

1-28-2015

# Toward improved evaluation of large scale hydrologic models: estimation and quantification of parameter uncertainty

Lijuan Jia

Follow this and additional works at: [https://digitalrepository.unm.edu/ce\\_etds](https://digitalrepository.unm.edu/ce_etds)

---

## Recommended Citation

Jia, Lijuan. "Toward improved evaluation of large scale hydrologic models: estimation and quantification of parameter uncertainty." (2015). [https://digitalrepository.unm.edu/ce\\_etds/18](https://digitalrepository.unm.edu/ce_etds/18)

This Dissertation is brought to you for free and open access by the Engineering ETDs at UNM Digital Repository. It has been accepted for inclusion in Civil Engineering ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact [disc@unm.edu](mailto:disc@unm.edu).

Lijuan Jia

Candidate

Civil Engineering

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

*Approved by the Dissertation Committee:*

Mark Stone

Chairperson

Julie Coonrod

Karl Benedict

Ramon Naranjo

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

**TOWARD IMPROVED EVALUATION OF LARGE SCALE  
HYDROLOGIC MODELS: ESTIMATION AND  
QUANTIFICATION OF PARAMETER UNCERTAINTY**

by

**LIJUAN JIA**

B.S., Civil and Environmental Engineering,  
Hebei University of Engineering, 2005  
M.S., Hydrology and Water Resources Engineering,  
Hohai University, 2008

DISSERTATION

Submitted in Partial Fulfillment of the  
Requirements for the Degree of

**Doctor of Philosophy  
Engineering**

The University of New Mexico  
Albuquerque, New Mexico

**December, 2014**

# **TOWARD IMPROVED EVALUATION OF LARGE SCALE HYDROLOGIC MODELS: ESTIMATION AND QUANTIFICATION OF PARAMETER UNCERTAINTY**

**by**

**Lijuan Jia**

B.S., Civil and Environmental Engineering,

Hebei University of Engineering, China, 2005

M.S., Hydrology and Water Resources Engineering,

Hohai University, China, 2008

Ph.D., Engineering, University of New Mexico, 2014

## **Abstract**

With the development of increasingly complex hydrologic models that use a wide range of parameters to represent hydrologic processes both in space and time, many challenges arise with respect to simulation and quantification of uncertainty. The goal of this research is to introduce strategies to effectively and efficiently estimate and quantify hydrologic responses. A robust framework for parameter estimation and uncertainty quantification is proposed. The procedure also considers temporal variations over a time-series. Specifically, two issues of traditional estimation schemes and uncertainty quantification methods were addressed: overparameterization and reduction of parameter uncertainty through quantitative information.

Parameters were categorized as distributed, inactive, or lumped by combining traditional concepts from identifiability and overparameterization with approaches from sensitivity analyses. This led to decreased dimensionality and thus less required computational demand. The framework takes into account climatic conditions over large scales. As a result, the modeler can investigate parameter uncertainty subbasin-by-subbasin as well as temporal variations. The result is a novel estimation scheme capable of subjectively investigating likelihood to extract quantitative information, improving communication of hydrologic simulation data, and ultimately improving reliability of hydrologic models.

The techniques proposed and demonstrated here were programmed within the MATLAB programming environment using the Linux platform. The hydrologic model used in this study was the Variable Infiltration Capacity (VIC) model. The finalized scripting environment will be made available to the modeling community.

# Table of Contents

<b>Chapter 1</b>	<b>Introduction.....</b>	<b>1</b>
1.1	Background.....	1
1.1.1	Heterogeneity and Homogeneity → Spatial and Temporal Variability .....	2
1.1.2	Sensitivity and Identifiability → Optimal Parameter Values vs. Optimal Parameter Sets; Overparameterization .....	3
1.1.3	Equifinality vs. Uniqueness → Optimal Sets vs. Optimality .....	4
1.1.4	Equal to vs. Equivalent → Uncertainty Quantification.....	5
1.1.5	Summary and Extended Definitions .....	6
1.2	Objectives .....	7
1.3	VIC Model Characteristics .....	9
1.4	Key Contributions in This Research.....	10
1.5	Outline of the Dissertation.....	11
<b>Chapter 2</b>	<b>Parameter Sensitivity and Dynamic Identifiability Analysis within a Spatially Distributed Framework for Large Scale Hydrologic Modeling: Case Study for the Gila River Basin .....</b>	<b>16</b>
2.1	Introduction .....	16
2.2	Background.....	17
2.3	Methods .....	20
2.3.1	Framework Overview.....	20
2.3.2	Hydrological Model .....	21

2.3.3 Objective Function .....	22
2.3.4 Sensitivity Analysis Approach .....	23
2.4 Case Study .....	24
2.4.1 Research Domain .....	24
2.4.2 Results.....	25
2.5 Discussion and Conclusion.....	32
<b>Chapter 3 Integrated Multi-Criteria Estimation under Parameter Uncertainty Quantification for Large Scale Distributed Hydrology Model.....</b>	<b>55</b>
3.1 Introduction .....	55
3.2 Background.....	57
3.3 Methodology for Quantification Parameter Uncertainty with Multi-Criteria Estimation.....	59
3.3.1 GLUE Method Characteristics .....	59
3.3.2 Pareto Ranking for Uncertainty Cluster Outputs .....	63
3.3.3 Distinction Between GLUE and Pareto Based Methods.....	64
3.4 Case Study .....	64
3.4.1 Single-Criteria Calibration .....	65
3.4.2 Multi-Criteria GLUE Analysis.....	67
3.4.3 Multi-Criteria Pareto Analysis .....	68
3.5 Conclusions and Discussions.....	68

<b>Chapter 4</b>	<b>Conclusions and Future Works</b> .....	<b>79</b>
4.1	Conclusions .....	79
4.1.1	Building a Framework to Effectively and Efficiently Demonstrate Estimation Scheme.....	79
4.1.2	Quantifying Hydrological Responses through Integrated Multi-Criteria Simulation .....	81
4.1.3	Integration of a Post-Processing Tool for Estimation/Uncertainty Analysis .....	82
4.2	Recommendations for Future Research.....	82
<b>Chapter 5</b>	<b>References</b> .....	<b>86</b>
<b>APPENDICES</b> .....		<b>107</b>
<b>Appendix A</b>	<b>An Approach to Measure Parameter Identifiability</b> .....	<b>108</b>
<b>Appendix B</b>	<b>Application of the Approach to Measure Identifiability - DYNIA</b> ...	<b>112</b>



## List of Tables

Table 1-1	VIC model primary components .....	13
Table 2-1	Selected basins for study domain .....	37
Table 2-2	Six parameters of VIC model for calibration and uncertainty analysis .....	37
Table 3-1	List probabilistic analysis with observed discharge, GLUE uncertainty outputs and Pareto uncertainty outputs .....	71

## List of Figures

- Figure 1–1 Parameter properties with identified and non identified output. (x-axis represents the feasible parameter range; y-axis represents selected objective function or simply means the errors between modeled and observed hydrologic behavior) ..... 15
- Figure 2–1 A flow chart for the new framework of parameter estimation for distributed large scale hydrologic model..... 38
- Figure 2–2 Research domain on the Gila River basin (The blank contour represents the Gila River domain, the color section with purple, green and yellow represents the three subbasins, Salt River basin, Upper Gila River basin and San Pedro River basin; the red triangular remarks the streamgauge for calibration discharge with USGS measurement for each subbasin.) ..... 39
- Figure 2–3 Model performance for Salt River basin. (a) Baseline simulations based on comparison of time-series monthly streamflow (1981.1-2000.12) by observed (blue line) vs. model calculated (purple dash line) as traditional optimal calibration scheme (with good matching especially on extreme flow). (b) Errors plot on each time step. (c) Modeled vs observed streamflow. It visually shows a bias estimation on this region. .... 41

- Figure 2–4 Model performance for San Pedro River basin. (a) Baseline simulations based on comparison of time-series monthly streamflow (1981.1-2000.12) by observed (blue line) vs. model calculated (green dash line) as traditional optimal calibration scheme (worse mimic on extreme flow). (b) Create errors plot on each time step. (c) Modeled vs observed streamflow. It visually shows a bias estimation on this region. .... 43
- Figure 2–5 Model performance for Upper Gila River basin. (a) Baseline simulations based on comparison of time-series monthly streamflow (1981.1-2000.12) by observed (blue line) vs. model calculated (yellow dash line) as traditional optimal calibration scheme. (b) Create errors plot on each time step. (c) Modeled vs observed streamflow. It visually shows a bias estimation on this region. .... 45
- Figure 2–6 Model performance with a class of outputs based on time-series streamflow simulations for each subbasin (more detail description). The color bar shows likelihoods to each simulations based on criteria function which given higher values with good simulation performance and lower values with bad performance. (a) Salt River basin; (b) San Pedro River basin; (c) Upper Gila River basin..... 47
- Figure 2–7 Parameter Sensitivity analysis for each basin. All of these global-population random simulations are divided into equal ten groups based on the ranking of objective functions. The nomalised cumulative distribution is calculated for each group with attained a higher value as better performing model

simulations. (a) Salt River basin; (b) San Pedro River basin; (c) Upper Gila River basin..... 49

Figure 2–8 Scatter plots for parameter sets based on model performance. (It is the outputs from Monte Carlo Uniform Random Simulations, the behavior parameter sets are equally divided into 10 groups ranked with model performance given higher value with better performing. The pink dots represent the best top 10% parameters comparing to the worst 10% parts with the light blue section). (a) Salt River basin; (b) San Pedro River basin; (c) Upper Gila River basin ..... 51

Figure 2–9 Parameter dynamic analysis for each basin. It is a dynamic measurement of identifiability (Appendix A) to each parameter based on time series hydrologic behavior. Take parameter  $b$  and  $d_2$  as representators. (a) Salt River basin; (b) San Pedro River basin; (c) Upper Gila River basin ..... 54

Figure 3–1 Traditional single-criteria streamflow simulations with respect to different objective functions. (a) The black points represent observed monthly discharge under study period. The blue and pink lines represent the best simulated streamflow compared to observed time series based on NSE and RE, respectively. (b) Modeled vs. observed streamflow for both objective functions. (c) Residual errors plot. .... 73

Figure 3–2 Normalized parameter sets space selected by optimal value due to NSE performance. Among them, the top 5% optimal clusters are drawn (each red

line represents a parameter set; the black line represents the optimal parameter set based on NSE)..... 74

Figure 3–3 Normalized parameter sets space selected by optimal value due to RE performance. Among them, the top 5% optimal clusters are drawn (each blue line represents a parameter set; the black line represents the optimal parameter set based on RE). Note: Each line represents one parameter set which could be given a measurement by model performance it will be meaningless with single point. .... 74

Figure 3–4 A hypothetical example trade-off between NSE and RE is shown. This trade-off is also referred to as the Pareto optimality. The performance drop into the lower left corner could satisfy both of the objective functions..... 75

Figure 3–5 GLUE likelihood measures with multi-criteria NSE&RE tradeoffs for Salt River basin. The application of GLUE in this research is subjectively selected the likelihood function as terminal criteria (defined in equation 3-4 and 3-5). Scatter plots also represent model performance with higher likelihood values by better simulations. All of those likelihood values spread through the parameter space should be added up to one..... 75

Figure 3–6 GLUE output associated confidence limits. Hydrograph of 95% percentile confidence prediction bounds estimated by GLUE with multi-criteria tradeoffs. The blue point indicates observed streamflow and the grey shaded area represents the prediction uncertainty resulted from GLUE estimation.

- The black line demonstrates the uncertainty bounds with specific (95%) confidence level..... 76
- Figure 3–7 Pareto front selected based on trade-offs between NSE and RE. The blue points represent model performances from all random simulations with respect to the adopted two objective functions. The red points are the optimal sets filtered by pareto ranking algorithm. The cluster simulations related to these red points are remarked as pareto uncertainty bounds. .... 77
- Figure 3–8 Pareto output uncertainty. The pink points are observed monthly discharge time series. The grey areas are extracted from pareto optimal space which represents the red points simulation cluster as shown in the last figure. .... 78
- Figure 4–1 Characteristics and flow chart for the framework of estimation-quantification uncertainty to large scale distributed model building in this research..... 85
- Figure A–1 Shows the definition and calculation to measure identifiability. Take parameter b on San Pedro River basin as an example..... 110
- Figure A–2 Shows the identifiability to same parameter b with different climatic gradient region. The parameter b has much higher identifiability in San Pedro River basin than Salt River basin..... 110
- Figure A–3 Measurement of identifiability with parameter b and  $d_1$  on San Pedro River basin. It demonstrates the parameter b has much higher identifiability than

parameter  $d_1$ . The cumulative distribution based on NSE with parameter  $d_1$  is almost straight line which means non-identified in this region..... 111

Figure B-1 Procedure to DYNIA (Wagener et al., 2003a)..... 113

## Chapter 1 Introduction

*Our models should be designed expressly to maximize the possibility of discovering that of which we are ignorant. -----Beck (2002)*

### 1.1 Background

Hydrologic models have become a powerful and reliable tool for simulating natural physical systems and estimating hydrologic impacts of changing watershed and climatic conditions. In order to explicitly recognize the inherent heterogeneity present in watersheds, complex large scale modeling schemes have been developed in recent years. These models numerically represent a range of environmental processes over temporal and spatial scales using a multitude of parameters (Ajami et al., 2004; Wagener et al. 2005; Wagener et al., 2010).

The increasing complexity of modern hydrologic models comes at the cost of increased parameter uncertainty and the inadequacy of classical approaches without considering uncertainty for evaluating model performance (Beven, 2000; Beven and Freer, 2001). It is necessary to perform a robust calibration and uncertainty assessment to sufficiently evaluate and improve upon hydrologic simulations and to conduct a



qualitative investigation of model behaviors. Accordingly, two questions arise with respect to improving model accuracy and computational efficiency in the processes of simulations and optimization procedures. First, how can we improve estimation efficiently and effectively for highly complex models? Second, what are the appropriate strategies to conduct a qualitative investigation for grouping or clustering simulations for distributed watershed models (Wagener et al., 2006)?

Before discussing the questions mentioned above, some terms need to be defined. These definitions will be used frequently in the following chapters to address the challenges surrounding the application of distributed hydrologic models.

### **1.1.1 Heterogeneity and Homogeneity → Spatial and Temporal Variability**

The true landscape defined by hydrologic models inherently displays a wide array of uncertainty as a result of spatial and temporal variability within all aspects of hydrologic processes (McDonnell et al., 2007; Lovett et al., 2006; Grayson et al., 2001). An initial step in the model development process is to determine whether a lumped or spatially distributed model is required. For example, due to the defined objectives by quantifying the infiltration, would a lump model be sufficient? If a relatively high degree of homogeneity is observed across the modeling domain, a lumped approach may be suitable. In this case, the parameters are considered to be uniform across the entire domain. In many cases, even when spatial variability (heterogeneity) exists across the landscape, the underlying processes can be sufficiently parameterized (e.g. area weighted) to produce sufficiently reasonable representations of the hydrologic processes.

As hydrologic sciences, numerical methods, and the availability of spatially explicit data continue to develop, it has become common to incorporate heterogeneity directly by representing basin characteristics in a distributed fashion and parameterizing the model on a cell-by-cell or subbasin-by-subbasin basis (McDonnell et al., 2007). This is accomplished through the use of spatially distributed inputs, such as digital elevation models (DEM), Geographic Information System (GIS), soil and vegetation mapping and remote sensing data (Lovett et al., 2006; Beven and Moore, 1992). Distributed models can be employed by coupling probabilistic methods and “grid” schemes with respect to hydrologic process interactions and by considering physical and dynamic variations in precipitation, land use / land cover, climatic conditions, and topography. Thus, distributed models can serve as an integrator of inputs to robustly deal with spatial variability.

### **1.1.2 Sensitivity and Identifiability → Optimal Parameter Values vs. Optimal Parameter Sets; Overparameterization**

A powerful approach for evaluating model performance is through sensitivity analyses (Bell et al., 2000). The analyses can easily be implemented to estimate the response to a range of parameters in order to evaluate the stability and specificity of the parameterization. High sensitivities demonstrate significant changes to model performance with small perturbation in the input parameter values. In contrast, insensitive parameters will produce minor differences to the behavior of hydrologic simulations as the input is varied. If model results are insensitive, it can be interpreted that the parameterization process does not capture the underlying hydrologic response modes in the sense of physical interpretations. This is referred to as an

“overparameterization” condition (Schwarz, 1978; O’Connell, 1990, 1998; Pokhrel, et al., 2008; Schoups et al., 2008; Whittaker et al., 2010).

In spite of the usefulness of this process, sensitivity analyses alone do not consider non-uniqueness, which could be recognized by identifiability analysis especially in highly complex distributed models. Parameters are considered to be identifiable when they demonstrate a distinct minimum region with a specific objective function (e.g. relative error) (Figure 1-1 (b)). In contrast, non-identifiable parameters (Figure 1-1 (a)) will lead to equally acceptable model performances through the whole feasible parameter range. An identifiability analysis helps the modeler better understand the characteristics of parameter without adequate prior information and also helps her or him recognize a reasonable parameter range and detect a proper model structure (Sorooshian and Gupta, 1985; Wagener, 1998) (tested in Chapter 2).

With the development of increasingly complex models and likewise additional non-identifiable parameters, traditional methods based on identification of unique optimal parameters are giving way to new methods aimed at identifying equal performing optimal parameter sets. Non-identified phenomena will progress to another hydrologic condition involving uncertainty analysis, termed as “equifinality” (as described below).

### **1.1.3 Equifinality vs. Uniqueness → Optimal Sets vs. Optimality**

Within traditional hydrologic model calibration approaches, a manual trial and error process is typically employed along with a sensitivity analysis. The approach assumes there is one optimal parameter set that uniquely matches the observed behavior. However, with the increasing degree of model complexity, modern approaches must also

consider parameter interactions. Each parameter can take a value across the entire feasible parameter space, which can lead to a large number of acceptable model parameterizations. The issue of non-uniqueness of optimum parameters or parameter sets is referred to as equifinality (Beven, et al., 1992; Beven & Binley, 1996; Beven, 2000; Yapo et al., 1998; Duan, et al., 2003; Abbaspour et al., 2004; Beven, 2006; Beven et al., 2011).

Accordingly, traditional calibration and simulation schemes based on a unique optimal solution are not adequate for dealing with highly complex models due to the equifinality condition. The uncertainty analysis should be quantified within a robust framework to significant the extent of influence to hydrologic impacts.

