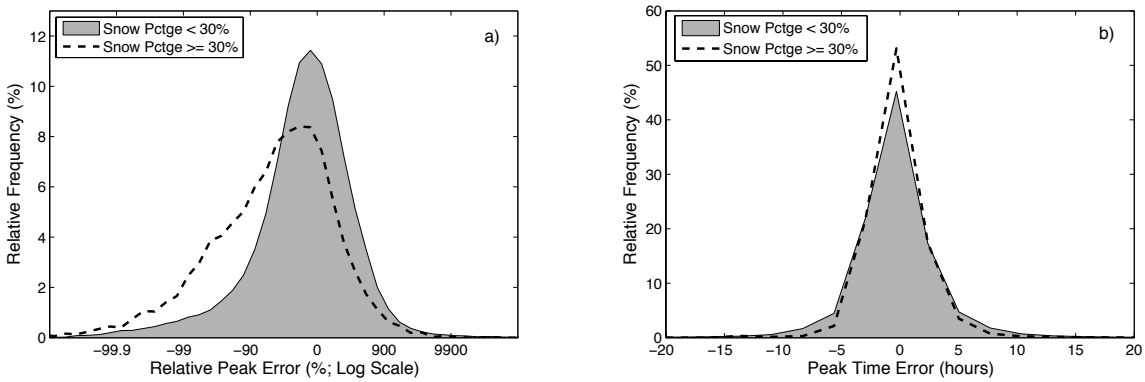

Figure 5.11: Same as Figure 5.9 but for mean snow percentage of annual precipitation.

An additional exercise on uncertainty due to physical representation is performed on the flow routing component of the model. A parameter based on the applicability of the kinematic wave approximation to the one-dimensional unsteady open channel flow equations (Figure 3.10) was devised. The percentage of pixels within a given basin where kinematic wave physics apply was computed, and a subjective threshold equal to 75% was employed to separate basins where the flow routing model is appropriate from those where it is not. Figure 5.12 presents the distribution of peak errors, where it can be seen that mainly the peak time error is affected.

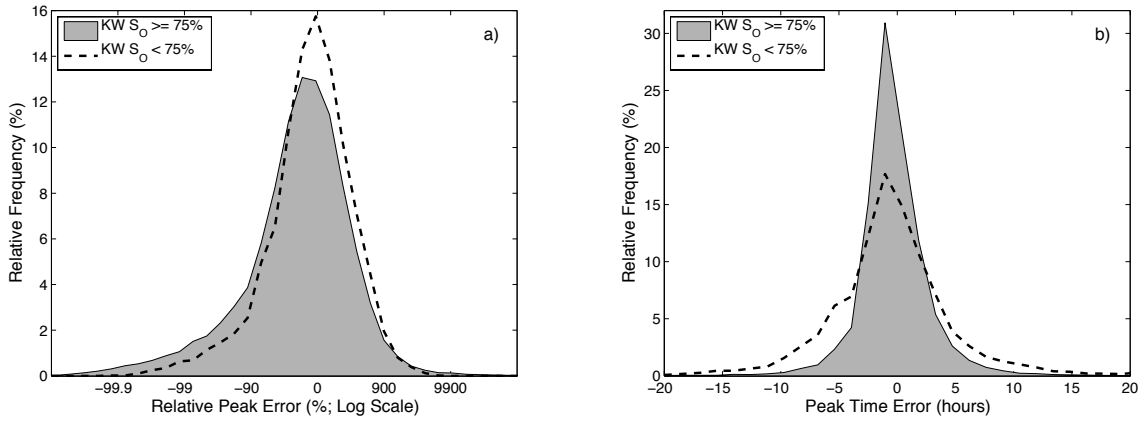

Figure 5.12: Same as Figure 5.9 but for kinematic wave model applicability criterion based on slope.

The results show consistency with the findings in Chapter 2, and illustrate the use of the parameter to explain uncertainty in the flow routing component.

Parameter estimation

As a last demonstration exercise, information related to the estimation of model parameters is employed to explain uncertainty. From the *a-priori* method for the estimation of flow routing parameters described in Chapter 4, it is possible to extract a simple index of uncertainty associated with the supervised extrapolation applied in some areas of CONUS. The index is computed from the accumulation of extrapolation flags assigned to each pixel for each of the variables of the estimation method, and is expressed as a normalized relative value. Figure 5.13 presents a map of the index for all gauged locations (i.e., from the original list of USGS stream gauges). The higher the value of the index, the higher the "amount" of extrapolation performed.

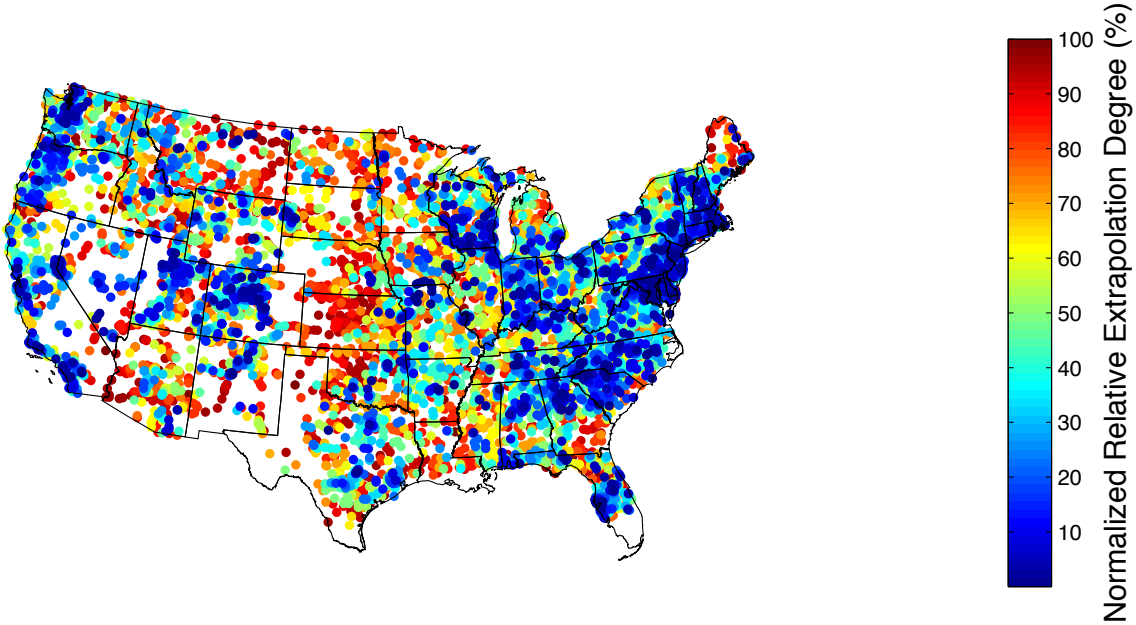

Figure 5.13: Extrapolation degree in kinematic wave parameter estimation.

Using the median of the index, a separation of basins with high degree of extrapolation from those with low degree of extrapolation was performed. Resulting distribution of peak errors are presented in Figure 5.14. Similar to the analysis on kinematic wave applicability, the peak time error displays the separation of the samples. This indicates that the parameter can be used to explain part of the uncertainty in peak time? skill.