#### **1.1.4 Equal to vs. Equivalent → Uncertainty Quantification**

The existing uncertainty of parameters raises a question with respect to traditional model evaluations: How do we evaluate the simulation results and quantify/communicate uncertainty to support decision making?

As mentioned above, the optimal parameter sets arising due to equifinality are equal with respect to their ability to create acceptable model performance. However, they are not equivalent parameterizations. As a result, we cannot simply perform a mathematical average to all acceptable outputs to produce a final single optimal simulation. Rather, we can develop and apply likelihood weights to discriminate the degree of accuracy to each acceptable parameter set (Kavetski et al., 2006). To achieve the uncertainty quantification, two questions need to be answered: (1) How to discriminate equally performance parameters (behavioral sets) from randomly parameter

space? (2) How to reveal the extent of parameterization and uncertainty by aggregating those optimums? These questions are explored through this research.

### **1.1.5 Summary and Extended Definitions**

In summary, as the application of complex hydrologic models grows in popularity, it is important to develop methods for accurately representing physical processes without sacrificing computational efficiency. For example, sensitivity and identifiability analyses can enhance the parameterization process by minimizing overparameterization of components (test in Chapter 2). The quantification of uncertainty can effectively define parametric modes and integrate available information reliably to inform decision-makers (test in Chapter 3).

This research focused on evaluation of three issues related to the terms explained above. First, within a distributed model, if a parameter is found to be insensitive through a sensitivity analysis, it could indicate that the model is overparameterized. Further, if a parameter has a low degree of identifiability but is found to be sensitive, it also could be recognized as overparameterized. To account for this condition, I defined these parameters as lumped parameters, which can be assigned a single calibrated value rather than considering distributed values across a grid. This strategy leads to reduced complexity by decreasing parameter dimensionality. Second, with respect to identifiability processes, the traditional approach is to explore the acceptable or optimal parameter regions as a standard to discriminate between identifiable and non-identifiable parameters, while ignoring the whole parameter space. In this research, I introduce an approach for identifying the parameter space using cluster

plots, and investigate the interaction between the identifiability of parameters and their underlying connection to hydrologic processes. Third, with respect to temporal uncertainty, I investigated approaches for varying parameters by testing their performance and identifiability in a transient or dynamic nature over a time series streamflow hydrograph.

Obviously, major challenges remain and room exists for improving the methods of parameterization, estimation, and improvement of hydrologic model performance. This research provides progress in this direction.

## 1.2 Objectives

Given the high degree of the complexity associated with distributed models that attempt to mimic the behavior of natural systems (especially to simulate streamflow in this research), there are a large number of parameters with high levels of uncertainty. *The overarching goal of this dissertation was to advance methods for understanding and describing uncertainty associated with parameters from two aspects: estimation schemes and uncertainty quantification. These concepts were analyzed from the perspectives of spatial and temporal heterogeneity.*

With broadly versatile applications, highly complex hydrologic models reflect various degrees of parameterization. To achieve an efficient and accurate simulation with respect to specified hydrological aspects, the selection of the appropriate degree of complexity by proper parameterization becomes critical. Traditional strategies identify inactive parameters through sensitivity analyses. In this study, an alternative is proposed

where identifiability and sensitivity analyses are integrated to reduce free parameters and represent them as lumped wherever possible. It is also argued that a framework for dealing with variability across climate regions can also enhance the estimation of parameter sets.

The strategies proposed here aim to address the overparameterization aspects of model calibration and to include a moderate level of complexity to the modeled system. This allows the modeler to shrink the parameter dimensionality and thus improve computational efficiency without diminishing accuracy of the hydrologic simulations.

Population-based search algorithms were employed widely in this research they proved to be powerful search optimization methods for large calibrated distributed model parameters. This was found to be especially true when investigating the interactions or correlations between parameters. In this research, the uncertainty analysis was quantified within a novel framework by addressing two key issues: how to discriminate behavioral parameter sets as optimal ones from a random parameter population? And how to reveal the extent of estimation by aggregating those optimums? The overarching aim is to constrain the acceptable optimal parameter sets by decreasing the extent of uncertainty and to extract qualitative information that is more reliable and understandable to decision-makers.

Finally, the temporal influence of parameter temporal uncertainty was investigated. This allowed for an estimation of the influence of parameterization on various hydrologic impacts over a time series.

### 1.3 VIC Model Characteristics

The Variable Infiltration Capacity (VIC) model is a large-scale, semi-distributed hydrologic model (Liang et al., 1994, 2003). VIC has been developed and is supported by researchers at the University of Washington. It balances both the water and surface energy within grid cells by varied spatial (2-, 1-, 1/2-, 1/4-, 1/8-, and 1/16-degree) and temporal scales (3-hourly, daily, and monthly) with respect to water and energy budget terms (meteorological data, maximum & minimum temperature, wind speed, elevation, soil properties, vegetation characteristics, evapotranspiration, runoff, snow water equivalent, soil moisture, net shortwave and longwave radiation, latent and sensible heat fluxes, ground flux, surface temperature). VIC is parameterized based on water balance and energy balance principle and it has been widely implemented from water resources management to land atmosphere interactions and climate change by incorporation with General Circulation Models (GCMs) (VIC manual, 2010). The significant characteristics of the VIC model are its ability to consider spatial information represented by vegetation maps and multiple soil layers (3-soil layers) with variable infiltration and non-linear baseflow (Liang et al., 1994; Liang et al., 1996; Nijssen et al., 1997). From the pool of parameters, we selected five parameters controlling the surface and subsurface simulation based on streamflow behavior. Table 1-1 lists the main model components with basic mathematical formulations.

The main features of VIC that are applicable to this research are as follows. The land surface is modeled as a grid of uniformly sized cells accounting for sub-grid heterogeneity (e.g. elevation, land cover) handled via statistical distributions. Inputs are



time series of daily or sub-daily meteorological drivers (e.g. precipitation, air temperature, wind speed). The model can use sub-daily meteorological data at intervals matching the simulation time step. VIC can consider spatial variability in precipitation, arising from either storm fronts/local convection or topographic heterogeneity and also it can subdivide the grid cell into a time-varying wet fraction (where precipitation falls) and dry fraction (where no precipitation falls). Land-atmosphere fluxes, and the water and energy balances at the land surface, are simulated at a daily or sub-daily time step. Land cover can be subdivided within each grid cell through any number of “titles” (Liang et al., 1994, 1996).

#### **1.4 Key Contributions in This Research**

The framework performed here addressed an effective and efficiency estimation-quantification uncertainty analysis scheme to highly complex distributed model system. This approach includes improved incorporation of parameter sensitivity and uncertainty quantification. This work provides the following contributions to hydrologic sciences and water resources engineering:

(1) Properties of parameter identification and overparameterization: we extend the applications of traditional definitions. Here, the parameters in distributed models are categorized as lumped and distributed parameters to decrease the model dimensionality using combined analysis of identifiability and sensitivity.

(2) Considering inherent characteristics with respect to parameter interaction: The demonstration of population-based search approaches (Monte Carlo Uniform

Random Search) were verified to be powerful for large distributed models. This was particularly evident when studying the interactions or correlations between parameters, which are usually ignored by traditional estimation schemes.

(3) Achieving a robust framework by consideration of trade-off quantification and subjective selection of likelihood: the qualitative information that directly considers uncertainty improves methods for communicating model results and uncertainty to decision-makers in applied hydro-sciences.

(4) Temporal variability of parameters: The analysis was extended to study varying hydrologic behavior through time series to investigate the influence of parameter temporal variability. The resulting methods can be implemented to improve selection of proper model structure, hydrograph segmentation, and tradeoff objective functions.

(5) Generation of a MATLAB script package: The methods and scripts generated for this work within the MATLAB programming platform have been combined into a runtime package. This work will be distributed and can be applied to a wide range of analysis related to climate change, wildfire affects, and other scenarios which require uncertainty outputs.

## **1.5 Outline of the Dissertation**

The research presented in this dissertation demonstrates a novel framework of model estimation and uncertainty quantification in order to address the accuracy and efficiency of complex, distributed hydrologic models. The research provides progress *towards improved evaluation of large-scale hydrologic systems*. This first chapter

described general background and motivation for the work. The second chapter introduces and illustrates the framework and provides a case study titled: *Parameter sensitivity and dynamic identifiability within a spatially distributed framework for large-scale hydrologic modeling: case study for the Gila River basin.*

The third chapter addresses the issues of qualitative information based on the outputs described in Chapter 2. Chapter 3 is titled *Integrated multi-criteria estimation under parameter uncertainty quantification for large scale distributed hydrology model.* An overall summary of this research and recommendations for future work are provided in Chapter 4.

Table 1-1 VIC model primary components

Processes	Governing model	Governing equations	Calibrated parameters	Parameter meanings	Notes
Water balance	Bucket model	$\frac{\partial S}{\partial t} = P - E - R$			It is continuous equation for each time-step
Evapotranspiration	Canopy evaporation-Penman-Monteith equation (Shuttleworth, 1993)	$\lambda_v E_p = \frac{\Delta(R_n - G) + \rho_a C_p (e_s - e_a)/r_a}{\Delta + \gamma}$			
	Vegetation transpiration (Blondin, 1991; Ducoudre et al., 1993)	$E_i = (1 - (\frac{W_i}{W_{im}})^{2/3}) E_p \frac{r_w}{r_w + r_o + r_c}$ Where, $r_c = \frac{r_{oc} g_T g_{vpd} g_{PAR} g_{sm}}{LAI}$			
Infiltration	Infiltration capacity (Zhao et al., 1980)	$i = i_m [1 - (1 - A)^{1/b}]$	b	Infiltration shape parameter	Infiltration capacity as a function of relative saturated gridcell area. Higher value gives lower infiltration.
Overland flow	St. Venant equation	$\frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} + \frac{\partial(vh)}{\partial y} = q$			

**Table 1-1 VIC model components (Cont'd)**

Unsatrate flow	One-dimensional Richard equation	$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left( D(\theta) \frac{\partial \theta}{\partial z} \right) + \frac{\partial K(\theta)}{\partial z}$			
Base flow	Arno model formulation (Franchini and Pacciani, 1991)	$Q_b = \begin{cases} \frac{D_s D_m}{W_s \theta_s} \cdot \theta_3, 0 \leq \theta_3 \leq W_s \theta_s \\ \frac{D_s D_m}{W_s \theta_s} \cdot \theta_3 + \\ (D_m - \frac{D_s D_m}{W_s}) \left( \frac{\theta_3 - W_s \theta_s}{\theta_s - W_s \theta_s} \right)^2 \\ \text{when } \theta_3 \geq W_s \theta_s \end{cases}$	Ws	The fraction of maximum soil moisture where non-linear baseflow occurs.	A higher value of Ws will raise the water content required for rapidly increasing, non-linear baseflow.
			Ds	The fraction of $D_{s_{max}}$ where non-linear baseflow begins.	A higher value of Ds, the baseflow will be higher at lower water content in the lowest soil layer.
			$D_{s_{max}}$	The maximum baseflow that occur from the lowest soil layer.	It depends on hydraulic conductivity.
Channel routing	St. Venant equation	$\frac{\partial h}{\partial t} + \frac{\partial (uh)}{\partial x} = q$			

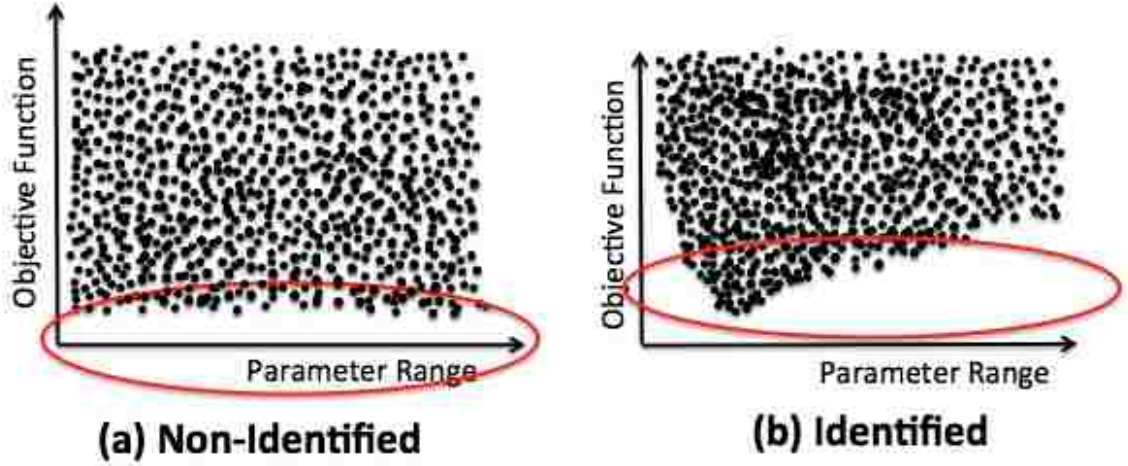


Figure 1–1 Parameter properties with identified and non identified output. (x-axis represents the feasible parameter range; y-axis represents selected objective function or simply means the errors between modeled and observed hydrologic behavior)

## **Chapter 2 Parameter Sensitivity and Dynamic Identifiability**

### **Analysis within a Spatially Distributed Framework for**

### **Large Scale Hydrologic Modeling:**

### **Case Study for the Gila River Basin**

#### **2.1 Introduction**

Hydrologic models provide an important translator function between meteorological, landscape, and subsurface processes (Gelhar, 1986; Beven, 1989; Wood, 1991; Liang et al., 1994; Arnold et al., 1998; Singh et al., 2002; Blöschl 2006; Gassman et al., 2007; Beven et al., 2012). Current models hold great potential for advancing understanding of complex physical processes across watersheds (Wood et al., 1992, 1997; Koster et al., 2000; Beven, 2001; Nijssen et al., 2001; Yao et al., 2001; Sivapalan, 2005). With the coexistence of deterministic and stochastic hydrological processes, a wide range of models using various formulations of underlying hydrologic processes are now available and widely implemented (Mo et al., 1997; Kochendorfer et al., 2005; Andreadis et al., 2006).

Physically-based distributed hydrologic models are often favored over more empirical approaches due to their ability to better represent underlying theory (Refsgaard

et al., 1996; Boyle et al., 2001). Such models consist of a variety of function categories, which are parameterized to represent the target physical systems through state variable responses (Refsgaard, 1997). Many advantages and disadvantages arise as a result of these approaches. As models have become more sophisticated and complex, our ability to accurately simulate certain physical processes has also grown (Beven, 2006; Schulz et al., 2006; McDonnell et al., 2007; Torch, et al., 2009). However, these advancements have increased the number of required model parameters; accordingly, the calibration scheme has grown in complexity and requires greater computer power (Duan et al., 2003; Cui et al., 2005; McCloskey et al., 2011; Ray et al., 2012). For example, spatial heterogeneity of rainfall-runoff processes can be investigated using distributed hydrologic models (Liang et al., 1994, 1996). Numerous studies have investigated approaches for specifying spatially explicit parameters within distributed models (Hundecha et al., 2004; Panday et al., 2004; Vrugt et al., 2005; McDonnell et al., 2007; Meins, 2013; Euser et al., 2014). However, the confounding interactions between spatially and temporally varying parameter identification remains a vexing problem. The objective of this study is to propose and demonstrate a framework for investigating parameter sensitivity in both spatial and temporal dimensions in order to reduce over-parameterization and increase computational efficiency.

## **2.2 Background**

The increased complexity and parameter dimensionality of hydrologic models introduces significant challenges to traditional calibration methods by considering interactions between physical processes and parameters (Bastideas, 1999). Taking into



account the parameterization of modes to represent surface and subsurface dynamics for hydro-model descriptions, parameters are generally implemented into two types: (1) conceptual parameters which should be investigated through calibration; and (2) physical-based parameters which are based on knowledge of hydrologic process. These two types create big challenges to estimation scheme induced from either the inherent uncertainty in properties or uncertainty stems from parameter interaction or correlation.

In traditional application of models, it is common to employ sensitivity analyses, calibration, and validation to pursue optimal model performance (Refsgaard,1996, 1997). By varying an input parameter value, the modeler can quantify the sensitivity of the output response and thus identify which parameters required the greatest attention. Few studies include more advanced analyses to investigate parameter interactions within sensitivity analyses. However, recent studies have focused on advancing parameter selection and the techniques have grown to be powerful to account for the evaluation of complex physical processes in the watershed, especially to insufficient information regions, including improving model properties, discretizing parameter categories, and capturing statistical input distributions (Abdulla et al., 1997; Liang et al., 2003; McDonnell et al., 2004; Wagener et al., 2007; Zeug et al., 2007; Muleta, 2007).

Improvements in model calibration have adopted reductions in the number of parameters caused by over-parameterization (Pokhrel et al., 2008; Schoups et al., 2008; Whittaker et al., 2010). Over-parameterization can result in misinterpretation of model performance by achieving model “fit” while improperly representing underlying processes (Schwarz, 1978; O’Connell, 1990, 1998;). Although personal computers and workstations have grown tremendously in power in recent years, computational

efficiency remains an important consideration when simulating large domain or working within a stochastic framework.

In most practical cases, distributed hydrologic models are considered to better represent the spatial heterogeneity of hydrologic processes and model parameters across a watershed – particularly for large scale simulations (Beven, 1996). Distributed approaches allow the modeler to represent small-scales physics at the grid-cell or sub-watershed scale with a specific resolution that is appropriate for the process and parameter heterogeneity (Beven, 1989). The application of distributed models has had far reaching implications for regional-scale physical processes such as investigations of climate change impacts (Richard, 2002; Christensen and Lettenmaier, 2006 ). The parameters are directly selected through relationship from soil properties and topography (Wagener, 2007; Brown et al., 2005; Samuel et al., 2011), hence reducing the number of conceptual parameters and decreasing the extent of uncertainty. However, these approaches require significant data and computational resources when applied to large-scale investigations (Shrestha et al., 2006).

Another important consideration of model parameters is that they are not only spatially variable, but they also have a dynamic response with time (Wagener et al., 2002b; Singh et al., 2008). This has often been addressed in previous studies by simply considering the scenario study or digital computation within geographic information systems (Miller et al., 2007). However, the time-dependent nature of parameter sensitivity and performance goes beyond physical meanings but also to be calibrated parameters with unstatic properties. That is, parameter sensitivity can vary across the temporal domain due to underlying physical processes such as a heavy rain event or a

drought – even within a given season. For example, NOAA adopts an ensemble method to statistically analyze extreme events in historical records (Hamill et al., 2013). Wagener et al. (2001) suggests the application of specific model structures to highlight streamflow simulations by partitioned hydrograph into quick and slow flows. Hence, to detect model performance, the modeler should not only consider process and parameter heterogeneity in space but also dynamic changes in time.

In this research, we propose a novel approach for investigating parameter sensitivity and identifiability (defined in Chapter 1) across both spatial and temporal domains to reduce over-parameterization and computational resources for large-scale simulations. The framework is demonstrated through application of the Variable Infiltration Capacity (VIC) model (Liang et al., 1994, 2003; Lohmann, et al., 1996) to the Gila River Basin in the southwest United States. Further, the proposed approach includes a mechanism for segregating the simulation domain into sub-regions based on prevailing climatological conditions (e.g. aridity). The framework was applied to address the following two questions: (1) Is parameterization and complexity of parameter representation variable across climatological conditions? and (2) Can temporal variability of parameter sensitivity and identifiability be characterized across a simulated time series?

## **2.3 Methods**

### **2.3.1 Framework Overview**

The following steps were completed to investigate the research questions. First, the VIC model was developed for the study region and a traditional calibration approach was applied to investigate model performance based on status-quo techniques. The

calibration was applied for three subbasins categorized by climatological condition (aridity). Second, a parameter sensitivity and identifiability analysis was completed for each subbasin in order to investigate the influence of climatological conditions on these procedures. Finally, the temporal nature of parameter sensitivity and identifiability was studied over a simulated time-series for each subbasin. Each step is expanded upon in the following sections.

### **2.3.2 Hydrological Model**

The VIC model is a large-scale, semi-distributed hydrologic model (Liang et al., 1996; Liang et al., 2003; Lohmann, et al., 1996 ) that has been widely applied over the past decade to study climate change and other hydrologic questions (e.g. Matheussen and Lettenmaier et al., 2000; Liang, et al., 1994; Abdulla et al., 1996; Nijssen et al., 1997; Nijssen et al., 2001; Payne et al., 2004; Christensen et al, 2004; Vano et al., 2010). The model is spatially distributed based on a grid cell framework and can be applied under different scale resolutions. Its underlying approach is based on the principle of a “water balance”. VIC is driven by daily inputs of precipitation, maximum and minimum air temperature, and wind speed. Additional model forcing data such as solar radiation, relative humidity, vapor pressure and vapor pressure deficit, are calculated in a preprocessing step within the model (Elsner et al., CIG report 2010).

For this study, the VIC model was implemented at a 1/8 degree latitude by 1/8 degree longitude resolution across the Gila River basin (Figure 2-2). Three subbasins were selected in order to investigate the influence of climatological conditions on parameter sensitivity and identifiability. The subbasins were identified based on their Dryness Index (DI) (Atkinson 2002), and included the Salt River Basin (wettest and 81

grids are included due to adopted resolution), Upper Gila River Basin (drier and 46 grids are included due to adopted resolution), and San Pedro River Basin (driest and 34 grids are included due to adopted resolution) (Table 2-1). The model was driven by observed meteorological data including daily precipitation, maximum and minimum temperature and wind speed. The data sets were selected from Maurer et al. (2002) and calibration was performed in 3-hrs time step with monthly outputs for analysis.

The model was parameterized using a modeling framework of three soil layers. The distribution of water through the three soil layers was allocated based on water flux and storage (evapotranspiration, runoff, baseflow, soil moisture, etc). Soil layer parameterization captured heterogeneous characteristics of the geology, soil types, topography, and vegetation. Five parameters were investigated for calibration:  $b$ ,  $D_s$ ,  $D_{smax}$ ,  $W_s$ ,  $d_1$ ,  $d_2$ . The parameters are defined in Table 2-2.

### 2.3.3 Objective Function

The Nash-Sutcliffe Efficiency (NSE) is a common optimization function (OF) for evaluating hydrologic model performance as a measure of how well streamflow is represented as follows:

$$NSE=1-\frac{\sum(Q_{obs}-Q_{sim})^2}{\sum(Q_{obs}-\bar{Q}_{obs})^2} \quad (\text{Eq. 2-1})$$

Where  $Q_{obs}$  and  $Q_{sim}$  are the observed and simulated streamflow, respectively;  $\bar{Q}_{obs}$  is the mean value of  $Q_{obs}$  (Nash and Sutcliffe 1970). NSE can range from  $-\infty$  to 1. An efficiency of 1 corresponds to a perfect match of simulated streamflow to the observed data. The further NSE departs from 1, the worse the model performance is considered. For more intuitive visualization, the NSE is transformed in this study as 1-

NSE and thus the resulting scale is from 0 to  $+\infty$ , where 0 represents best performance (no error) and higher values represent poor performance.

#### **2.3.4 Sensitivity Analysis Approach**

The sensitivity analysis was applied using the VIC model based on sub-grids and an evaluation of the model performance using a Uniform Random Search (URS) scheme. The URS was applied under different sets of parameters within a Monte Carlo (MC) framework; not a unique global optimum parameter set for the entire watershed as is traditionally applied with this type of exercise. Finally, a Regional Sensitivity Analysis (RSA) was combined with an improved Dynamic Identifiability Analysis (DYNIA) to quantify the variability of parameters and model sensitivity both spatially and temporally.

The URS approach was applied to estimate the parameter sensitivities in the VIC model based on the three climatic regions (wet, dry, and intermediate). The upper and lower bounds of the parameter space were initially selected based on prior knowledge, field measurements, or literature review (Meyer et al., 2007) as the basis of Monte Carlo sampling. A uniform prior distribution was then applied to investigate data outliers (Table2-2).

The RSA method (Spear and Hornberger, 1980; Freer et al. 1996; Wagener et al., 2003a) was used to measure the sensitivity of the individual parameters with respect to the specific OF. The RSA approach can retrieve the information from a cumulative distribution for each analyzed parameter. It works on the feasible parameter space, which is created from the Monte Carlo URS results (Rose et al., 1991). The original RSA can simply separate the parameter population into two groups: behavioral and non-behavioral parameters. Based on the improved RSA approach proposed by Freer et al. (1996), the

parameter space was split into 10 groups of equal size, ranked by their OF values. The cumulative distributions for each group were then plotted for investigation. The differences among the distributions by slope show the extent of parameter sensitivity where steep slopes represent higher sensitivity. The results also can be used to visually evaluate the identifiability of a given parameter. Because the cumulative sum of a uniform distribution is a straight line, departure from a straight line represents a higher degree of identifiability in that region.

Theoretically, any parameter with a physical meaning that is parameterized to represent the underlying system should be both sensitivity and identifiable. A lack of these characteristics represents a failure of parameterization. To resolve this problem, the DYNIA method was adopted to identify periods of high and low identifiability of model parameterization in time series. The procedure was applied through the following six steps: (1) build the feasible parameter space through a Monte Carlo URS framework with a single OF; (2) select the optimal parameter sets under a defined threshold (e.g., the best top 20%); (3) sort the posterior parameters equally sized groups (e.g. ten groups) based on the OF; (4) draw the cumulative distribution curves for each group; (5) calculate the identifiability based on the RSA method for each curve; (6) plot the results as a colored grid over a time series.

## **2.4 Case Study**

### **2.4.1 Research Domain**

The Gila River drains approximately 160,000 km<sup>2</sup> of southwestern New Mexico and southern Arizona within a mostly arid watershed. The watershed experienced

significant human development in the 20<sup>th</sup> Century – particularly in the form of water diversions for urban and agricultural use along with construction of flood control structures. The degree of hydrologic alterations present in watersheds like the Gila present another major challenge for development and calibration of hydrologic models. In order to minimize the impacts of diversions on our study results, simulations were conducted on watersheds located above USGS stream gages 09498500 (Salt River), 09430500 (Upper Gila River), and 09471000 (San Pedro River) which are all located upstream from major reservoirs or diversion projects and that represent relatively natural flows.

The characteristic of the three subbasins are summarized in Table 2-1 and their locations are shown in Figure 2-2. Dryness Index (DI) represents a ratio between potential evaporation and precipitation (reference) and can be used as a simple indicator of the local physical environment. A higher DI value represents a more arid catchment. Based on the dryness index, the Salt River basin is the wettest, the San Pedro River basin is dry, and the Upper Gila River basin is intermediate.

## **2.4.2 Results**

### ***2.4.2.1 Baseline – Traditional Calibration***

Figure 2-3, Figure 2-4 and Figure 2-5 demonstrate the model performance based on simulated vs. observed streamflow for the three subbasins using a traditional calibration method. As traditional calibration method procedure, the first 6 years (1975.1.1-1980.12.31) were used as a warm-up period and the simulation process spanned the next 20 years (1981.1.1-2000.12.31). The calibration approach involved first defining the objective function and then adjusting input parameters to minimize error



over the entire time series. The calibration approach can be adjusted based on different periods in the time series and different time scales. Here, all six calibration parameters were adjointly adjusted to find the best model performance while considering parameter interactions. This was accomplished using a Monte Carlo Uniform Random Search (MCURS) method to reach the global minimum errors between simulated and observed hydrological behavior. Plots (a) in each figure show the best model simulations through selected time series. To be clear a visual simulation of the runoff volume with respect to selected objective function, the bias plot was shown in (c) within each figure. The extent of model performance works pretty well in wetter regions, comparing to the higher bias in drier regions with much worse simulations. During the low yielding time period, base flow is the dominate component during modeled simulations with minor errors which are shown in plot (b) of the Figure 2-3, 2-4 and 2-5.

Figure 2-6 shows a selected class of simulations accomplished by tuning all 6 calibration parameters at the same time based on the NSE objective function. In avoid to embed information due to input spread variance, we categorized the random simulations as equal ten groups based on rank of performance instead of good or bad criteria. Select the best performance or reasonable mimics to physical reality as representation to group a class of time series simulations. The bounds of “cluster” could be an indicator of uncertainty or effectiveness. The best performance produced a NSE value of 0.8 in the Salt River basin with lower variance compared to other two river basins where a maximum NSE of 0.4 was achieved. It creates significantly perturbations for model performance with respect to catchment system through broad spatial domain. The

parameterization should have varying level of complexity or effectiveness across climatological conditions.

In summary from Figure 2-3, 2-4, 2-5 and 2-6, the parameters that are calibrated to perform model evaluations can create substantial bias. It shows that there are significant errors between observed and modeled streamflow and varies extent of bias based on climate conditions. The wetter region (Salt River basin) performs better than the drier region (San Pedro River basin) with a larger forecast streamflow. In the low-yielding catchments (such as San Pedro River basin) with many zero flow days, it can not provide enough information related to the region's soil moisture status, so that there is a lack of effective mechanism to simulate baseflow and ET. This is the potential reason for bias (Nijssen et al., 1997; Abdulla and Lettenmaier, 1997; and Wooldridge et al., 2003). Thus, the capability of hydrologic models can potentially build a relationship to the climate scenarios (such as arid index) by selected interaction parameters. These forecast biases could be corrected by uncertainty quantification. This represents a potential mechanism to reduce the bias by parameter selections.

#### ***2.4.2.2 Parameter Sensitivity and Identifiability Analysis***

The modified RSA approach was used to identify the sensitivity of the different parameters and to evaluate their relative importance with respect to model behavior. The degree to which the models are sensitive to each parameter can be visualized and the results can then be used to eliminate those that are insensitive and to identify sensitive parameters to focus on for future uncertain analyses. A uniform distribution was assumed for the prior distributions to the Pareto sets. The applications of those methods and

outcomes are achieved by using the toolbox MCAT (Monte Carlo Analysis Toolbox), which was created by Thorsten Wagener et al. and my building of MATLAB scripts.

The RSA results are shown in Figures 2-7 a, b, and c for each of the climatic regions in the form of a cumulative OF distribution. With non-behavioral parameters, the cumulative distribution will approximate a diagonal line. Hence, the extent of parameter sensitivity to model behavior is represented by deviation of the cumulative curves away from the diagonal line. The color of the lines represent the binned performing sets with the best performing parameter sets displayed in pink and the worst performers in light blue. Through visual inspection of each of the six parameters over the three subbasins, the most and least sensitive parameters can be identified.  $d_1$  was found to be the least sensitivity parameter and it can be classified as over-parameterization because its value does not impact model response. Demaria and Wagener et al. (2007) also concluded that the base flow component is typically over-parameterized for water balance systems using a daily time step. The most sensitive parameters included  $b$ ,  $W_s$  and  $d_2$ . The results of the RSA produced curves reveal that model more accurately captures observed conditions under wetter conditions (Salt River) than for the other two regions.

Figures 2-8 a, b, and c are scatterplot representations of the Monte Carlo URS simulations and can be used to evaluate identifiability of the selected model parameter sets based on different climatic regions. The pink diamond designates the “best” optimal value in searching the feasible parameter space. Considering the non-uniqueness of parameter sets, the optimal solution sets through the range of potential parameter values can be investigated visually. The parameters are considered identifiable when one or more solution set performs well based on the selection criteria. Based on Figures 2-8 a, b,

and  $c$ , it is observed that the model parameters  $b$ ,  $W_s$ , and  $d_2$  are identified with a critical minimum value among the scatter points based on the OF. Conversely, parameters parameter  $D_s$  and  $D$  are considered non-identified because no significant optimal sets over the range of parameter values were identified. The results also reveal that the parameter identifiability in the wet region (Salt River basin) was lower than in the drier regions (San Pedro River basin and Upper Gila River basin).

The changes in model parameter identifiability based on climate regions described above can be evaluated based on the physical interpretation of the model parameters. Parameter  $b$  defines the shape of the Variable Infiltration Capacity curve. It describes the amount of available infiltration capacity as a function of relative saturated grid cell area. In other words, it separates the infiltration and direct runoff. With a higher value of  $b$ , the model yields higher surface runoff and lower infiltration. In Figure 2-8, parameter  $b$  is highly identifiable for the two dry basins and the minimum criteria values can be used to identify an optimal parameter space. On the other hand,  $b$  is poorly identified within the wet basin and the optimal sets are widely distributed over the feasible parameter space.

The parameter  $D_s$  and  $W_s$  are both related to the drainage component. For example,  $W_s$  is the fraction of the maximum soil moisture (of the lowest soil layer) where non-linear baseflow occurs. A higher value of  $W_s$  will raise the water content required for rapidly increasing, non-linear baseflow. In other words, with a smaller value of  $W_s$ , the baseflow generation will increase and eventually it will contribute to total runoff. However, under dry conditions little runoff is generated and instead evaporation without drainage will dominate. In the wet condition, the third layer is more likely to experience

infiltration which will contribute to total runoff through increased baseflow. That is why the lowest boundary layer as expressed through  $W_s$  moves left when the region becomes wetter.

Parameter  $d_2$  represents the baseflow component. Spectrum has specific meaning in time series analysis in the Salt River basin as compared to the other two subbasins (Figure 2-8). Accordingly, Figure 2-8 b and c reveal a much broader extent of data spread nearly evenly through all the parameter space, which is different from the parabolic curve observed for the Salt River basin. Thus the  $d_2$  parameter was over predicted by streamflow in drier areas. Conversely in the dry areas,  $d_2$  could be seen as insensitive, but with respect to the specific hydrologic behavior – streamflow, it will be a calibration parameter to improve model performance; especially to the lower values. This provides a potential explanation for better model performance for the Salt River Basin as shown in Figure 2-3.

The case study can be used to identify when it is necessary to categorize the model parameters as distributed (identified) or lumped (non-identified) parameter types. Here we conclude that parameters  $b$ ,  $W_s$  and  $d_2$  should be represented as distributed parameters and  $D_s$  and  $D_{smax}$  should be lumped. Based on selected simulation behavior,  $d_2$  is identified as a key parameter for improving model performance.

#### **2.4.2.3 Dynamic Identifiability Analysis**

The degree to which parameter identifiability varies through the simulated time domain was investigated using a Dynamic Identifiability Analysis (DYNIA), shown here in Figures 2-9 a, b, and c. They demonstrate the nature of the DYNIA based on parameters  $b$  and  $d_2$ . The plots reveal the quantitative relation between hydrologic

characteristics (e.g. high flows) and the parameterization characteristics for each of the subbasins. The grey shading indicates the degree of identifiability for the parameter at the designated time step and parameter value. A darker color represents higher identifiability. The results reveal marked variation in the of parameter identifiability between low flow and high flow conditions and as a function of aridity condition.

By comparing DYNIA results across the three subbasins, it is clear that parameter identifiability is dynamic and related to the aridity. Wide confidence limits were observed for the Salt River basin, which indicates that the range of parameters that produce an equivalent OF are widely distributed through the parameter space. In contrast, the San Pedro and Upper Gila subbasins revealed relatively narrow confidence limits. Thus the region of parameter space producing high-performing results are concentrated in a small space range. Thus, attention should be paid to calibration of the VIC model under the wetter conditions because greater uncertainties are present.

The primary motivation behind the DYNIA was to evaluate whether parameter space changes dynamically over the course of the simulation period in response to shifting meteorological and hydrological conditions. If they do exist, such dynamic variations would represent a departure from the status-quo approach of time-consistent parameterization. First, we investigate parameter b and observe less consistency in identifiability for the lower value range for the Salt River subbasin as compared to the Upper Gila and San Pedro subbasins. Therefore, it is justified to simulate hydrologic characteristics within a climatic domain.

Next, it is observed that there are two shifts in the range of optimal values within the Salt River; both associated with wet periods in the streamflow record (around time

step 30-50 and 115-150). During these periods, there is a marked shift in the identified parameter space towards larger values. Likewise, identifiability drops within lower parameter space at these same points. These distinct shifts with respect to the identifiability range highlights the inadequacies and inflexibility of typical model structure components. A more adaptable variable parameterization process based on dynamic watershed conditions as a function of time could provide an improved modeling framework.

Another limitation of static model parameterization is revealed through the temporal nature of the parameterization confidence limits. For example, consider time steps 135-155 on the Salt River simulation results. At this point, the confidence limits expand during the occurrence of a large flow event. This behavior reveals that simulation results are highly uncertain during this extreme flow event. Such behavior could be addressed by considering an alternative OF to improve identifiability or by calibrating the model specifically for this sub-period of the modeling time domain (Gharari et al., 2013).

Similarly, the parameter  $d_2$  shows a distinct area of identifiability by lining up the upper section of the feasible parameter space with narrow confidence limits within the San Pedro subbasin, which is in contrast to conditions observed in the Salt River subbasin.

## 2.5 Discussion and Conclusion

Hydrologic models have grown in complexity and sophistication over the past several decades. High performance computers and advanced GIS capabilities have led to great advancements in hydrologic modeling approaches. Large scale hydrologic studies

are now commonplace and at the same time becoming more sophisticated through improved understanding of model behavior. With the coexistence of deterministic and stochastic hydrological processes, the growing complexity of models is accompanied with a high degree of parameter uncertainty. The framework presented here (Figure 2-1) for evaluating parameter uncertainty and sensitivity takes into consideration both basin characteristics (climatological condition) and temporal variation in parameter identifiability. The method can also result in reduced computational demand, which can improve the efficiency of running uncertainty analyses.