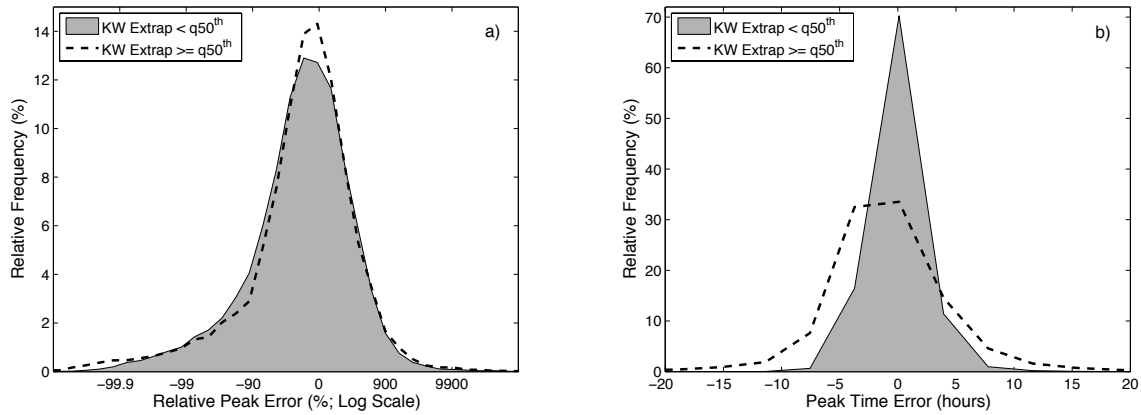

Figure 5.14: Same as Figure 5.9 but for kinematic wave GAMLSS extrapolation.

5.3.4 Multi-dimensional modeling of peak flow error

The first experiment conducted for the error modeling is the fit of a 1-D linear peak error model based on simulated peak flow alone. This model, referred to hereafter as PEM1D, will serve as a baseline for the multi-dimensional peak error model. Figure 5.15 presents the density scatter plots of the reference peak flow ratio and the modeled peak flow ratio as a goodness-of-fit assessment of the PEM1D statistical error model. The evaluation is done for the training (panel a) and validation (panel b) datasets. The training sample consists of 38,707 peak flows from 1,349 basins, while the validation sample consists of 36,789 from 1,331 basins. It can be seen that the PEM1D is able to simulate the peak flow error with a correlation coefficient of 0.64. However, significant

spread in the comparison can be observed, which indicates that the variability of peak errors cannot be explained by the simulated peak flow alone. Very similar features of the scatter are displayed in the evaluation of the validation sample, whose correlation coefficient is slightly higher. This demonstrates the robustness of the regression fit.

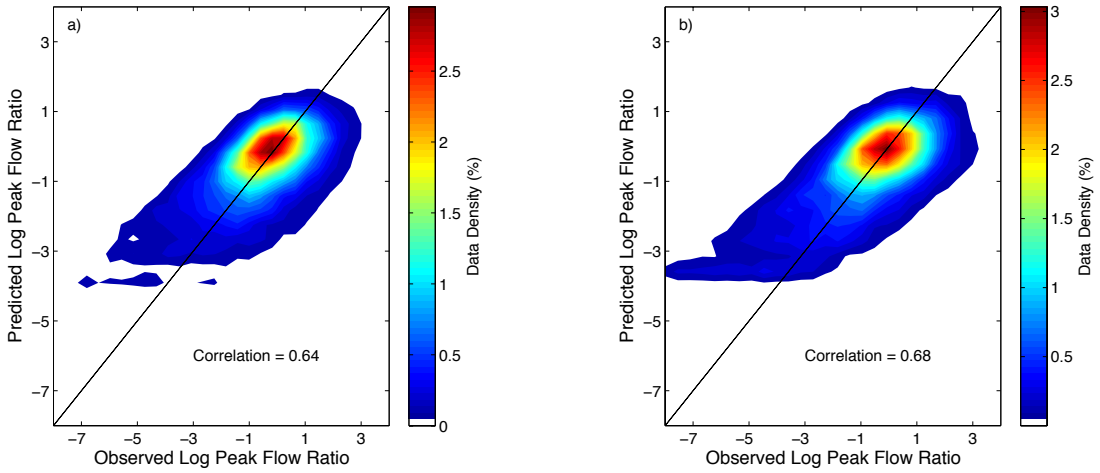

Figure 5.15: Evaluation of the PEM1D (peak flow error model based on simulated peak flow alone): a) training dataset, and b) validation dataset.

Introduction of geophysical factors

The collection of geophysical variables considered in this research was employed in the fitting of the multi-dimensional peak flow error model. This included geomorphologic, hydro-climatic, soil and land surface variables (see Chapter 3 for a discussion of these variables). The snow percentage of annual precipitation, radar coverage, kinematic wave applicability and extrapolation index parameters discussed in Section 5.3.3 were also included. The multi-dimensional error model, referred to hereafter as PEMMD, builds upon the PEM1D. That is, PEMMD consists of the simulated peak flow and geophysical parameters. The initial set of candidate parameters included a total of 40 variables, from which 30 were employed in PEMMD. The two most important variables were the simulated peak flow and the 30-yr mean annual

precipitation. The importance of the simulated peak flow is not surprising, given that the PEM1D could explain the majority of the error. The relevance of the hydro-climatic variable is also expected and consistent with the exercise performed with GAMLSS for the estimation of kinematic wave parameters in Chapter 4. Other variables retained by the algorithm were drainage area, mean temperature, radar coverage, and the kinematic wave applicability, and extrapolation index, among others. The snow percentage of annual precipitation parameter was not retained. This was anticipated in Section 5.3.3 where it was noted the significant correlation that this factor displayed with the radar coverage. Figure 5.16 illustrates the process of fitting the error model and evolution of the goodness-of-fit by incrementally including explanatory variables (i.e., degrees of freedom). The plot shows that the majority of the explanatory power is achieved with approximately five variables.

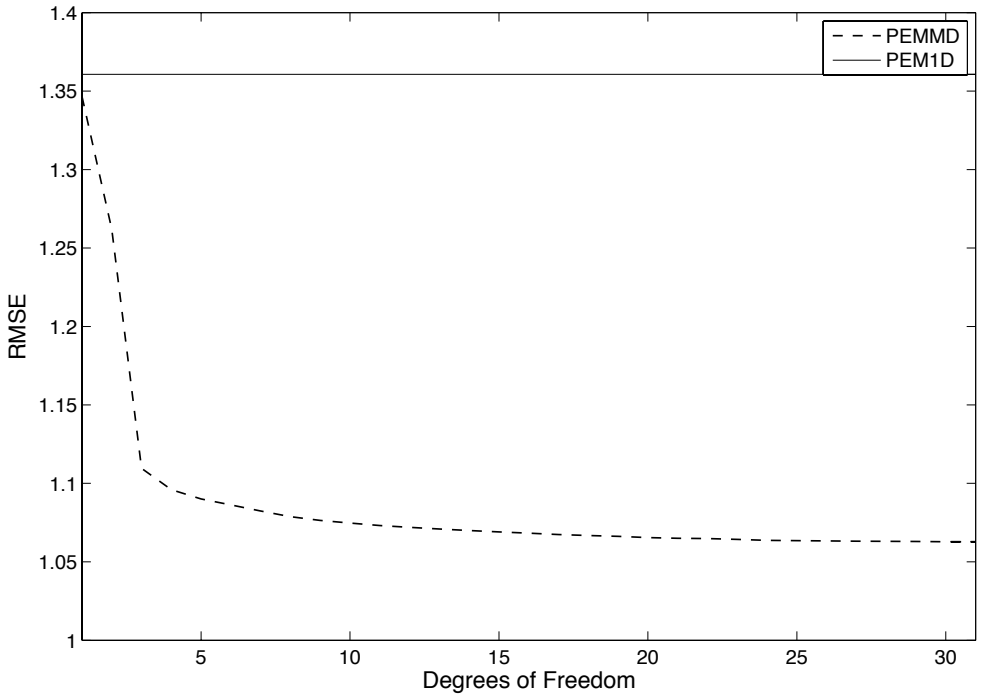

Figure 5.16: Progression of the goodness-of-fit, as indicated by the Root Mean Squared Error (RMSE), during the stepwise multi-linear regression. The RMSE of the PEM1D baseline model is included as benchmark.