In this study, a novel approach has been proposed for evaluating parameter uncertainty that takes into account basin-wide variability, cross-subbasin variability, and temporally based (dynamic) uncertainty based on aridity. The goal of this framework is to consider the spatial and temporal heterogeneity of physical properties and processes in order to optimize model performance and computational efficiency. With a thorough analysis and case study in the Gila River basin, this model is demonstrated as a robust and flexible approach for improving hydrologic simulation efficiency.

Although the framework can consider the condition where lumped models are used, the main point of this framework is to address spatial heterogeneity within distributed models. As applied using the VIC model, uncertainty arises through the spatial distribution from: grid interpolation of rainfall; spatial probability distributions of dynamic soil moisture storage capacity; elevation bands for representation of topographic variability; and spatial geomorphic and vegetation conditions.

The first research question addressed through the proposed and demonstrated framework was whether parameter sensitivity and identifiability vary across



climatological conditions? Based on the analysis results presented in Figures 2-7 and 2-8, there is a clear argument that the application of climatic-based subbasins can improve parameter sensitivity and identifiability. By selecting subbasins based on a dryness index within the Gila River basin, the variable behavior of parameters directly pointed to their physical meanings. Parameters  $b$ ,  $W_s$  and  $d_2$ , which relate to drainage and baseflow components, show good identifiability in the drier basin. However, the parameters were not well identified in the wetter Salt River basin. This result verified what was observed in the baseline simulation calibrations where better model performance was involved in the Salt River Basin and worse performance was observed in the drier San Pedro and Upper Gila Basins. This result was consistent with Abdulla and Lettenmaier (1997) who also pointed to better VIC performance in humid catchments than dry ones using regional parameters.

The study results also revealed that parameter  $d_1$  was over-parameterized and thus dimensionality can be reduced by setting it to a constant value. A similar conclusion was drawn by Demaria and Wagener et al. (2007). Again, this result can be explained through the physical interpretation of the parameter and the processes that dominate the watershed. To insure consistency in the mathematical formulation when evaporation is evaluated under extremely dry conditions, the VIC model evaluates the top thin layer which is typically set to 10 cm thick. Here we conclude that adding additional complexity in this term via spatial variability does not improve model performance.

The improved accuracy of model predictions through additional model complexity and parameterization comes at the cost of increasing model uncertainty. By incorporating both sensitivity and identifiability analyses, the modeler can extract clues

about the role of climatic conditions on influencing the importance of spatially distributed parameters. This allows the modeler to identifying which parameters are best defined as lumped or distributed based on parameter identifiability. Insensitive parameters suggest less dependence for achieving good model performance but also result in high model uncertainty. For example we conclude that parameters  $D_s$  and  $D_{smax}$  can be calibrated on a lumped basis for the study region. Evaluation on a distributed basis will not improve model performance but will increase the computational load. This is an important consideration when undertaking uncertainty analyses that require thousands of simulations.

The second research question aimed to evaluate if temporal variability of parameter sensitivity and identifiability can be characterized across a simulated time series? This question was investigated through the evaluation of the parameter dynamic analysis in order to estimate the variability in time-series simulation. Strong evidence was provided to suggest that temporal variability could indeed be characterized. For example, the parameter  $b$  clearly exhibited a different extent of sensitivity during different prevailing hydrologic condition. During low flow periods, the parameter  $b$  showed a high degree of identifiability over low  $b$  values. Conversely, high identifiability of  $b$  was observed for high values when flows were high. The dynamic variability of parameters leading tied to flow regimes has been observed with the SWAT model (Cibin et al. 2010). This result indicates the traditional calibration method with a single optimal value simulation will not satisfy the inherent dynamic nature of hydrologic processes.

The framework that was proposed and demonstrated here aimed to comprehensively evaluate the roles of spatial and temporal heterogeneity model

parameters and their contribution to model performance and uncertainty. The results revealed that the spatial and temporal identifiability can be used to identify which parameters should be spatially lumped or set as time-constant in order to optimize model performance and computational demand. Further, the novel approach of characterizing subbasins by climatological conditions to further study parameter identifiability proved to be a useful construct. The resulting framework represents an objective and flexible approach to improve the process of evaluating parameter sensitivity and uncertainty.

The contribution of this framework includes: (1) The application of a climatic-based parameter evaluation scheme can reduce the level of model complexity with respect to the assignment of spatially varying parameters. The case-study in the Gila River basin demonstrated the connections between this approach and the physical interpretation of the model parameters. (2) The dynamic identification of temporally varying model parameters revealed high variability of parameters across the time series. This result highlights the inadequacy of the standard modeling approach and can be used to balance tradeoffs associated with a multi-model framework (e.g. including fast and slow processes). (3) The framework allows for more scientifically defined assignment of the initial feasible parameter space for prior distributions as compared to the standard approach that includes more subjective selections. (4) The framework allows for a visual examination of performance and identifiability that is a flexible approach for achieving calibration.

**Table 2–1 Selected basins for study domain**

<b>Sub-basins</b>	Salt River	Upper Gila River	San Pedro River
<b>Streamgauge (USGS)</b>	09498500	09430500	09471000
<b>Location</b>	33°37'10'' 110°55'15''	33°03'42'' 108°32'15''	31°37'33'' 110°10'26''
<b>Drainage Areas (km<sup>2</sup>)</b>	11152	4828	3196
<b>Dryness Index DI=Ep/P</b>	0.8	1.1	1.16
<b>Defined</b>	Wettest	Drier	Driest

**Table 2–2 Six parameters of VIC model for calibration and uncertainty analysis**

<b>Parameter</b>	<b>Range</b>	<b>Unit</b>	<b>Description</b>
b	0.001-0.8	N/A	Define variable infiltration curve shape
Ds	0.001-0.2	N/A	Fraction of Dsmax where non-linear baseflow begins
Dsmax	0-30	mm/day	Maximum velocity of baseflow
Ws	0.001-1	N/A	Fraction of maximum soil moisture where non-linear baseflow occurs
d <sub>1</sub>	0.05-0.35	m	Soil depth of the first soil layer
d <sub>2</sub>	0.1-1	m	Soil depth of the second soil layer

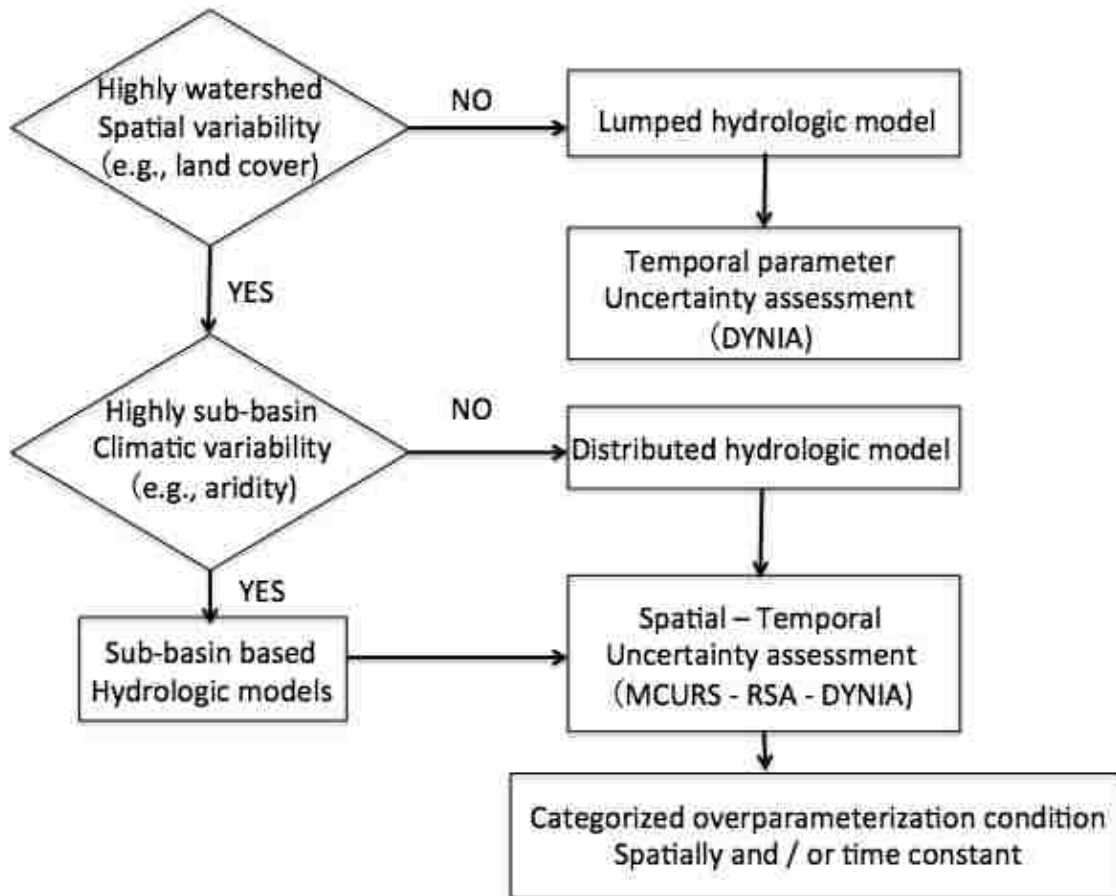
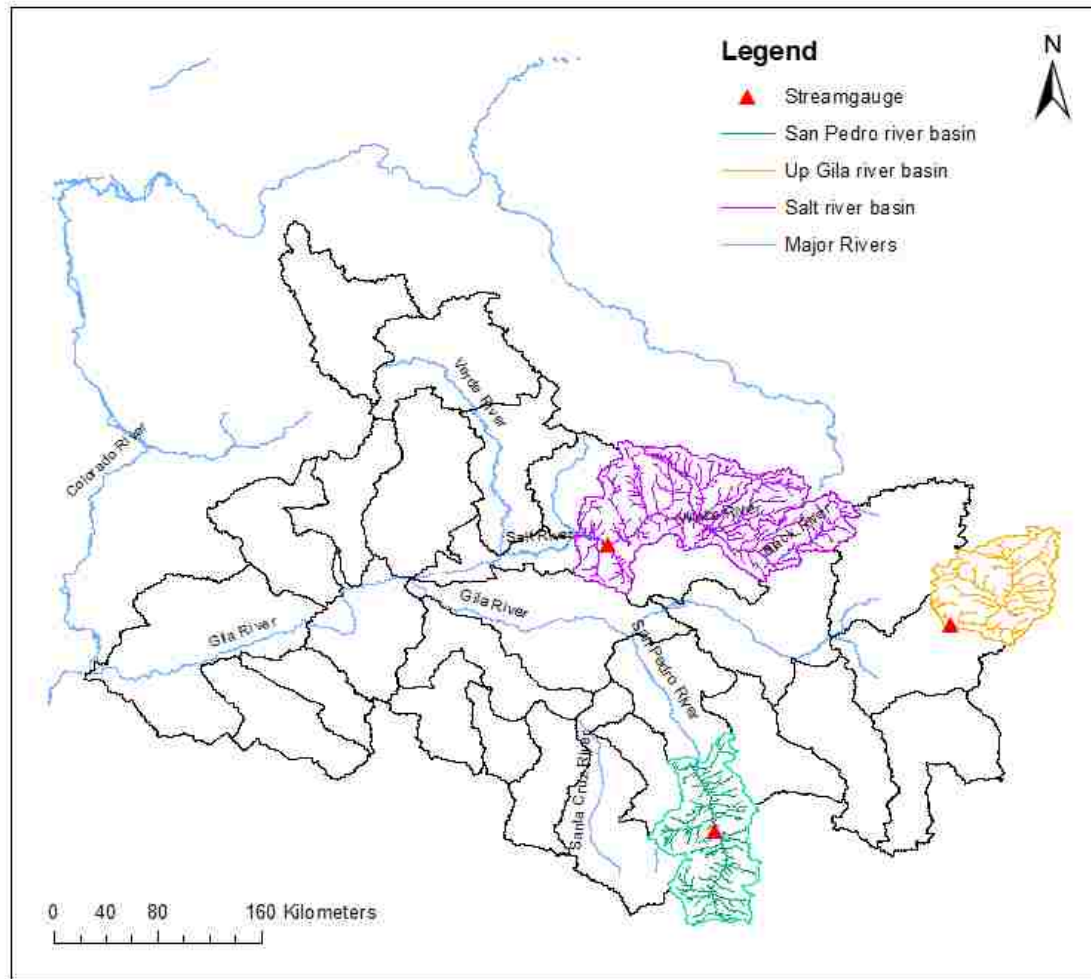
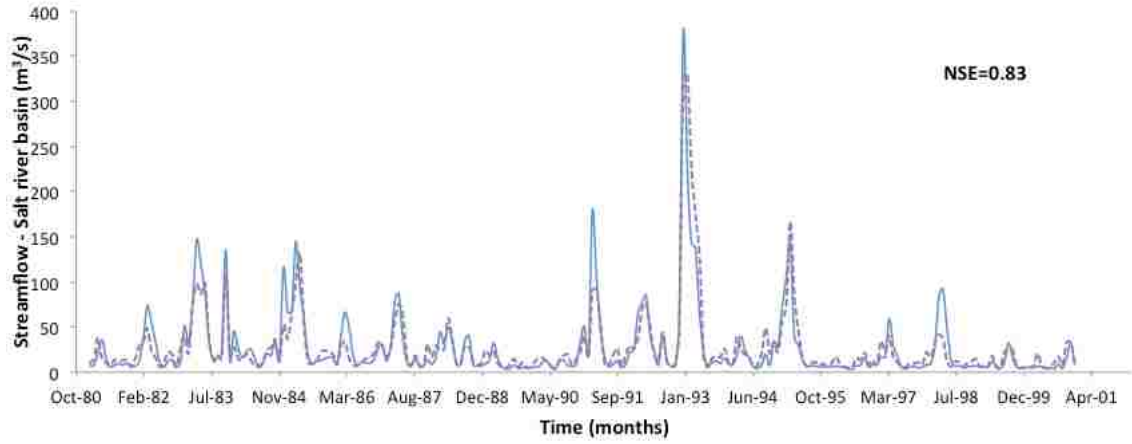


Figure 2-1 A flow chart for the new framework of parameter estimation for distributed large scale hydrologic model

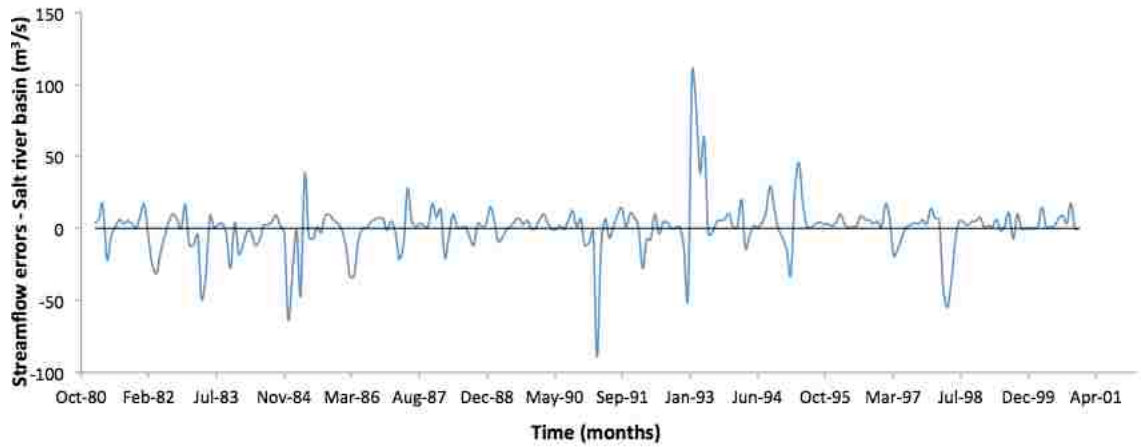


**Figure 2–2 Research domain on the Gila River basin (The blank contour represents the Gila river domain, the color section with purple, green and yellow represents the three subbasins, Salt River basin, Upper Gila River basin and San Pedro River basin; the red triangular remarks the streamgauge for calibration discharge with USGS measurement for each subbasin.)**

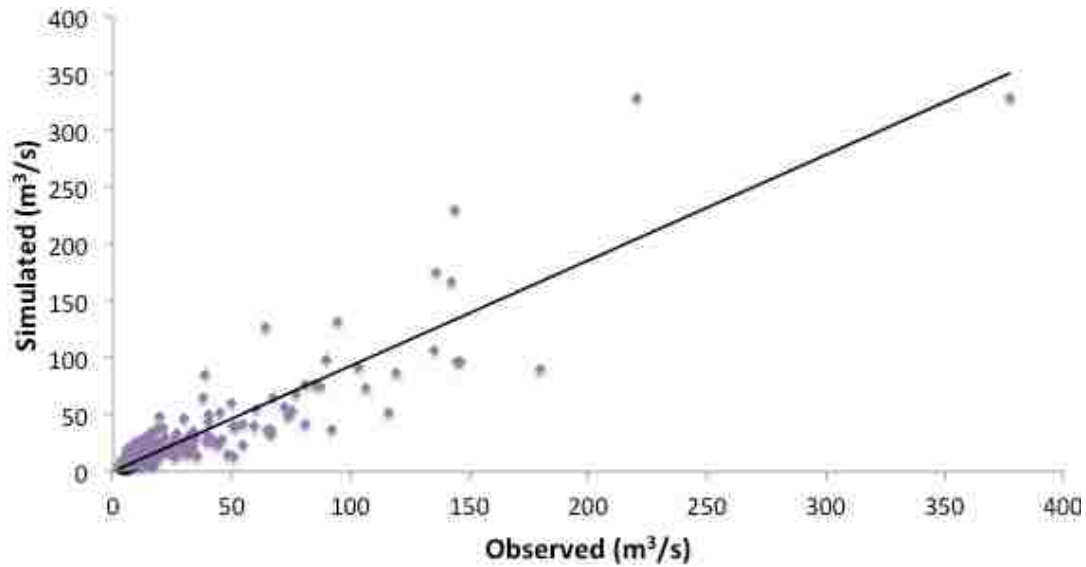
(a)



(b)



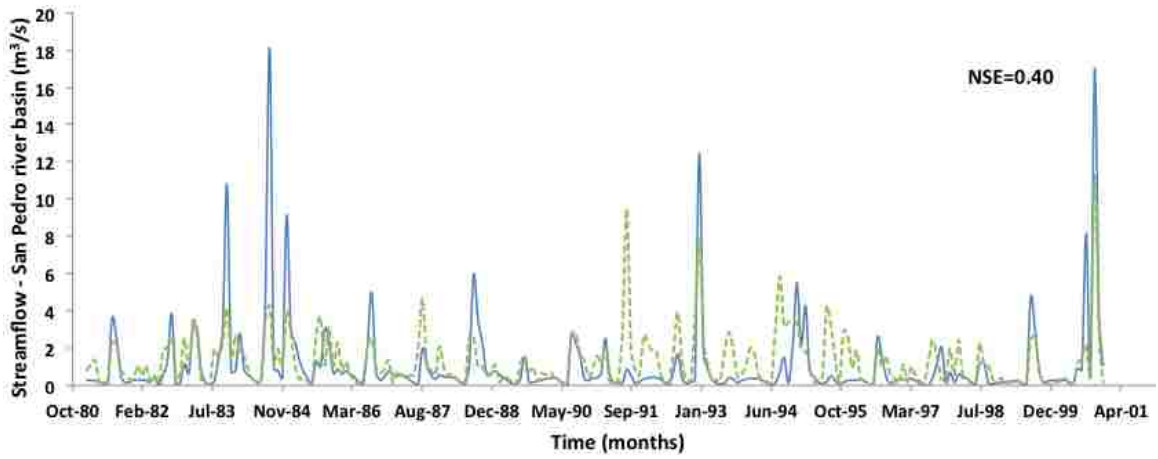
(c)



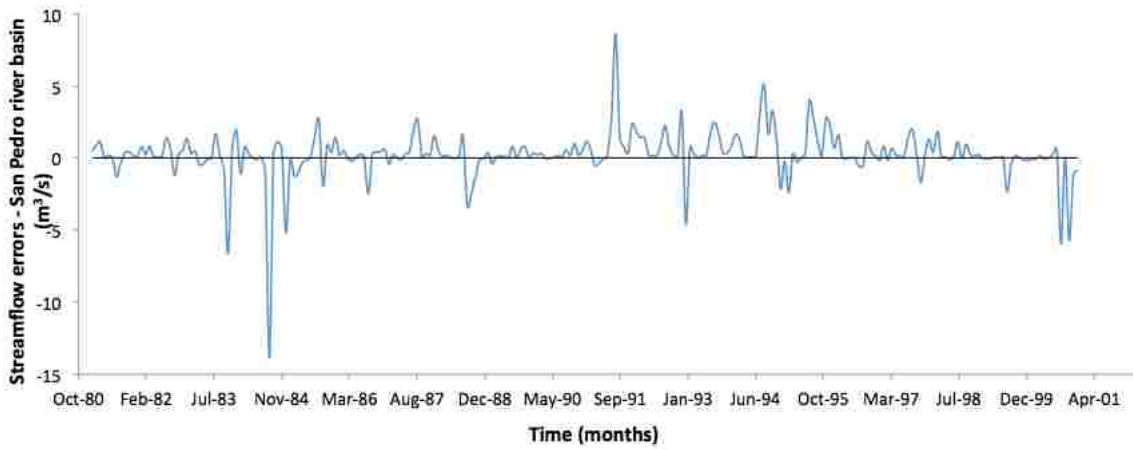
**Figure 2–3 Model performance for Salt River basin. (a) Baseline simulations based on comparison of time-series monthly streamflow (1981.1 -2000.12) by observed (blue line) vs. model calculated (purple dash line) as traditional optimal calibration scheme (with good matching especially on extreme flow). (b) Errors plot on each time step. (c) Modeled vs observed streamflow. It visually shows a bias estimation on this region.**



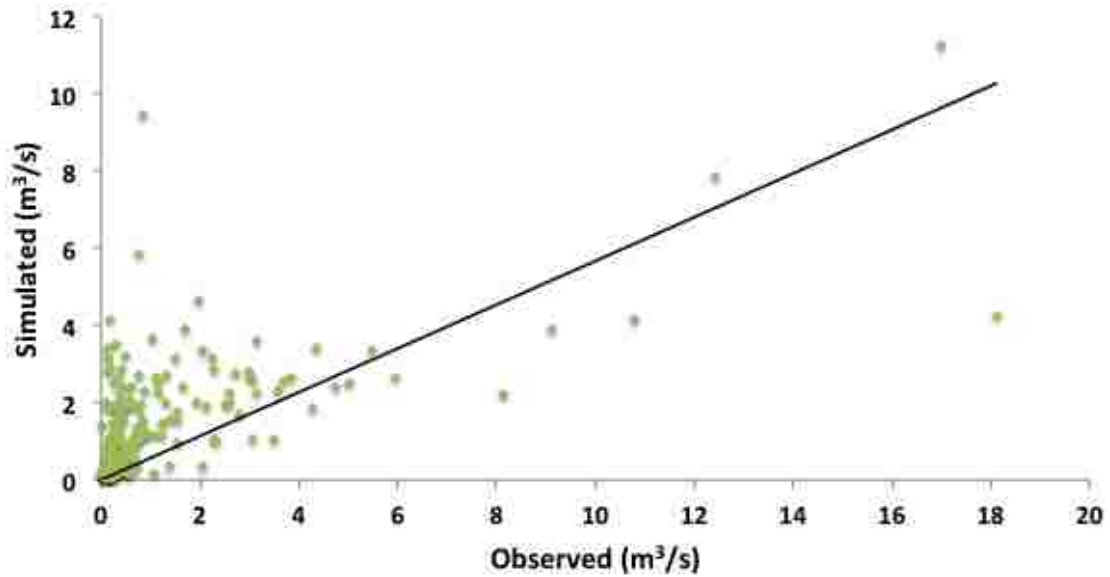
(a)



(b)

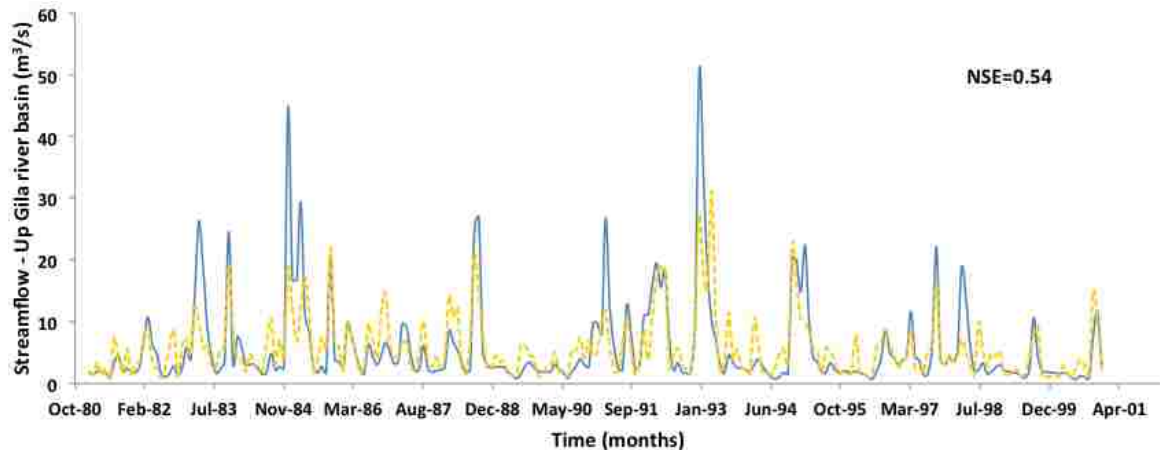


(c)

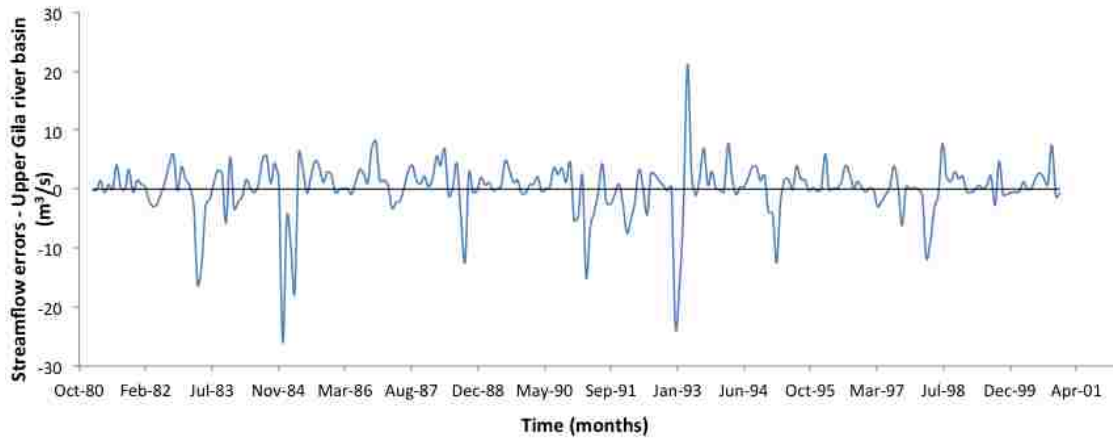


**Figure 2–4 Model performance for San Pedro River basin. (a) Baseline simulations based on comparison of time-series monthly streamflow (1981.1-2000.12) by observed (blue line) vs. model calculated (green dash line) as traditional optimal calibration scheme (worse mimic on extreme flow). (b) Create errors plot on each time step. (c) Modeled vs. observed streamflow. It visually shows a bias estimation on this region.**

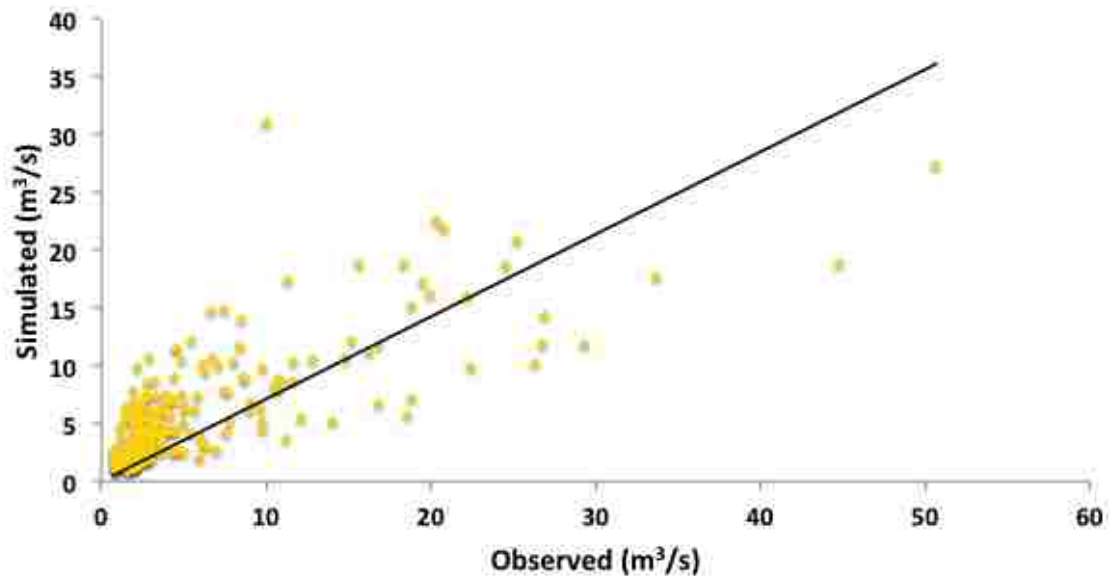
(a)



(b)

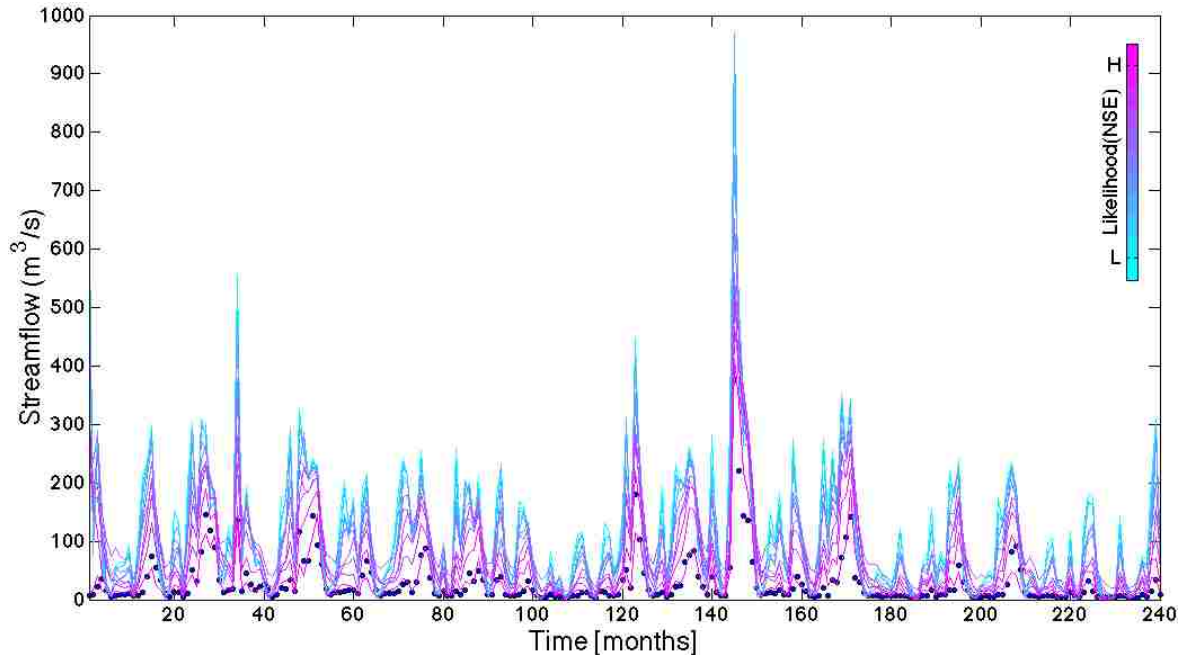


(c)

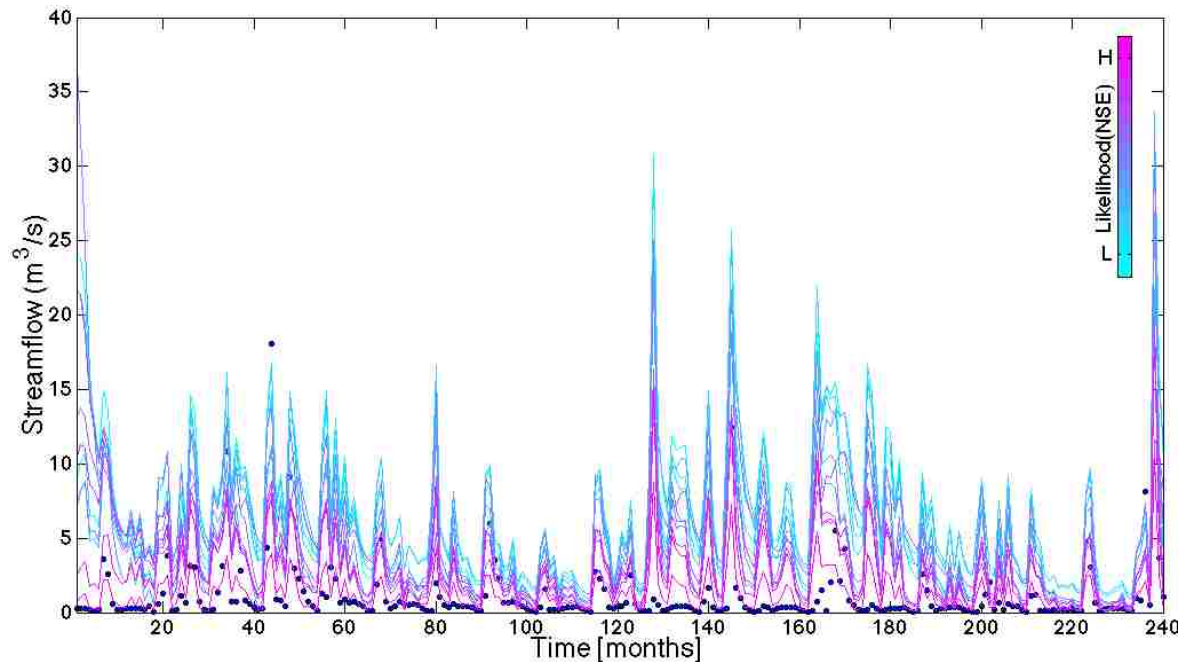


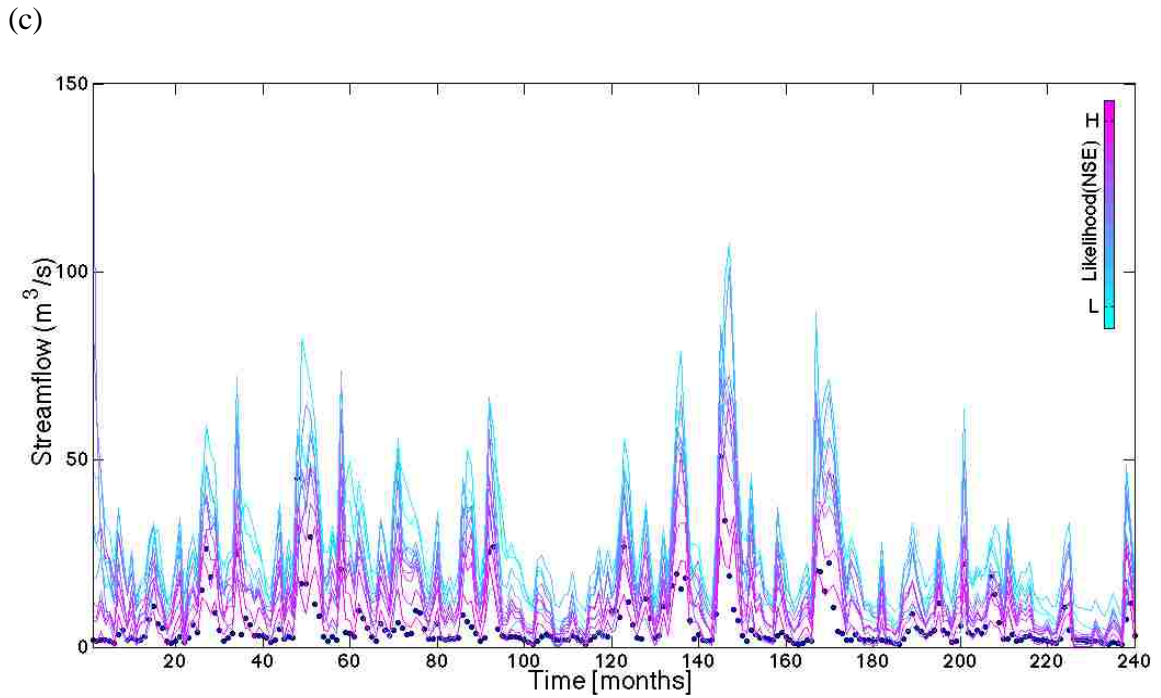
**Figure 2–5 Model performance for Upper Gila River basin. (a) Baseline simulations based on comparison of time-series monthly streamflow (1981.1-2000.12) by observed (blue line) vs. model calculated (yellow dash line) as traditional optimal calibration scheme. (b) Create errors plot on each time step. (c) Modeled vs. observed streamflow. It visually shows a bias estimation on this region.**