Lastly, the evaluation of the PEMMD through density scatter plots of the reference peak flow ratio and the modeled peak flow ratio for both samples is presented in Figure 5.17. It can be observed that the error model was greatly improved by the introduction of geophysical factors. The correlation coefficient increased from 0.64 to 0.8, and the scatter was significantly reduced. Consistent features and level of improvement can be observed for the validation sample.

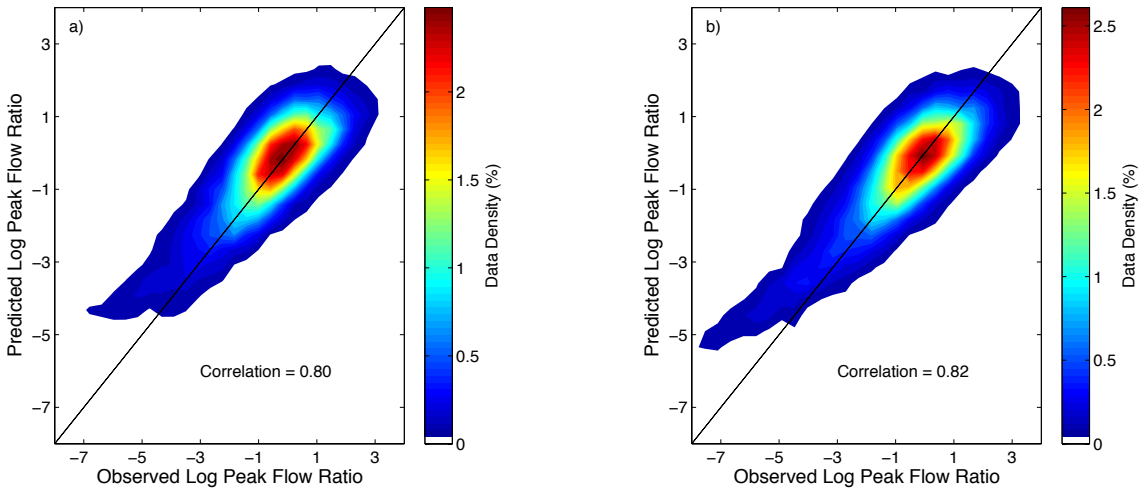

Figure 5.17: Same as Figure 5.15 but adding geophysical parameters to the error model.

5.4 Summary and conclusions

This chapter presented a series of experiments to illustrate the association of geophysical variables and errors in peak flow simulations produced by the hydrologic model described and configured in Chapters 3 and 4. The dataset consisting of 75,496 peak flow events from 2,680 basins over CONUS was divided into a training and validation dataset for a statistical modeling exercise. An initial assessment of peak flow simulation skill was presented, along with an illustration of a conditional probabilistic peak flow model. The spatial variability of peak flow errors over CONUS was presented and discussed. The explanatory power of the parameters of basin physical

structure was demonstrated through simple separation of statistical distribution and using a multi-linear regression algorithm to fit a peak flow error model.

The initial evaluation of the simulation of peak flows showed significant skill of the hydrologic model. This highlights the successful development and configuration discussed in Chapters 3 and 4. The map of relative peak flow error over CONUS displayed clear patterns, pointing to the negative impact that poor radar coverage has on the hydrologic modeling over regions of complex topography. On the other hand, the peak time error did not show any significant patterns, which is consistent with what it was found in Chapter 4.

Clear associations between geophysical parameters and peak flow error were observed in the analysis. Furthermore, an error model based on these geophysical parameters was successfully fitted and validated. This exercise demonstrated the utility of variables that describe macro scale patterns of hydrologic physical structure in explaining uncertainty in peak flow simulation. Moreover, the resulting error model can be used to generate grids of estimated peak flow error over the entire CONUS surface. Such datasets will have the potential of guiding future efforts to improve the hydrologic model.

Chapter 6. Summary, Concluding Remarks and Perspectives

6.1 Dissertation summary

Quantification of uncertainty is becoming a key aspect for operational hydrologic forecasting frameworks due to its usefulness in decision-making situations and growing awareness of the limitations in deterministic systems. However, describing uncertainty is not an easy task due in part to the many challenges associated with its inherent multidimensional nature. In this research, an exercise to characterize the spatial variability of peak flow errors over the Conterminous United States was performed to characterize uncertainty in the hydrologic modeling system. The main goal of this study was to establish the basis for a probabilistic forecasting framework at gauged and ungauged locations.

A large catchment sample approach was employed to perform the required analysis. This kind of data intensive exercise warrants methodologies to automate data analysis and simplify quantitative evaluations. Several algorithms were developed and utilized to satisfy this need. Different geospatial datasets describing geophysical attributes of watershed structure were collected and discussed. The modeling framework entitled Ensemble Framework For Flash Flood Forecasting (EF5) was briefly described, along with a detailed description of the hybrid conceptual/physics-based mathematical representation of hydrologic processes. Also, the calibration-free approach to hydrologic modeling adopted for the main experiments of this research was discussed. To support the argument for the calibration-free modeling approach, an Observing-Systems Simulation Experiment (OSSE) was included to illustrate the

impact that data fitting in model calibration can have on the spatial consistency of simulation skill.

An innovative methodology was devised to generate *a-priori* estimates for the parameters of the widely-used kinematic wave approximation to the Saint-Venant's, one-dimensional, unsteady, open channel flow equations for hydrologic flow routing. The approach is based on an analysis of the conditional distribution of rating curve parameters over the Conterminous United States given a set of geophysical basin characteristics, including geomorphology, hydro-climatology, pedology and land cover/land use.

Lastly, a series of experiments to illustrate the association of some geophysical variables and errors in peak flow simulations was presented. An illustration of a conditional probabilistic peak flow model was included. The spatial variability of peak flow errors over CONUS was also presented and discussed. The explanatory power of the parameters of a basin's physical structure was demonstrated through simple separation of statistical distribution and using a multi-linear regression algorithm to fit a peak flow error model.

6.2 Overall conclusions and remarks

The multi-objective approach to evaluating model performance explored in Chapter 2 proved useful in the disentanglement of the individual impact of the two sources of uncertainty characterized herein. The choice of metrics of skill (or error) and the scale at which they are integrated are critical to properly describe uncertainty and identify its sources. Simplifying the modeling problem by focusing on the dominant processes helps in the selection of these metrics and facilitates the diagnostic process.