(a)



(b)

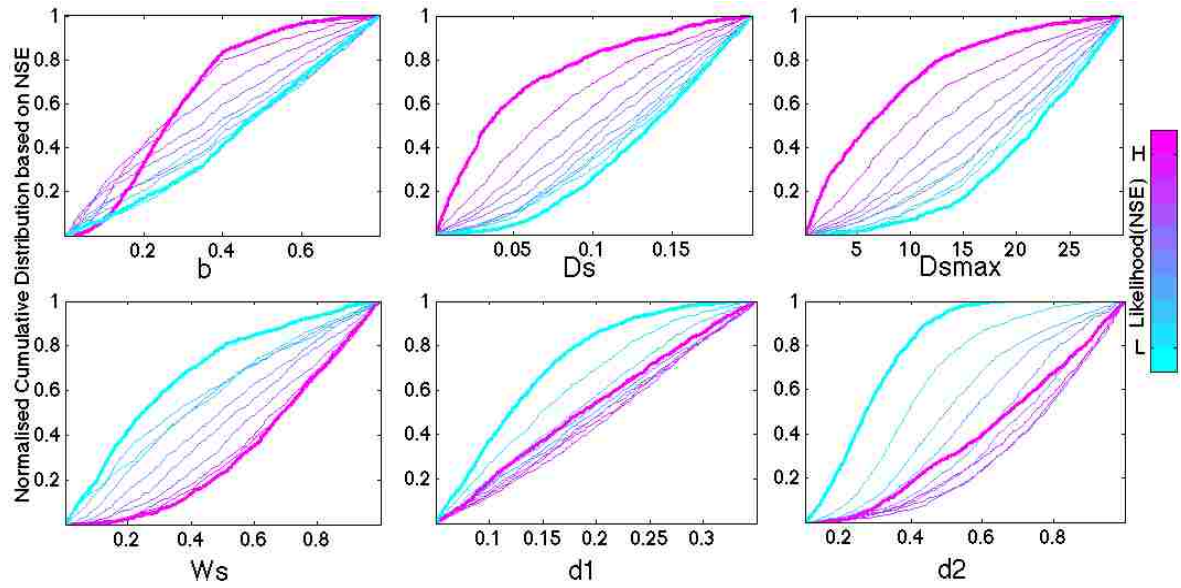




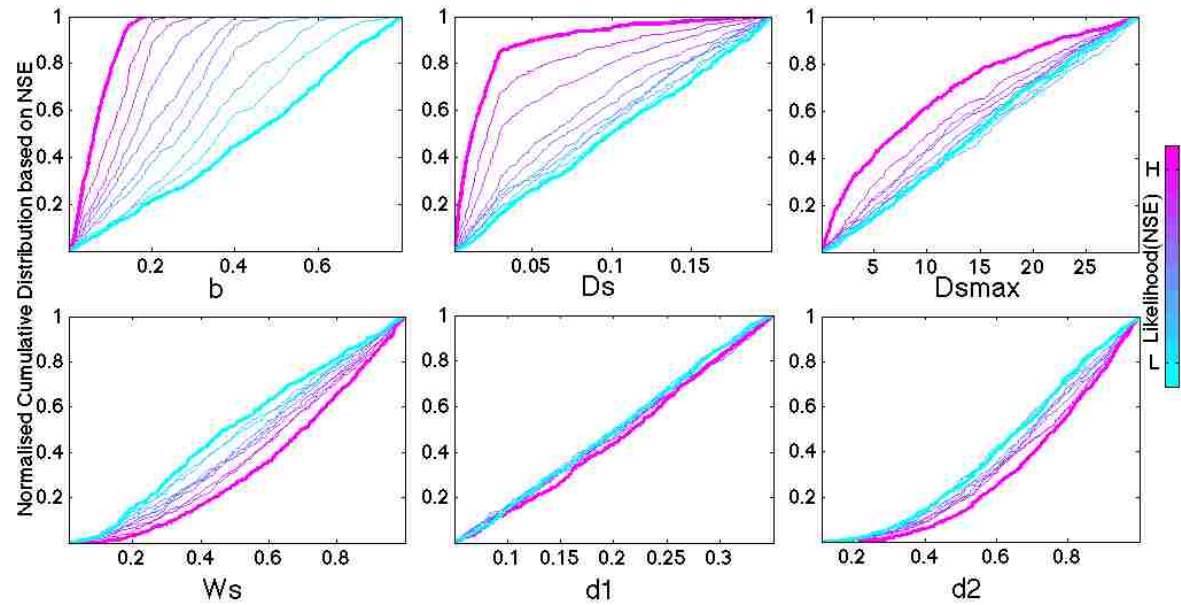
**Figure 2–6 Model performance with a class of outputs based on time-series streamflow simulations for each subbasin (more detail description). The color bar shows likelihoods to each simulations based on criteria function which given higher values with good simulation performance and lower values with bad performance.**

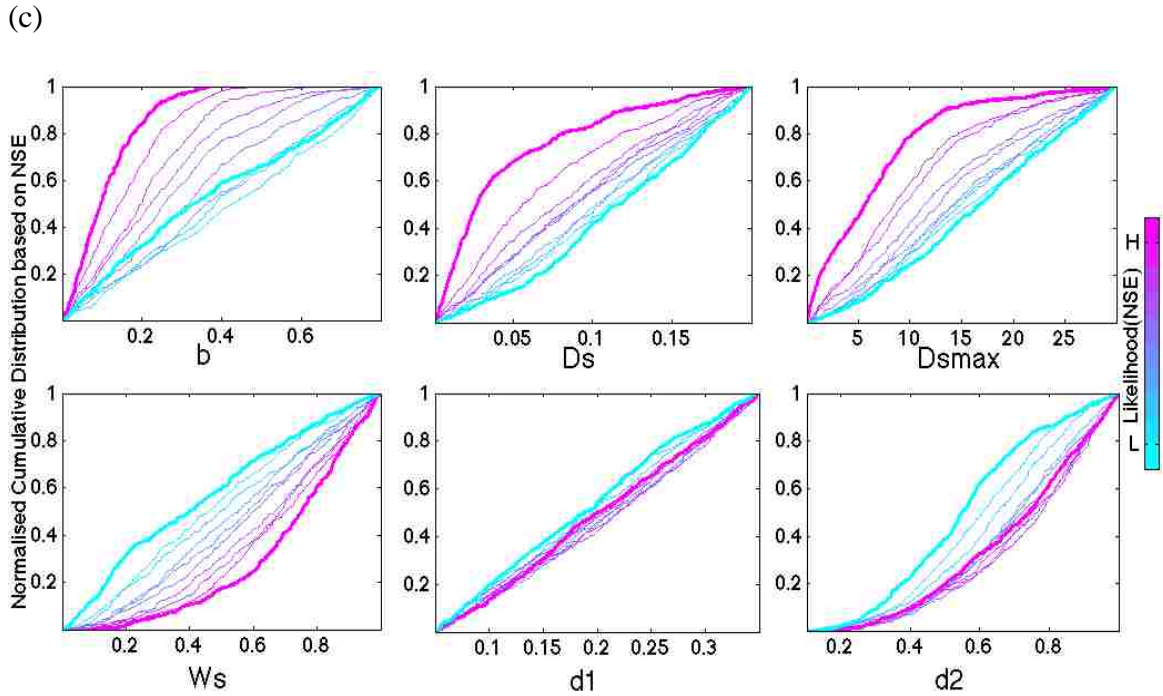
**(a) Salt River basin; (b) San Pedro River basin; (c) Upper Gila River basin.**

(a)



(b)

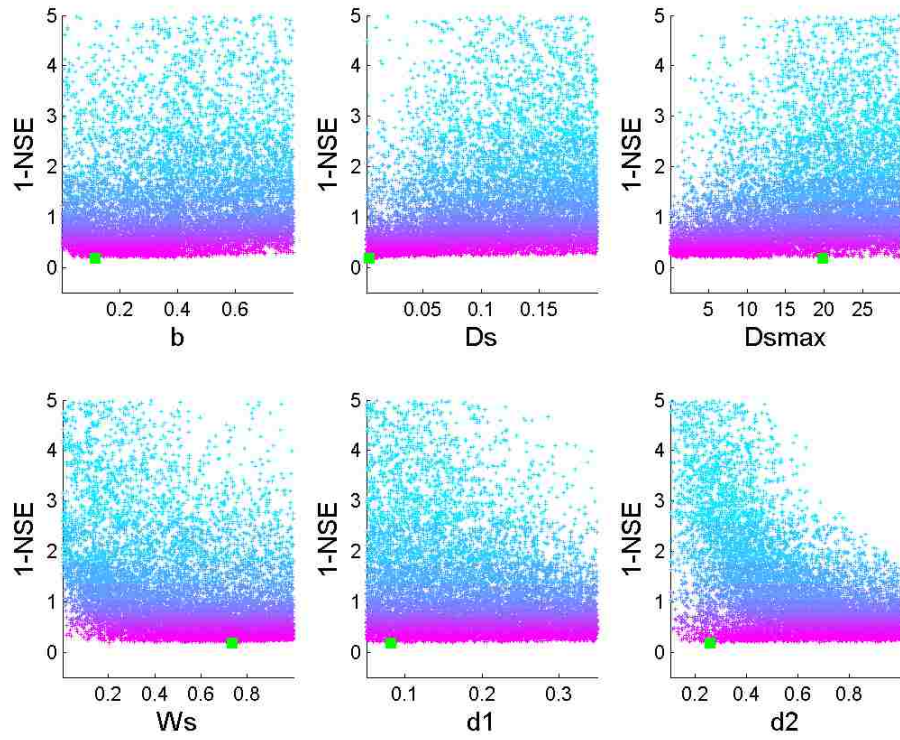




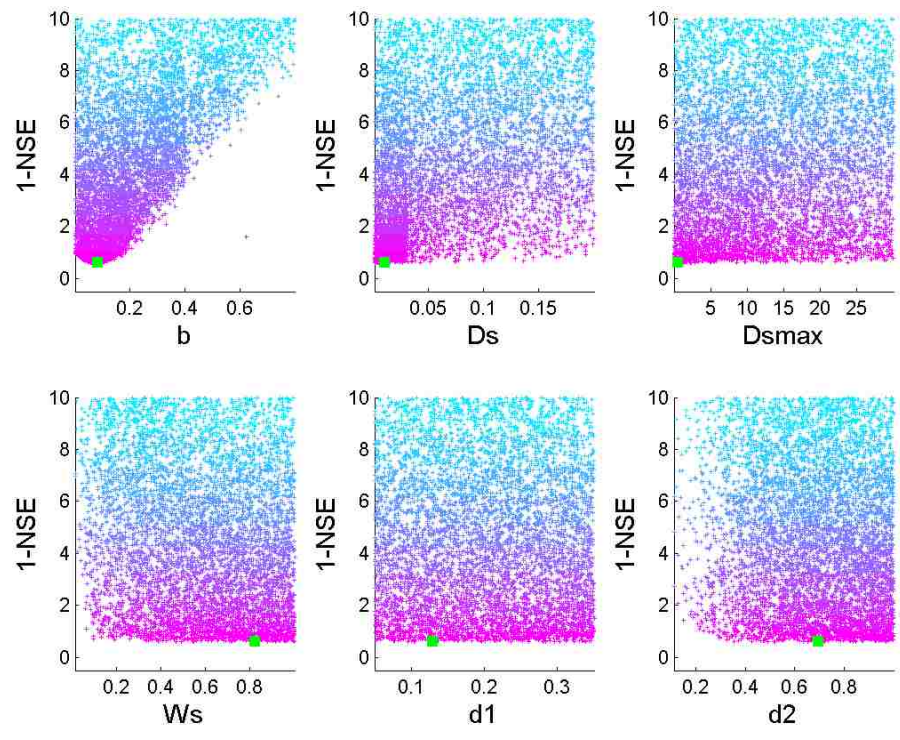
**Figure 2–7 Parameter Sensitivity analysis for each basin. All of these global-population random simulations are divided into equal ten groups based on the ranking of objective functions. The nomalised cumulative distribution is calculated for each group with attained a higher value as better performing model simulations.**  
**(a) Salt River basin; (b) San Pedro River basin; (c) Upper Gila River basin**



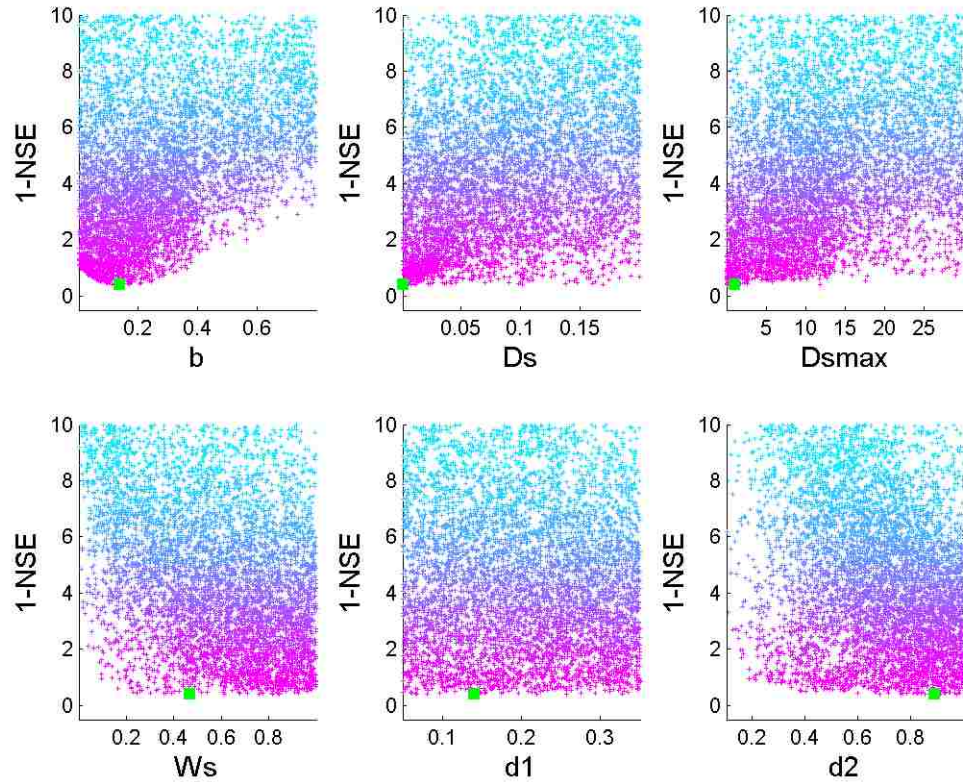
(a)



(b)

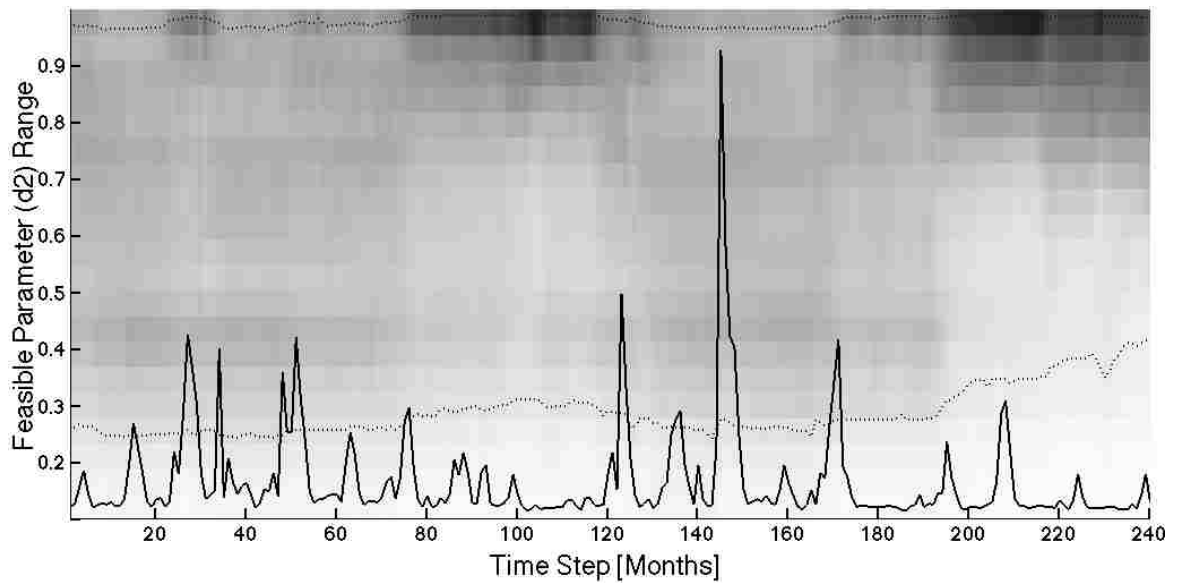
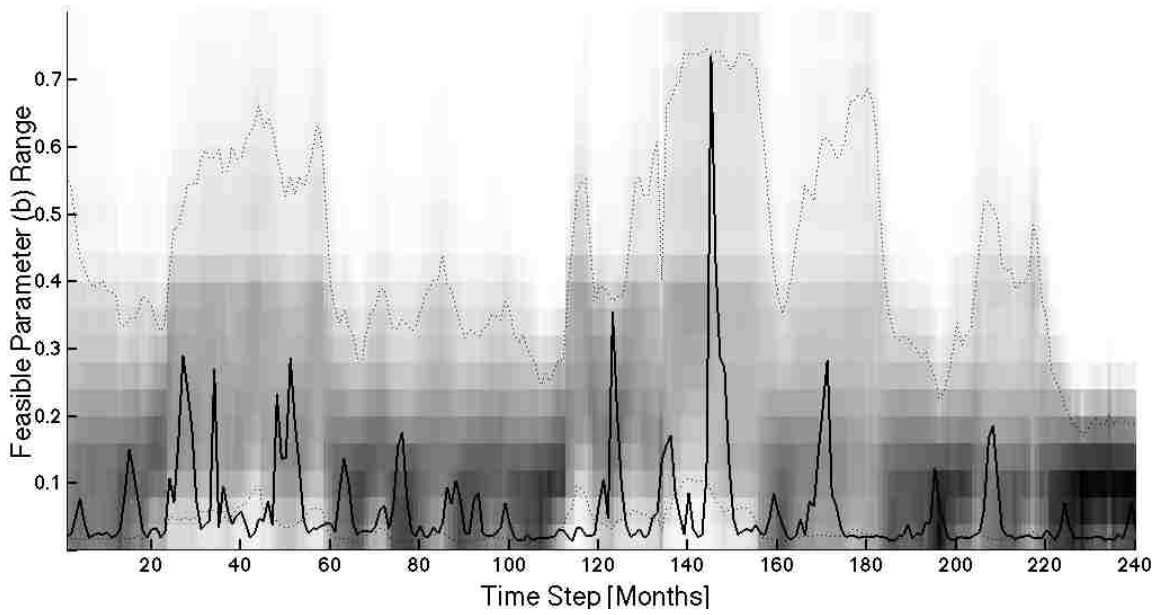


(c)

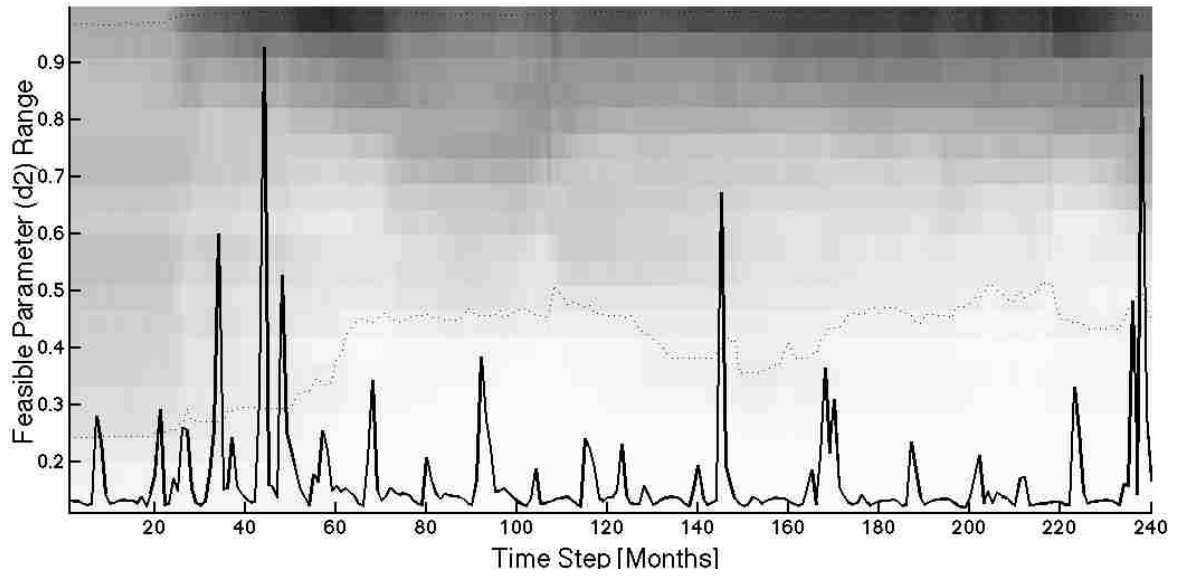
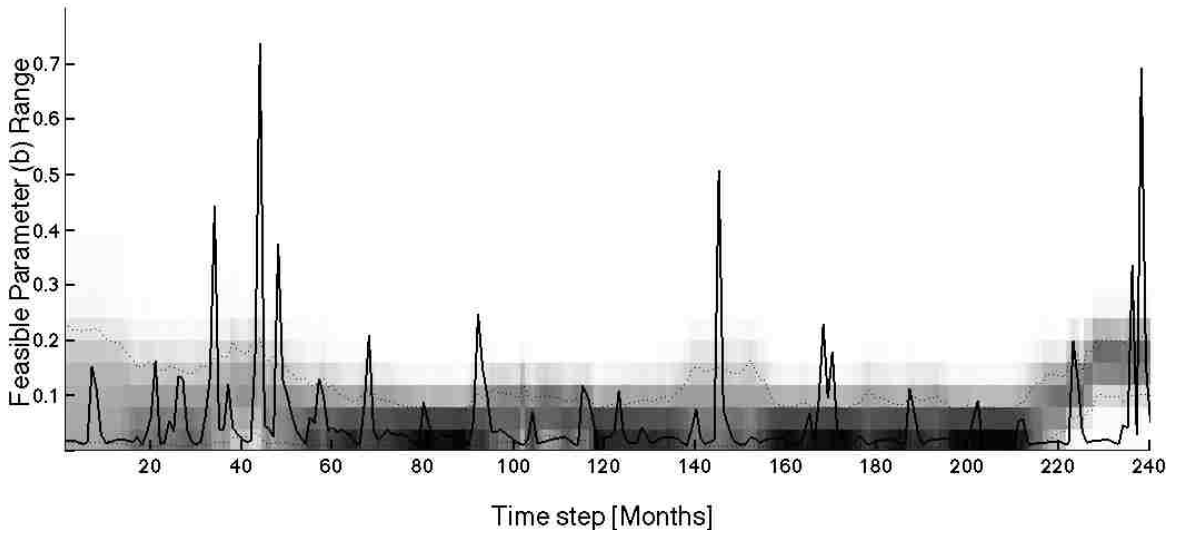


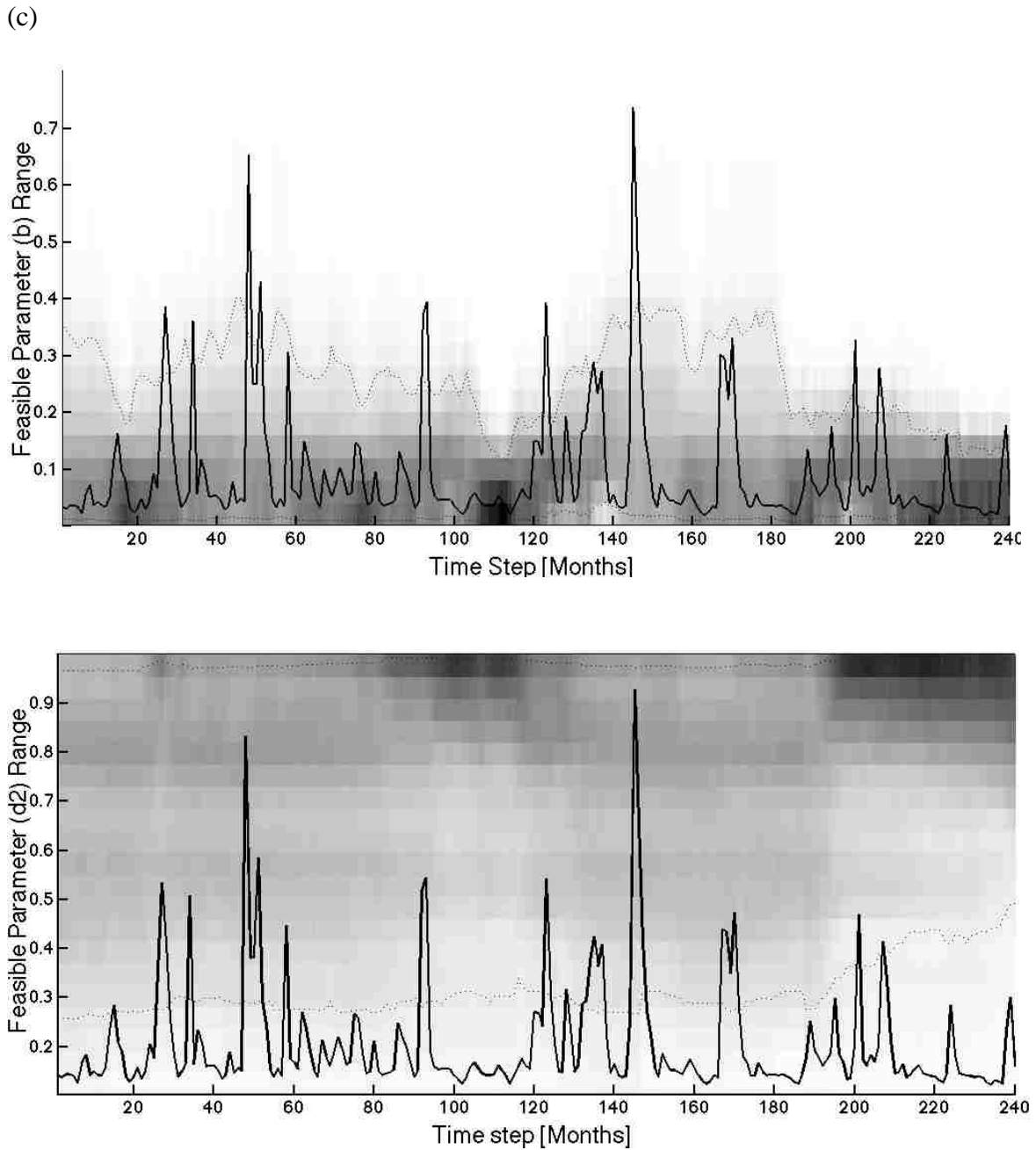
**Figure 2–8 Scatter plots for parameter sets based on model performance. (It is the outputs from Monte Carlo Uniform Random Simulations, the behavior parameter sets are equally divided into 10 groups ranked with model performance given higher value with better performing. The pink dots represent the best top 10% parameters comparing to the worst 10% parts with the light blue section). (a) Salt River basin; (b) San Pedro River basin; (c) Upper Gila River basin**

(a)



(b)





**Figure 2–9 Parameter dynamic analysis for each basin. It is a dynamic measurement of identifiability (Appendix A) to each parameter based on time series hydrologic behavior. Take parameter b and d2 as representators. (a) Salt River basin; (b) San Pedro River basin; (c) Upper Gila River basin**

## **Chapter 3 Integrated Multi-Criteria Estimation under Parameter Uncertainty Quantification for Large Scale Distributed Hydrology Model**

### **3.1 Introduction**

Hydrologic models have proven to be a powerful tool for improving process understanding and providing predictive capabilities. Physically based models can be used to study hydrological behavior across a range of spatial and temporal scales. As the complexity of such models is increased, so are the number of model parameters required to represent the physical environment and processes (Franks et al., 1999). To achieve accurate simulations with respect to specified hydrologic characteristics, it is essential to select and apply an appropriate parameterization scheme. Distributed hydrologic models utilize various integrated routines to represent the storage and flux of water and energy within and between grid cells. Versatile parameterization frameworks represent competing degrees of parameterization (e.g. based on climatic conditions) and various methods for estimating hydrologic behavior (e.g. peak streamflow, water budget, or flood inundation) (Kollat et al., 2012).

Within the various parameterization approaches, it is commonly understood that different combinations of model parameter sets can result in a models that perform equally well with respect to a given objective. This concept is known as “equifinality” and was first emphasized in the field of hydrology by Beven and others (Beven & Binley, 1992; Beven, 1993 and Freer et al., 1996). Thus, the equifinality concept assumes that equally good model performance can result from a branch of a parameter set with multiple possible combinations rather than the traditional hypothesis inherent in traditional calibration of uniqueness of a parameter set to optimize model performance (Beven 2000).

A framework of trade-off measurements has been explored within the research area of model simulation in order to extract information based on the equifinality parameter sets (Gupta et al., 1998; Efstratiadis et al., 2010; Kollat et al., 2012; ). Such approaches consider multiple spatial variables, multiple responses based on temporal patterns, and multi-criteria by varying the terminal criteria (Blasone et al., 2008). However, the existence of trade-offs amongst these approaches reveals that a unique global optimization cannot reproduce the entire hydrological behavior with a single selected model performance measurement (Kollat, Reed and Wagener, 2012; Beven, et al., 2012).

It is necessary to extract and quantify uncertainty information across the range of acceptable parameter space in order to assess the conditions that produce behavioral parameter sets. Behavioral parameter sets are those that perform at an acceptable level when comparing model results with a measurement of performance. In order to provide a robust framework, two questions must be resolved. First, how can we discriminate

behavioral parameter sets as optimal ones from within the random parameter population? As described for the term equifinality, the branch of optimal parameters can all be considered to have an equal (but not equivalent) performance. That raises a second question: how can we reveal the extent of uncertainty by aggregating those optimums?

The goal of this chapter was to propose and demonstrate an integrated, multi-criterion simulation routine based on parameter uncertainty quantification for large scale distributed hydrologic models. The ability to explore different objective functions to create optimal model performance was applied using several hydrological behavior features using time-series simulations. Two objective functions (Nash-Sutcliffe efficiency and Relative Error) were used to investigate model performance based on simulated peak streamflows and water volume, which are two crucial features for simulating streamflows.

In order to evaluate and extract information from the acceptable range of potential parameter sets, the model uncertainty was quantified using two approaches. The first approach was based on the Generalized Likelihood Uncertainty Estimation (GLUE) approach, which was based on a likelihood algorithm using multi-criteria features to evaluate the acceptable parameter space. The second approach was based on the concept of Pareto optimality and a resulting Pareto front. The uncertainties of the simulations were also investigated by both methodologies.

## **3.2 Background**

Many studies have recognized the condition of “overparameterization” where one or more model parameter is effectively inactive for representing the physical processes in a hydrological system (Schwarz, 1978; O’Connell, 1990, 1998; Demaria et



al., 2007). Overparameterization leads to increased computational requirements (Feyen et al., 2008; Whittaker et al., 2010) and can increase model uncertainty (Schoups et al., 2008). In the context of large-scale distributed hydrologic models, the simulations require a large number of parameters in order to characterize spatial heterogeneities across grid scales (Wheater et al., 1999; Decharme et al., 2005; Richard et al., 2013). This condition greatly amplifies the challenges arising from overparameterization. Thus, it is possible to reduce the model dimensionality to a required level through sensitivity analyses. This reduces the dependency of model scales and locations (climatic gradient, mentioned in Chapter 2, Jia 2014) and improves the identifiability of model parameters with optimization techniques.

A robust estimation scheme will reduce the overparameterization while addressing the following issues. First, it should be possible to emphasize key hydrologic aspects (e.g. peak streamflows and volume) by assigning various objective functions to characteristics of interest (Yapo et al., 1998; Boyle et al., 2000; Wagener et al., 2003a). Second, the scheme should be able to consider the tradeoffs amongst multiple objectives. For example, multi-objective analysis can be combined into a representative, conventional single-objective estimation (Seibert, 2000; Blasone, 2008 and Efstratiadis, 2010). Third, the approach should be capable of incorporating multiple indicators of model performance to retrieve maximum information from time-series simulations results (e.g. high flow, low flow and flashiness) (Kollat, Reed, and Wagener, 2012).

To a great extent, quantification of uncertainty can be used to identify feasible parameter space. The GLUE method is the most common approach for quantifying uncertainty in this regard (Beven and Binley 1992). Within the GLUE approach,

likelihood measures are used to assign higher weights to better performing models. Alternative approaches include the simple uniform random sampling method (used in Chapter 2 of this dissertation) (Uhlenbrook et al., 1999), Markov Chain Monte Carlo methods with Genetic Algorithms (Thiemann et al., 2001), Bayesian model averaging methods (BMA) (Ajami et al., 2007), Gaussian mixture model parameters estimation by Expectation maximum algorithm (Bilmes, 1998), and many other techniques for assessing uncertainty with parameter estimation procedures (Efstratiadis 2010).

### **3.3 Methodology for Quantification Parameter Uncertainty with Multi-Criteria Estimation**

#### **3.3.1 GLUE Method Characteristics**

##### **3.3.1.1 GLUE Framework**

The GLUE methodology was proposed by Beven and Binley (1992). The method is based on Monte Carlo simulations with randomly chosen parameter values from a priori probability distributions and the application of Bayes Theorem. It is a robust method for calibration and quantifying uncertainty in hydrologic and environmental modeling related to model parameterization and outputs (Freer et al., 1996; Beven, 1998). Compared to other approaches, GLUE is flexible and it provides a simple approach to distinguish the global uncertainty by exploring interactions of parameters.

The GLUE approach is based on the assumption that uncertainties arise from the equifinality phenomenon (Beven and Freer, 2001). Thus, the uncertainties associated with a set of parameter values (behavioral) are being assigned a likelihood based on the

acceptable criteria of a selected hydrologic behavior. The method is employed using Monte Carlo simulations with a large number of parameter sets, which are chosen randomly from prior distributions (usually uniform distributions). Based on a specific model performance criteria (objective function) model results are evaluated with respect to fit between the predicted and observed responses. A likelihood value is assigned to each set of parameter values with higher values given to better performing parameters. The sum of all likelihood values is equal to one. A threshold value is determined for which parameter sets performing worse than the criteria are described as non-behavioral and the likelihood for such sets is set to zero (Beven, 1992).

### ***3.3.1.2 Objective Functions (Termination Criterion)***

An objective function is used to evaluate model performance and in hydrologic modeling this is typically achieved by comparing simulated and observed results. Traditionally, it is common to select an objective function (OF) as a measurement to aggregate the residual variance through the application of mathematical functions (Gupta et al., 1998). The minimum or maximum values of OFs represent the optimal model parameters. Inevitably, even by development of advanced powerful automatic calibration algorithm, it is necessary to combine with visual inspections to evaluate specific aspects of model performance by examining time series hydrographs (Boyle, Gupta and Sorooshian, 2000). Previous research (Wagner et al., 2001) has shown that a single-criteria approach cannot fit all response model components and may fail to match some physical processes between prediction and observed hydrological behaviors (Efstratiadis, Koutsoyiannis, 2010). Within this study, we demonstrate a novel application by combining two objective functions (multi-criteria) in order to evaluate model

performance with respect to both peak streamflows and water volumes – both of which are important for examining behavior and performance of the hydrologic model.