Geographical information systems and remote-sensing platforms have made it possible to obtain measurements of the geophysical characteristics of the land surface at useful scales for hydrologic analysis and modeling. Moreover, the availability of these geospatial datasets has increased to cover necessary extents to support *large sample hydrology* analysis. The datasets collected for CONUS herein show significant variability in space, which enables one to define explanatory macro scale patterns of hydrologic characteristics.

The calibration OSSE example illustrated the negative impact that data fitting at a basin outlet can have on the simulation skill at upstream locations. This is a critical result with implications for the use of distributed hydrologic models, particularly with predictions at ungauged locations. Furthermore, the analysis on internal states highlighted the usefulness of describing the spatial heterogeneity of physical structure parameters of the land surface. More importantly, the demonstration in the OSSE supports the calibration-free modeling approach chosen for this research.

The results of this work demonstrate the value of *a-priori* parameter estimation in a successful configuration of a hydrologic modeling system. The skill of the flow routing simulations, considering that no calibration was performed, is remarkable. More importantly, the skill shows consistency, as indicated by the large sample verification. Attaining such level of skill and consistency is crucial in extending forecasting capabilities to ungauged locations. The initial evaluation of the simulation of peak flows showed significant skill of the hydrologic model.

Overall, this contribution illustrates the advantages of investigating the utility of geophysical variables whose availability in the form of geospatial datasets is increasing.

The approach to uncertainty characterization featured herein combines the power of statistical multi-dimensional analysis and physical theory to explain uncertainty in peak flow simulation.

Lastly, results of this research represented evidence to accept the hypotheses stated in Chapter 1. The first hypothesis state that "*if simulated streamflow errors are correlated with at least one of the properties of basins, then at least one of those properties could be used to explain the spatial variability of simulated streamflow errors*", and the second hypothesis state that "*If hypothesis 1 is true and if the association between simulated streamflow errors and the property (or properties) of the basin can be modeled, then simulated streamflow errors could be predicted in basins whose property (or properties) is (or are) encompassed in the training dataset*". Clear associations between geophysical parameters and peak flow error were observed in the analysis. Furthermore, an error model based on these geophysical parameters was successfully fitted and validated.

6.3 Future and ongoing work

The work presented herein consists of a methodology to evaluate and characterize errors in model simulations, which enables probabilistic forecasting capabilities. However, it is worthwhile to discuss that this is only a first step in what would represent a fully probabilistic forecasting framework. Several research endeavors are either underway or planned to be initiated in the near future to address different aspects of the probabilistic forecasting framework.

In terms of modeling uncertainty in the hydrologic modeling system, it is an objective of this research to introduce characterization of the uncertainty as early as

possible in the modeling algorithm. In other words, the overall goal is to model uncertainty explicitly for every component of the hydrologic modeling system. Such a hydrologic modeling strategy represents an analogy to the Probabilistic Quantitative Precipitation Estimates (PQPE) proposed by Kirstetter et al. (2015). Indeed, a Probabilistic Hydrologic Modeling System (PHMS) would necessarily have to be able to directly ingest PQPE.

Ongoing research on a data assimilation approach based on model sensitivity entitled the Forward Sensitivity Method (FSM; Lakshmivarahan and Lewis 2010) might hold the key for a methodology to develop a PHMS. In this ongoing work, an analogy of CREST physics has been developed in which discontinuities in the model structure have been approximated by continuous mathematical functions. In this manner, *first-order sensitivity functions* for the hydrologic model can be computed via differentiation of model equations. Therefore, it is possible to employ functions to estimate probability based on sensitivity, such as in a case discussed by Lewis et al. (2006):

$$P_1(x_1) = \sum_{x_1(i) \in S_{M(x_1)}} \frac{1}{Det\left[D_M\left(x_0(i)\right)\right]} P_0\left(x_0(i)\right)$$

where $P_1(x_1)$ is the probability of the state x (e.g., streamflow) at $t = 1$, which depends on the probability of the state x at the previous time $t = 0$, $P_0(x_0)$, assumed to be known, and $D_M(x_0(i))$ is the model *Jacobian Matrix* with respect to initial conditions. This *Jacobian Matrix* corresponds to the sensitivity equations derived for the hydrologic model, as discussed above.

6.4 Research perspective

Bridging geophysical science, risk management and society to mitigate the impact of natural disasters is perhaps the most important and compelling challenge in this field. As stated in a book published by the International Association of Hydrological Sciences that focuses on the future of water resources: "*The water management sector is one arena where socio-economic factors interact closely with physical and environmental factors, and this clearly needs to be reflected within scientific modeling and political planning*" (Oki et al. 2006). The decision-making process in geophysical hazard applications includes (roughly) the observation of relevant characteristics and process of the associated geophysical phenomena, the representation of the physical system and prediction implementation, the analysis and interpretation of the forecast products, and the communication to emergency managers and the general public. This process is traditionally designed and currently implemented as a series of individual, and sometimes independent, sub-processes that take care of particular tasks. The objective in these discretized systems is to be able to construct a "whole" by putting together what every sub-process produces. However, if no communication between parts at any point of their design, development and implementation takes place, issues in constructing the "whole" arise, impairing the system to fulfill its purpose. It is therefore the perspective of this work to find holistic approaches to address the decision-making process as a continuum.

References

Adhikari, P., Y. Hong, K. Douglas, D. Kirschbaum, J. Gourley, R. Adler, and G. Robert Brakenridge, 2010: A digitized global flood inventory (1998–2008): compilation and preliminary results. *Natural Hazards*, **55**, 405-422 10.1007/s11069-010-9537-2

Akantziliotou, C., R. Rigby, and D. Stasinopoulos, 2002: The R implementation of generalized additive models for location, scale and shape. *Proc.* Proceedings of the Statistical modelling in Society: 17th International Workshop on Statistical Modelling, Chania,

Arnold, J., P. Allen, R. Muttiah, and G. Bernhardt, 1995: Automated base flow separation and recession analysis techniques. *Groundwater*, **33**, 1010-1018

Arnold Jr, C. P., and C. H. Dey, 1986: Observing-systems simulation experiments: Past, present, and future. *Bulletin of the American Meteorological Society*, **67**, 687-695

Ashley, S. T., and W. S. Ashley, 2008: Flood fatalities in the United States. *Journal of Applied Meteorology and Climatology*, **47**, 805-818

Bales, J. D., 2003: Effects of Hurricane Floyd inland flooding, September,ÄìOctober 1999, on tributaries to Pamlico Sound, North Carolina. *Estuaries*, **26**, 1319-1328

Bedient, P. B., W. C. Huber, and B. E. Vieux, 2008: *Hydrology and floodplain analysis* Prentice-Hall,

Beven, K., 2006: Searching for the Holy Grail of scientific hydrology: $Q_t = H(S, R, \Delta t)A$ as closure. *Hydrology and Earth System Sciences Discussions*, **10**, 609-618

Beven, K., 2013: So how much of your error is epistemic? Lessons from Japan and Italy. *Hydrological Processes*, **27**, 1677-1680

Beven, K., and A. Binley, 1992: Future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, **6**, 279-298

Beven, K. J., 2011: *Rainfall-runoff modelling: the primer*. John Wiley & Sons,

Boyle, D. P., H. V. Gupta, and S. Sorooshian, 2000: Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resour. Res.*, **36**, 3663-3674 10.1029/2000wr900207

Briedenbach, J. P., and J. S. Bradberry, 2001: Multisensor precipitation estimates produced by the National Weather Service River Forecast Centers for hydrologic applications. *Proc.* 2001 Georgia Water Resources Conference, Institute of Ecology, University of Georgia, Athens, Ga,

Burnash, R. J. C., R. L. Ferral, and R. A. McGuire, 1973: A general streamflow simulation system - Conceptual modeling for digital computers. In *Report by the Joint Federal State River Forecasts Center, Sacramento, California*,

Carney, M., and P. Cunningham, 2006: Evaluating Density Forecasting Models.

Carpenter, T., and K. Georgakakos, 2006: Intercomparison of lumped versus distributed hydrologic model ensemble simulations on operational forecast scales. *Journal of Hydrology*, **329**, 174-185 10.1016/j.jhydrol.2006.02.013

Carpenter, T. M., and K. P. Georgakakos, 2004: Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed hydrologic model. *Journal of Hydrology*, **298**, 202-221 10.1016/j.jhydrol.2004.03.036

Chow, V. T., D. R. Maidment, and L. W. Mays, 1988: *Applied hydrology*. McGraw-Hill, Inc.,

Clark, M., D. Rupp, R. Woods, X. Zheng, R. Ibbitt, A. Slater, J. Schmidt, and M. Uddstrom, 2008: Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. *Advances in Water Resources*, **31**, 1309-1324 10.1016/j.advwatres.2008.06.005

Clark, M. P., D. Kavetski, and F. Fenicia, 2011: Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resources Research*, **47**,

Clark, R. A., J. J. Gourley, Z. L. Flamig, Y. Hong, and E. Clark, 2014: CONUS-Wide Evaluation of National Weather Service Flash Flood Guidance Products. *Weather and Forecasting*, **29**, 377-392

Cleveland, W. S., E. Grosse, and W. M. Shyu, 1992: Local regression models. *Statistical models in S*, 309-376

Colby, J. D., K. A. Mulcahy, and Y. Wang, 2000: Modeling flooding extent from Hurricane Floyd in the coastal plains of North Carolina. *Global Environmental Change Part B: Environmental Hazards*, **2**, 157-168

Cosby, B., G. Hornberger, R. Clapp, and T. Ginn, 1984: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resources Research*, **20**, 682-690

Costa, J. E., 1987: Hydraulics and basin morphometry of the largest flash floods in the conterminous United States. *Journal of Hydrology*, **93**, 313-338

DeChant, C. M., and H. Moradkhani, 2012: Examining the effectiveness and robustness of sequential data assimilation methods for quantification of uncertainty in hydrologic forecasting. *Water Resources Research*, **48**, 10.1029/2011wr011011

Di Baldassarre, G., and A. Montanari, 2009: Uncertainty in river discharge observations: a quantitative analysis. *Hydrol. Earth Syst. Sci.*, **13**, 913-921 10.5194/hess-13-913-2009

Dingman, S. L., 2009: *Fluvial hydraulics*. oxford university press New York,

Dresback, K. M., J. G. Fleming, B. O. Blanton, C. Kaiser, J. J. Gourley, E. M. Tromble, R. A. Luettich, R. L. Kolar, Y. Hong, and S. Van Cooten, 2013: Skill assessment of a real-time forecast system utilizing a coupled hydrologic and coastal hydrodynamic model during Hurricane Irene (2011). *Continental Shelf Research*, **71**, 78-94

Duan, Q. 2003. Global Optimization for Watershed Model Calibration. In *Calibration of Watershed Models*, edited by Q. Duan, S. Sorooshian, H. V. Gupta, A. N. Rousseau and R. Turcotte. Washington, DC: American Geophysical Union.

Duan, Q., H. V. Gupta, S. Sorooshian, A. N. Rousseau, and R. Turcotte, 2003: *Calibration of watershed models*. American Geophysical Union,

Duan, Q., V. K. Gupta, and S. Sorooshian, 1993: Shuffled complex evolution approach for effective and efficient global minimization. *Journal of optimization theory and applications*, **76**, 501-521

Evensen, G., 1992: Using the Extended Kalman Filter With a Multilayer Quasi-Geostrophic Ocean Model. *J. Geophys. Res.*, **97**, 17905-17924 10.1029/92jc01972

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143-10162 10.1029/94jc00572

Evensen, G., 2003: The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53**, 343-367 10.1007/s10236-003-0036-9

Evensen, G., 2009: The ensemble Kalman filter for combined state and parameter estimation. *Control Systems Magazine, IEEE*, **29**, 83-104

Falcone, J. A., D. M. Carlisle, D. M. Wolock, and M. R. Meador, 2010: GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States: Ecological Archives E091-045. *Ecology*, **91**, 621-621

Feldman, A., 1995: HEC-1 flood hydrograph package. *Computer Models of Watershed Hydrology*, 119-150

Feldman, A., 2000: Hydrologic Modeling System HEC-HMS Technical Reference Manual. U.S. Army Corps of Engineers,

Finnegan, N. J., G. Roe, D. R. Montgomery, and B. Hallet, 2005: Controls on the channel width of rivers: Implications for modeling fluvial incision of bedrock. *Geology*, **33**, 229-232

Fischer, G., F. Nachtergaele, S. Prieler, H. Van Velthuizen, L. Verelst, and D. Wiberg, 2008: Global agro-ecological zones assessment for agriculture (GAEZ 2008). *IIASA, Laxenburg, Austria and FAO, Rome, Italy*,

Fry, J. A., G. Xian, S. Jin, J. A. Dewitz, C. G. Homer, Y. LIMIN, C. A. Barnes, N. D. Herold, and J. D. Wickham, 2011: Completion of the 2006 national land cover database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, **77**, 858-864

Fulton, R., 2002: Activities to improve WSR-88D Radar Rainfall Estimation in the National Weather Service. *Proc.* 2nd Federal Interagency Hydrologic Modeling Conference, Las Vegas, Nevada,

Georgakakos, K., D. Seo, H. Gupta, J. Schaake, and M. Butts, 2004: Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. *Journal of Hydrology*, **298**, 222-241 10.1016/j.jhydrol.2004.03.037

Gesch, D., G. Evans, J. Mauck, J. Hutchinson, and W. J. Carswell Jr, 2009: The national map: Elevation. *US geological survey fact sheet*, **3053**,

Gneiting, T., F. Balabdaoui, and A. E. Raftery, 2007: Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **69**, 243-268

Gourley, J., and B. Vieux, 2006: A method for identifying sources of model uncertainty in rainfall-runoff simulations. *Journal of Hydrology*, **327**, 68-80 10.1016/j.jhydrol.2005.11.036

Gourley, J. J., J. M. Erlingis, Y. Hong, and E. B. Wells, 2012: Evaluation of tools used for monitoring and forecasting flash floods in the United States. *Weather and Forecasting*, **27**, 158-173

Gourley, J. J., Z. L. Flamig, Y. Hong, and K. W. Howard, 2014: Evaluation of past, present and future tools for radar-based flash-flood prediction in the USA. *Hydrological Sciences Journal*, **59**, 1377-1389

Gourley, J. J., Y. Hong, Z. L. Flamig, A. Arthur, R. Clark, M. Calianno, I. Ruin, T. Ortel, M. E. Wieczorek, and P.-E. Kirstetter, 2013: A unified flash flood database across the United States. *Bulletin of the American Meteorological Society*, **94**, 799-805

Gourley, J. J., Y. Hong, Z. L. Flamig, L. Li, and J. Wang, 2010: Intercomparison of Rainfall Estimates from Radar, Satellite, Gauge, and Combinations for a Season of Record Rainfall. *Journal of Applied Meteorology and Climatology*, **49**, 437-452 10.1175/2009jamc2302.1

Gourley, J. J., and B. E. Vieux, 2005: A method for evaluating the accuracy of quantitative precipitation estimates from a hydrologic modeling perspective. *Journal of Hydrometeorology*, **6**, 115-133

Gupta, H., C. Perrin, R. Kumar, G. Bloschl, M. Clark, A. Montanari, and V. Andréassian, 2013: Large-sample hydrology: a need to balance depth with breadth. *Hydrology and Earth System Sciences Discussions*, **10**, 9147-9189

Gupta, H. V., M. P. Clark, J. A. Vrugt, G. Abramowitz, and M. Ye, 2012: Towards a comprehensive assessment of model structural adequacy. *Water Resources Research*, **48**,

Gupta, H. V., S. Sorooshian, T. S. Hogue, and D. P. Boyle. 2003. Advances in Automatic Calibration of Watershed Models. In *Calibration of Watershed Models*, edited by Q. Duan, S. Sorooshian, H. V. Gupta, A. N. Rousseau and R. Turcotte. Washington, DC: American Geophysical Union.

Gupta, H. V., T. Wagener, and Y. Liu, 2008: Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrological Processes*, **22**, 3802-3813 10.1002/hyp.6989

Hamill, T. M., 2001: Interpretation of Rank Histograms for Verifying Ensemble Forecasts. *Monthly Weather Review*, **129**, 550-560 10.1175/1520-0493(2001)129<0550:iorhfv>2.0.co;2

Hamill, T. M. 2006. Ensemble-based atmospheric data assimilation. In *Predictability of weather and climate*, edited by T. Palmer and R. Hagedorn. Cambridge ; New York: Cambridge University Press.

Hey, T. 2012. The Fourth Paradigm: Data-Intensive Scientific Discovery. In *E-Science and Information Management*: Springer.

Hong, Y., R. F. Adler, D. B. Kirschbaum, and G. Huffman, 2010: Capacity Building for Disaster Prevention in Vulnerable Regions of the World: Development and Implementation of A Prototype Global Flood/Landslide Prediction System. *Disaster Advances* **3** pp14-19

Hong, Y., and J. J. Gourley, 2014: *Radar Hydrology: Principles, Models, and Applications*. CRC Press,

Hong, Y., K.-l. Hsu, H. Moradkhani, and S. Sorooshian, 2006: Uncertainty quantification of satellite precipitation estimation and Monte Carlo assessment of the error propagation into hydrologic response. *Water Resources Research*, **42**, 10.1029/2005wr004398

Hrachowitz, M., H. H. G. Savenije, G. Blöschl, J. J. McDonnell, M. Sivapalan, J. W. Pomeroy, B. Arheimer, T. Blume, M. P. Clark, U. Ehret, F. Fenicia, J. E. Freer, A. Gelfan, H. V. Gupta, D. A. Hughes, R. W. Hut, A. Montanari, S. Pande, D. Tetzlaff, P.

A. Troch, S. Uhlenbrook, T. Wagener, H. C. Winsemius, R. A. Woods, E. Zehe, and C. Cudennec, 2013: A decade of Predictions in Ungauged Basins (PUB)—a review. *Hydrological Sciences Journal*, **58**, 1198-1255 10.1080/02626667.2013.803183

Huber, W., and V. Singh, 1995: EPA Storm Water Management Model-SWMM. *Computer models of watershed hydrology.*, 783-808

Huffman, G. J., D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker, 2007: The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, **8**, 38-55

Jazwinski, A., 1970: *Stochastic Processes and Filtering*. Dover Publications, 1-376.

Juston, J. M., A. Kauffeldt, B. Q. Montano, J. Seibert, K. J. Beven, and I. K. Westerberg, 2013: Smiling in the rain: Seven reasons to be positive about uncertainty in hydrological modelling. *Hydrological Processes*, **27**, 1117-1122

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, **82**, 35-45

Kampf, S. K., and S. J. Burges, 2007: A framework for classifying and comparing distributed hillslope and catchment hydrologic models. *Water Resources Research*, **43**,

Kazezyilmaz-Alhan, C., and M. Medina, 2007: Kinematic and Diffusion Waves: Analytical and Numerical Solutions to Overland and Channel Flow. *Journal of Hydraulic Engineering*, **133**, 217-228 doi:10.1061/(ASCE)0733-9429(2007)133:2(217)

Kirstetter, P.-E., G. Delrieu, B. Boudevillain, and C. Obled, 2010: Toward an error model for radar quantitative precipitation estimation in the Cévennes–Vivarais region, France. *Journal of Hydrology*, **394**, 28-41 10.1016/j.jhydrol.2010.01.009

Kirstetter, P.-E., Y. Hong, J. Gourley, S. Chen, Z. Flamig, J. Zhang, M. Schwaller, W. Petersen, and E. Amitai, 2012: Toward a framework for systematic error modeling of spaceborne precipitation radar with NOAA/NSSL ground radar-based national mosaic QPE. *Journal of Hydrometeorology*, **13**, 1285-1300

Kirstetter, P. Ä., J. J. Gourley, Y. Hong, J. Zhang, S. Moazamigoodarzi, C. Langston, and A. Arthur, 2015: Probabilistic precipitation rate estimates with ground-based radar networks. *Water Resources Research*, **51**, 1422-1442

Komma, J., G. Bloschl, and C. Reszler, 2008: Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting. *Journal of Hydrology*, **357**, 228-242 10.1016/j.jhydrol.2008.05.020

Koren, V., S. Reed, M. Smith, Z. Zhang, and D. J. Seo, 2004: Hydrology laboratory research modeling system (HL-RMS) of the US national weather service. *Journal of Hydrology*, **291**, 297-318

Koren, V., J. Schaake, Q. Duan, M. Smith, and S. Cong, 1998: PET Upgrades to NWSRFS, Project Plan. In *Unpublished Report*,

Koren, V., M. Smith, Q. Duan, Q. Duan, H. Gupta, S. Sorooshian, A. Rousseau, and R. Turcotte. 2003. Use of a priori parameter estimates in the derivation of 9 spatially consistent parameter sets of rainfall-runoff models. In *Calibration of watershed models*, edited by Q. Duan, S. Sorooshian, H. Gupta, H. Rosseau and R. Turcotte: AGU.

Koren, V., M. Smith, D. Wang, and Z. Zhang, 2000: Use of soil property data in the derivation of conceptual rainfall-runoff model parameters. *Proc.* 15th Conference on Hydrology, AMS, 2, Long Beach, CA, 103–106

Lakshmivarahan, S., and J. M. Lewis, 2010: Forward Sensitivity Approach to Dynamic Data Assimilation. *Advances in Meteorology*, **2010**, 1-12 10.1155/2010/375615

Lewis, J. M., S. Lakshmivarahan, and S. Dhall, 2006: *Dynamic data assimilation: a least squares approach*. Cambridge University Press,

Li, Y., D. Ryu, A. W. Western, and Q. Wang, 2013: Assimilation of stream discharge for flood forecasting: The benefits of accounting for routing time lags. *Water Resources Research*, **49**, 1887-1900

Liu, Y., and H. V. Gupta, 2007: Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, **43**, 10.1029/2006wr005756

Liu, Y., A. H. Weerts, M. Clark, H. J. Hendricks Franssen, S. Kumar, H. Moradkhani, D. J. Seo, D. Schwanenberg, P. Smith, A. I. J. M. van Dijk, N. van Velzen, M. He, H. Lee, S. J. Noh, O. Rakovec, and P. Restrepo, 2012: Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. *Hydrology and Earth System Sciences*, **16**, 3863-3887 10.5194/hess-16-3863-2012

Maggioni, V., R. H. Reichle, and E. N. Anagnostou, 2012: The Impact of Rainfall Error Characterization on the Estimation of Soil Moisture Fields in a Land Data Assimilation System. *Journal of Hydrometeorology*, **13**, 1107-1118

Maggioni, V., M. R. Sapiano, R. F. Adler, Y. Tian, and G. J. Huffman, 2014: An Error Model for Uncertainty Quantification in High-Time-Resolution Precipitation Products. *Journal of Hydrometeorology*, **15**, 1274-1292

Stepwise Regression Fit R2011a. The MathWorks Inc., Natick, Massachusetts.

McMillan, H., J. Freer, F. Pappenberger, T. Krueger, and M. Clark, 2010: Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions. *Hydrological Processes*, **24**, 1270–1284 10.1002/hyp.7587

Miller, D., and R. A. White, 1998: A Conterminous United States Multi-Layer Soil Characteristics Data Set for Regional Climate and Hydrology Modeling. Earth Interactions,

Mockus, V., 1961: Watershed lag. U.S. Dept. of Agriculture, Soil Conservation Service,

Montanari, A., and D. Koutsoyiannis, 2012: A blueprint for process-based modeling of uncertain hydrological systems. *Water Resources Research*, **48**, W09555

Montgomery, D. R., and K. B. Gran, 2001: Downstream variations in the width of bedrock channels. *Water Resources Research*, **37**, 1841-1846

Moore, R. J., 1985: The probability-distributed principle and runoff production at point and basin scales. *Hydrological Sciences*, **30**, 273-297

Moradkhani, H., S. Sorooshian, H. V. Gupta, and P. R. Houser, 2005: Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Advances in Water Resources*, **28**, 135-147 10.1016/j.advwatres.2004.09.002

Nash, J., 1957: The form of the instantaneous unit hydrograph. *IAHS Publ*, **45**, 114-121

Nash, J., and J. V. Sutcliffe, 1970: River flow forecasting through conceptual models part I - A discussion of principles. *Journal of Hydrology*, **10**, 282-290

Neal, J., P. Atkinson, and C. Hutton, 2007: Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements. *Journal of Hydrology*, **336**, 401-415 10.1016/j.jhydrol.2007.01.012

Nie, S., J. Zhu, and Y. Luo, 2011: Simultaneous estimation of land surface scheme states and parameters using the ensemble Kalman filter: identical twin experiments. *Hydrology and Earth System Sciences*, **15**, 2437-2457 10.5194/hess-15-2437-2011

NWS, 2010: Southeast United States Floods, September 18-23, 2009.

Oki, T., C. Valeo, and K. Heal, 2006: *Hydrology 2020: an integrating science to meet world water challenges*. IAHS Press,

Pappenberger, F., and K. J. Beven, 2006: Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resources Research*, **42**, n/a-n/a 10.1029/2005wr004820

Petersen-Overleir, A., 2004: Accounting for heteroscedasticity in rating curve estimates. *Journal of Hydrology*, **292**, 173-181 10.1016/j.jhydrol.2003.12.024

Pettyjohn, W. A., and R. J. Henning, 1979: Preliminary estimate of regional effective ground-water recharge rates in Ohio.

Poff, N., 1996: A hydrogeography of unregulated streams in the United States and an examination of scale-dependence in some hydrological descriptors. *Freshwater Biology*, **36**, 71-79

Pokhrel, P., H. Gupta, and T. Wagener, 2008: A spatial regularization approach to parameter estimation for a distributed watershed model. *Water Resources Research*, **44**, W12419

Ponce, V., 1986: Diffusion Wave Modeling of Catchment Dynamics. *Journal of Hydraulic Engineering*, **112**, 716-727 doi:10.1061/(ASCE)0733-9429(1986)112:8(716)

Ponce, V. M., 1991: Kinematic wave controversy. *Journal of Hydraulic Engineering*, **117**, 511-525

Reed, S., J. Schaake, and Z. Zhang, 2007: A distributed hydrologic model and threshold frequency-based method for flash flood forecasting at ungauged locations. *Journal of Hydrology*, **337**, 402-420 10.1016/j.jhydrol.2007.02.015

Reichle, R. H., W. T. Crow, and C. L. Keppenne, 2008: An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water Resources Research*, **44**, 10.1029/2007wr006357

Reusser, D. E., T. Blume, B. Schaefli, and E. Zehe, 2009: Analysing the temporal dynamics of model performance for hydrological models. *Hydrology and Earth System Sciences*, **13**, 999-1018

Rigby, R., and D. Stasinopoulos, 2001: The GAMLSS project: a flexible approach to statistical modelling. *Proc.* New trends in statistical modelling: Proceedings of the 16th international workshop on statistical modelling, 337-345

Rigby, R. A., and D. M. Stasinopoulos, 2005: Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **54**, 507-554

Sauer, V. B., and R. W. Meyer, 1992: Determination of Error in Individual Discharge Measurments. U.S. Geological Survey,

Schumm, S. A., 1956: Evolution of drainage systems and slopes in badlands at Perth Amboy, New Jersey. *Geological Society of America Bulletin*, **67**, 597-646

Seo, D. J., A. Seed, and G. Delrieu, eds. 2010. *Radar and multisensor rainfall estimation for hydrologic applications*. Edited by F. Y. Testik and M. Gebremichael. Vol. 191, *Rainfall: State of the Science, Geophys. Monogr. Ser.* Washington, D.C.: AGU.

Sivapalan, M., K. Takeuchi, S. Franks, V. Gupta, H. Karambiri, V. Lakshmi, X. Liang, J. McDonnell, E. Mendiondo, and P. O'connell, 2003: IAHS Decade on Predictions in Ungauged Basins (PUB), 2003‚Äì2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal*, **48**, 857-880

Smith, P. J., K. J. Beven, and J. A. Tawn, 2008: Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach. *Water Resources Research*, **44**, 10.1029/2006wr005205

Soil Survey Staff, 1994: State Soil Geographic Database (STATSGO) data users guide. In *USDA Natural Resources Conservation Service Misc. Publ. 1492*,

Soil Survey Staff, 1996: Soil Survey Laboratory methods manual. In *Soil Survey Investigations Rep. 42, Version 2.0*, National Soil Survey Center, Soil Conservation Service, U.S. Department of Agriculture,

Sorooshian, S., and J. A. Dracup, 1980: Stochastic parameter estimation procedures for hydrologie rainfall-runoff models: Correlated and heteroscedastic error cases. *Water Resour. Res.*, **16**, 430-442 10.1029/WR016i002p00430

Sorooshian, S., Q. Duan, and V. K. Gupta, 1993: Calibration of rainfall-runoff models: Application of global optimization to the Sacramento Soil Moisture Accounting Model. *Water Resour. Res.*, **29**, 1185-1194 10.1029/92wr02617

Stasinopoulos, D. M., and R. A. Rigby, 2007: Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, **23**, 1-46

Van Cooten, S., K. E. Kelleher, K. Howard, J. Zhang, J. J. Gourley, J. S. Kain, K. Nemunaitis-Monroe, Z. Flamig, H. Moser, A. Arthur, C. Langston, R. Kolar, Y. Hong, K. Dresback, E. Tromble, H. Vergara, R. A. Luettich, B. Blanton, H. Lander, K. Galluppi, J. P. Losego, C. A. Blain, J. Thigpen, K. Mosher, D. Figurskey, M. Moneypenny, J. Blaes, J. Orrock, R. Bandy, C. Goodall, J. G. W. Kelley, J. Greenlaw, M. Wengren, D. Eslinger, J. Payne, G. Olmi, J. Feldt, J. Schmidt, T. Hamill, R. Bacon, R. Stickney, and L. Spence, 2011: The CI-FLOW Project: A System for Total Water Level Prediction from the Summit to the Sea. *Bulletin of the American Meteorological Society*, **92**, 1427-1442 10.1175/2011bams3150.1

Vergara, H. 2011. Assimilating Streamflow with the Ensemble Kalman Filter for Improving Flood Prediction. M.Sc. Thesis, Civil Engineering and Environmental Sciences, The University of Oklahoma, Norman.

Vergara, H., Y. Hong, J. J. Gourley, E. N. Anagnostou, V. Maggioni, D. Stampoulis, and P.-E. Kirstetter, 2014: Effects of Resolution of Satellite-Based Rainfall Estimates on Hydrologic Modeling Skill at Different Scales. *Journal of Hydrometeorology*, **15**, 593-613

Vrugt, J. A., and W. Bouten, 2003: Toward improved identifiability of hydrologic model parameters: The information content of experimental data. *Water Resources Research*, **38**, 10.1029/2001wr001118

Vrugt, J. A., C. J. F. Braak, H. V. Gupta, and B. A. Robinson, 2008: Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? *Stochastic Environmental Research and Risk Assessment*, **23**, 1011-1026 10.1007/s00477-008-0274-y

Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten, 2005: Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. *Water Resources Research*, **41**, 10.1029/2004wr003059

Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian. 2003. A Shuffled Complex Evolution Metropolis Algorithm for Estimating Posterior Distribution of Watershed Model Parameters. In *Calibration of Watershed Models*, edited by Q. Duan, S. Sorooshian, H. V. Gupta, A. N. Rousseau and R. Turcotte. Washington, DC: American Geophysical Union.

Vrugt, J. A., H. V. Gupta, S. C. Dekker, S. Sorooshian, T. Wagener, and W. Bouten, 2006: Application of stochastic parameter optimization to the Sacramento Soil Moisture Accounting model. *Journal of Hydrology*, **325**, 288-307 10.1016/j.jhydrol.2005.10.041

Vrugt, J. A., and B. A. Robinson, 2007: Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging. *Water Resources Research*, **43**, 10.1029/2005wr004838

Vrugt, J. A., C. J. F. ter Braak, C. G. H. Diks, B. A. Robinson, J. M. Hyman, and D. Higdon, 2009: Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling. *International Journal of Nonlinear Sciences and Numerical Simulation*, **10**, 273-290 10.1515/ijnsns.2009.10.3.273

Wagener, T., D. P. Boyle, M. J. Lees, H. S. Wheater, H. V. Gupta, and S. Sorooshian, 2001: A framework for development and application of hydrological models. Copernicus GmbH,

Wagener, T., M. Sivapalan, P. Troch, and R. Woods, 2007: Catchment classification and hydrologic similarity. *Geography Compass*, **1**, 901-931

Wang, J., Y. Hong, L. Li, J. J. Gourley, S. I. Khan, K. K. Yilmaz, R. F. Adler, F. S. Policelli, S. Habib, and D. Irwn, 2011: The coupled routing and excess storage (CREST) distributed hydrological model. *Hydrological Sciences Journal*, **56**, 84-98

Wang, X., and C. H. Bishop, 2005: Improvement of ensemble reliability with a new dressing kernel. *Quarterly Journal of the Royal Meteorological Society*, **131**, 965-986

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, **130**, 1913-1925

Wigmosta, M. S., B. Nijssen, P. Storck, and D. Lettenmaier, 2002: The distributed hydrology soil vegetation model. *Mathematical models of small watershed hydrology and applications*, 7-42

Wigmosta, M. S., L. W. Vail, and D. P. Lettenmaier, 1994: A distributed hydrology-vegetation model for complex terrain. *Water Resour. Res.*, **30**, 1665-1679 10.1029/94wr00436

Williams, G. P., 1978: Bank-full discharge of rivers. *Water Resources Research*, **14**, 1141-1154

Woolhiser, D. A., R. Smith, and D. C. Goodrich, 1990: *KINEROS: a kinematic runoff and erosion model: documentation and user manual*. US Department of Agriculture, Agricultural Research Service,

Wu, L., D.-J. Seo, J. Demargne, J. D. Brown, S. Cong, and J. Schaake, 2011: Generation of ensemble precipitation forecast from single-valued quantitative precipitation forecast for hydrologic ensemble prediction. *Journal of Hydrology*, **399**, 281-298 10.1016/j.jhydrol.2011.01.013

Xie, X., and D. Zhang, 2010: Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. *Advances in Water Resources*, **33**, 678-690 10.1016/j.advwatres.2010.03.012

Yao, C., Z. Li, Z. Yu, and K. Zhang, 2012: A priori parameter estimates for a distributed, grid-based Xinanjiang model using geographically based information. *Journal of Hydrology*, **468**, 47-62

Yilmaz, K. K., H. V. Gupta, and T. Wagener, 2008: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. *Water Resour. Res.*, **44**, W09417 10.1029/2007wr006716

Zhang, J., and Coauthors, 2015: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc*, **In review**,

Zhang, J., K. Howard, C. Langston, S. Vasiloff, B. Kaney, A. Arthur, S. Van Cooten, K. Kelleher, D. Kitzmiller, F. Ding, D.-J. Seo, E. Wells, and C. Dempsey, 2011: National Mosaic and Multi-Sensor QPE (NMQ) System: Description, Results, and Future Plans. *Bulletin of the American Meteorological Society*, **92**, 1321-1338 10.1175/2011bams-d-11-00047.1

Zhao, R., X. Liu, and V. Singh, 1995: The Xinanjiang model. *Computer models of watershed hydrology.*, 215-232

Zhao, R., Y. Zhang, L. Fang, X. Liu, and Q. Zhang, 1980: The Xinganjiang Model. *Proc.* Hydrological Forecasting Proceedings Oxford Symposium, IASH 129, 351-356