Two common objective functions were applied in this study. The Nash-Sutcliffe Efficiency (NSE) is considered to evaluate peak streamflows and the Relative Error (RE) is used to emphasize the water balance in simulated time period. The OFs are defined as follows:

$$NSE = 1 - \frac{\sum(Q_{obs} - Q_{sim})^2}{\sum(Q_{obs} - \bar{Q}_{obs})^2} \quad (\text{Eq. 3-1})$$

$$RE = \left| \frac{R_{sim} - R_{obs}}{R_{obs}} \right| \times 100\% \quad (\text{Eq. 3-2})$$

where  $Q_{obs}$  and  $Q_{sim}$  are the observed and simulated streamflow, respectively;  $\bar{Q}_{obs}$  is the mean value of  $Q_{obs}$ ;  $R_{sim}$  and  $R_{obs}$  are the simulated and observed average annual streamflow, respectively.

NSE can range from  $-\infty$  to 1 with an efficiency of 1 corresponds to a perfect match of simulated streamflow to the observed data. On the other hand, RE is spread over a range of 0 to  $+\infty$  where 0 represents perfect model performance. In order to facilitate visual analysis when evaluating model performance, NSE was transformed to 1-NSE with same patterns (0 to  $+\infty$ ) as RE when plotted.

When investigating model performance based on multiple criteria, the modeler must have the ability to weigh trade-offs of model performance based on each criterion. Take Figure 2-8a (Random parameter scatter plot based on criteria values) as an example, following the Monte Carlo streamflow simulation runs outputs from Salt River basin (as shown in Chapter 2), model performance was analyzed but with a different OF, RE. The

goal of using a multi-criteria strategy is to identify the best overall model performance from multiple perspectives as provided by the combination of both NSE and RE parameter estimation processes. Figure 3-4 shows a trade-off curve between NSE and RE parameter estimation approaches. The lower-left corner provides the optimum values for both functions.

### 3.3.1.3 *Building a Relationship of the OF With Likelihood Function for GLUE*

The key feature for the application of the GLUE method in this research is the building of a likelihood function that couples the multi-criteria model performance – as opposed to the traditional use of statistical likelihood. The multi-criteria approach aims to optimize model performance based on emphasis of multiple aspects of the underlying hydrologic behavior (e.g. peak streamflow and water volume) (Gupta et al., 1998; Boyle et al., 2000).

The principle of multi-criteria theory has been proposed and applied to hydrological and environmental simulations by Gupta and others (Gupta, et al. 1998; Yapo et al., 1998; Duan et al., 2013). The theory can be expressed as follows:

$$\text{Minimize } F(\theta) = \{f_1(\theta), \dots, f_m(\theta)\} \text{ subject to } \theta \in \Theta \quad (\text{Eq. 3-3})$$

Where  $f_m(\theta)$  represents one of the model residuals or OFs;  $\theta$  is the vector of calibrated parameters; and  $\Theta$  is the feasible parameter space, which can be identified through the prior uncertainty of parameter distributions.

An aggregating function (Efstratiadis and Koutsoyiannis, 2010) is needed to quantify the combined uncertainty associated with each OF. The function should also allow for an analysis of the trade-offs between multiple OFs (e.g. NSE and RE) in order

to achieve high model performance from multiple perspectives. The aggregation function proposed in this study combines NSE and RE OFs as follows (Blasone, R., et al., 2008):

$$NSERE = L(NSERE) = F(\theta_i) = \frac{1-NSE(\theta_i)}{\min\{1-NSE(\theta_i)\}} + \frac{RE(\theta_i)}{\min\{RE(\theta_i)\}} \quad (\text{Eq. 3-4})$$

The likelihood measurement could be defined as subjective selections connected to terminal criteria (Beven et al., 2007;).

### 3.3.2 Pareto Ranking for Uncertainty Cluster Outputs

The concept of Pareto optimality was proposed in the Nineteenth Century by Italian economist Vilfredo Pareto (1848-1923) (Pareto, 1896). According to the principle, given an initial allocation of goods amongst a closed set of participants, a change to a different allocation that improves conditions for at least one individual without harming other individuals is called a Pareto Improvement. The allocation is defined as “Pareto optimal” when no further improvements can be made via reallocation of goods.

This concept can be extended to the context of optimal model performance. That is, we seek a condition at which point all residuals are simultaneously minimized. As shown in Equation (3-3), a Pareto set is a sub-set of the feasible parameter space that simultaneously minimizes all residuals. It was assumed there are no objective functions that can satisfy all aspects of the time series hydrograph. Thus, a multi-criteria approach was adopted. The performance of an OF will vary depending which aspect of the simulation is emphasized (Gupta et al., 1998). The minimum values of OFs can be scattered in separate regions throughout the feasible parameter space. The area (plotted on a two-dimensional graph) between OFs’ optimal values is categorized as the Pareto space (Gupta et al., 1998; Das, 1999). The Pareto front is selected via a solver by searching the Pareto space and querying the solutions based on rank levels. The first Pareto optimal

front is identified based on interrogation of the whole feasible parameter space and assigned the rank one. The top ranked performer is removed and the interrogation is repeated to identify the second ranked solution amongst the remaining population. The process is repeated until all of parameter sets are ranked.

### **3.3.3 Distinction Between GLUE and Pareto Based Methods**

In this chapter, a multi-criteria GLUE approach and a Pareto approach were used in conjunction to evaluate model performance and uncertainty. From a general perspective, these two approaches have relative strengths and weaknesses. With respect to GLUE, quantitative information can be extracted by a specified confidence level in a percentile format. This approach is convenient and the results are easily conveyed to managers. The likelihood is the key factor that influences the uncertainty. In contrast, outputs from the Pareto approach are used to formulate cluster ranges of simulations with minimum and maximum values constituting the bounds on each time step. The solver used to identify the Pareto front is crucial for ensuring appropriate trade-offs with Pareto optimality. Considering required computational resources, GLUE is more efficient and flexible as compared to the optimizing search used to generate the Pareto front.

## **3.4 Case Study**

The model investigation reported in this chapter was based on VIC simulations of the Salt River basin, which is a tributary of the Gila River located primarily in eastern Arizona (see Figure 2-2). The watershed has a drainage area of 35,000 km<sup>2</sup> and the elevation ranges between 300 m and 3500 m above mean sea level. The annual precipitation within the Salt River basin (Chapter 2) for the period under study (1981.1-

2000.12) ranged from 370 to 900 mm, with a mean of 625 mm and a standard deviation of 134. Dominant land cover in the basin includes cotton, alfalfa, fruit and vegetables.

Using the same general techniques as those described in Chapter 2, a VIC model was developed for the basin on a 1/8 degrees by 1/8 degrees basis. Input data were obtained as same in Chapter 2. A classical single-criteria calibration method was used to generate an initial model and to demonstrate general model performance. This was followed by the application of the GLUE and Pareto uncertainty techniques using the methods described above. The results from these analyses are presented in this section.

### **3.4.1 Single-Criteria Calibration**

Traditional calibration approaches compare simulated and observed results for a parameter of interest (e.g. streamflow) using a single objective function. In this study, we first investigated model performance by applying to separate single-criteria objective functions. For each investigation, the optimal parameter set was determined by adjusting input parameters until the highest degree of fit could be achieved for both peak flows (evaluated using the NSE OF) and water volume (evaluated using the RE OF). These characteristics were also investigated separately.

As described in Chapter 2, Section 2.2.1 showed the traditional calibration based on the Nash-Sutcliffe Efficiency (NSE) criteria. Here we investigated the same simulation results but we also include results for a second objective function, relative error (RE). The simulations were created based on Monte Carlo Random Search and around  $10^4$  unique parameter sets are selected for each objective function. Figure 3-1 demonstrates the streamflow simulation time-series based on optimal parameter sets using NSE and RE objective functions in traditional calibration methods. As expected the



simulation results based on calibration using the NSE OF and peak flow as the evaluation parameter, performed much better with respect to matching large flow events (Figure 3-1 a). Conversely, the simulation results based on the RE OF using water volume as the evaluation parameter, produced simulation results with a smooth time-series pattern and which better matched low-flow conditions (Figure 3-1 a). From the residual error plots, we can make the same conclusions (Figure 3-1 c). As mentioned in Chapter 2, the model simulation system has some bias with over-predicted streamflow, especially the significant bias based on the objective function RE (Figure 3-1 b).

Insights can be gained with respect to the parameter sets resulting in good model performance (e.g. top 5%) by normalizing and plotting the parameter sets (Figures 3-2 and 3-3). The parameter “equifinality” phenomenon can be visualized by inspecting the parameter distributions and examples of parameter sets for different objective functions. We can see parameter sets spread across the range of feasible parameter space that produce equally good model performance. This represents an alternative approach for visualizing the identifiability analysis discussed in Chapter 2 (Figure 2-8).

Based on the NSE (Figure 3-2) and RE (Figure 3-3) objective functions, the top 5% parameter sets from the Monte Carlo simulations were selected and plotted. The points corresponding to each parameter represent the range of parameter space. The lines represent the parameter sets. By comparing Figures 3-2 and 3-3, it can be seen that the optimal parameter sets are influenced by the selected objection function. However, general characteristics are similar between the two OFs. The dominant parameters  $b$ ,  $W_s$  and  $d_2$ , show relatively high identifiability with different extents of uncertainty. Within the optimal model performance regions,  $b$  is focusing in the range of 0.15-0.5 with some

spread above 0.5 when considering the NSE OF. Based on the RE OF, the lower range of b approaches zero and most points fall below 0.4 to 0.45. No points were observed above 0.625. Ws demonstrates less uncertainty when considering the top performing parameter sets base on the RE OF as compared to the NSE OF. The parameter space also showed slightly less uncertainty for the RE OF compared to the NSE OF for the parameter d2. Insights can be gained by investigating the connections between these observations and the underlying physical mechanisms described by each parameter

### 3.4.2 Multi-Criteria GLUE Analysis

A multi-criteria GLUE analysis was completed by transforming the results into a single criteria framework, which was used to assign weights to each parameter set. The results were investigated using a best-fit criterion rather than the traditional likelihood function as described by Beven (1996) and defined by Franks (1998). The effects of conditioning based on the ranked behavioral parameter response are shown in Figure 3-5. The plots consider the interaction of multiple parameters by representing the likelihood value with respect to each parameter set. Each point represents the likelihood value with respect to each parameter set. The higher likelihood values are weighted as simulations with better fit of the observed streamflow time-series.

The likelihood weighted output from the GLUE technique can be used to generate confidence bounds for simulated streamflows. For example, Figure 3-6 shows the 95% confidence bounds, which encompass most of the observed discharge values. Another example of the utility of this analysis is contained in Table 3-1. The uncertainty of the model's forecast of water volume over the simulation was quantified to as  $(1.89 \pm 0.02) \times 10^{10} m^3$  at the 95% confidence interval.

### 3.4.3 Multi-Criteria Pareto Analysis

The multi-criteria Pareto analysis allows for the interrogation of model performance based on multiple OF and to identify high-performing parameter sets along a Pareto front, which represents a tradeoff between the two OFs. Figure 3-7 shows the resulting OFs from model results generated through Monte Carlo random simulations across the range of feasible parameter space. Although there is a general correlation between the two OFs, the challenge of tradeoffs is more obvious when observing the high-performing points in red (bottom left corner). These points represent top performing simulations with respect to both OFs.

As a consequence, the results from the multi-criteria Pareto analysis can be used to quantify uncertainty using the high-performing parameter sets. Figure 3-8 demonstrates the Pareto output uncertainty as applied to the simulation time series. Rather than a single time series dataset, the Pareto output provides a range of potential streamflow simulations. Further, the total water total water volume in the research period ranges from  $1.40 \times 10^{10}$  to  $2.37 \times 10^{10} \text{m}^3$  (Table 3-1).

## 3.5 Conclusions and Discussions

The goal of this chapter was to propose and demonstrate an integrated, multi-criterion estimation routine based on parameter uncertainty quantification for large scale distributed hydrologic models. This was accomplished using two methods: (1) a multi criterion within a GLUE framework (rather than the traditional likelihood function); and (2) investigation of tradeoffs using a Pareto front framework.

Exploring model performance based on classical single-criteria calibration approaches revealed that the simulated discharge time-series could not reproduce peak streamflow and water volume characteristics. As expected, the NSE OF performed well in simulating peak streamflow but poorly with respect to water volume. Likewise, the model performance was accurate when forecasting water volume when using the RE OF but poorly in describing peak streamflow.

To address the multi-criteria trade-offs issue, the two objective functions were transformed into a single criteria within the GLUE methodology. This technique was compared with a Pareto optimum methodology. These approaches represent different strategies to extract optimal parameter sets based on the concept of equifinality. Within the GLUE method, the optimal sets were formulated by assigning higher likelihood values to parameters with good performance. Optimal sets were observed with the Pareto set by ranking parameter sets base on different levels along the Pareto front. The quantification of uncertainty reveals the tradeoffs between simulating different aspects of the hydrologic system. The GLUE analysis revealed relatively low uncertainty and in the time series, as compared to the Pareto front results. However, it overestimated the water volume compared to the historical observation. In contrast, the uncertainty bands generated from the optimal Pareto sets were overall larger. However, the observed volume of water was within the predicted range of values.

The results of the case study demonstrated the utility of the uncertainty analyses for generating probabilistic confidence intervals. The differences in the extent of uncertainty generated by the two approaches can be related to the underlying physical processes as explored in Chapter 2. Recall, from the identifiability analysis of parameter

d2 (the second layer soil depth) presented in Chapter 2, the parameter was found to be slightly identifiable with respect to the optimal aspects through the whole feasible parameter space. However, it showed a high degree of identifiability across the whole behavioral criteria space. This result introduces an overestimate of streamflow.

Within the GLUE framework, the likelihood weighted to one particular parameter will vary depending on the values assigned to other parameters. As a result, the likelihood is associated with a set of parameters rather than a single parameter. Thus, it is only considered as an acceptable parameter within the series of weights assigned across the set of parameters. Such uncertainties are reflected within the confidence bounds. These impacts potentially contribute to the overestimate of streamflow yet narrow uncertainty bounds with the GLUE results as compared to the Pareto outcomes.

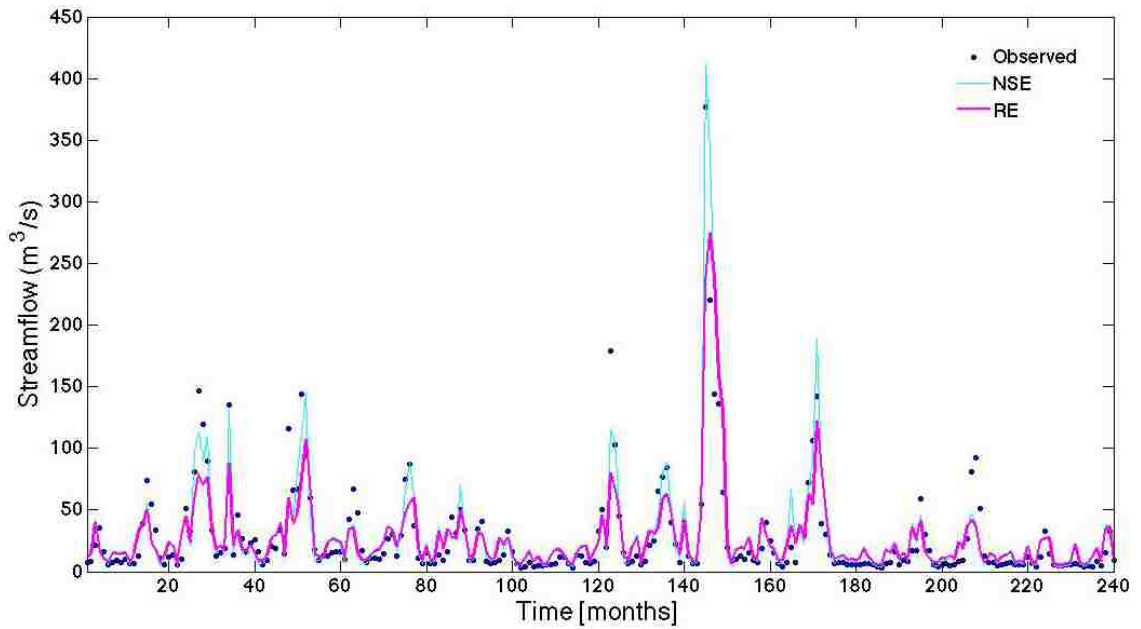
The issues of equifinality with parameter interactions are an important source of uncertainty for hydrologic simulations. An appropriate parameterization with higher identifiability would help reduce the degree of uncertainties.

Overall, the results of this study revealed reliable and reasonable probabilistic hydrologic predictions by quantifying parameter uncertainty. Concerning the practicality to water management, the model results showed an overall overestimate of water yields beyond the specified confidence levels.

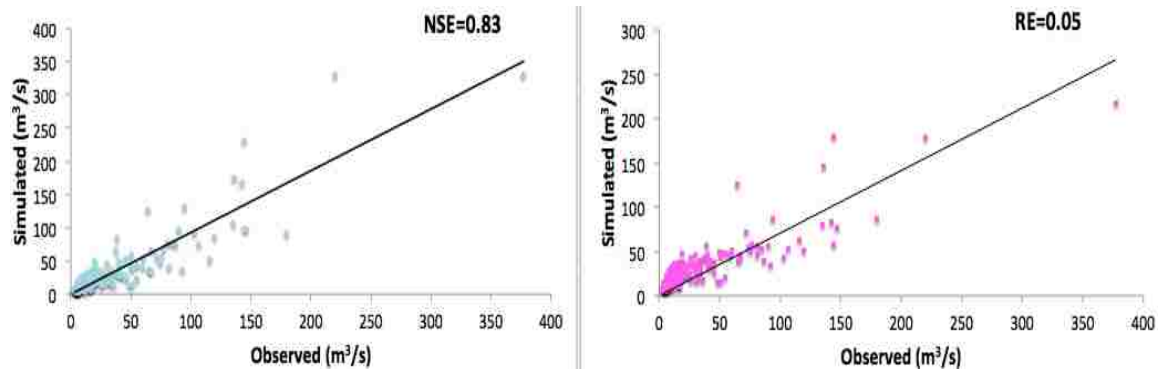
**Table 3–1 List probabilistic analysis with observed discharge, GLUE uncertainty outputs and Pareto uncertainty outputs**

<b>Adopted Approaches</b>	<b>Standard Deviation</b>	<b>Total Volume (m<sup>3</sup>)</b>	<b>Uncertainty Extent (m<sup>3</sup>)</b>
Observed		$1.75 \times 10^{10}$	
Pareto uncertainty	$2.82 \times 10^9$	$1.95 \times 10^{10}$	$(1.4 \sim 2.37) \times 10^{10}$
GLUE uncertainty	$2.12 \times 10^9$	$1.89 \times 10^{10}$	$(1.89 \pm 0.02) \times 10^{10}$

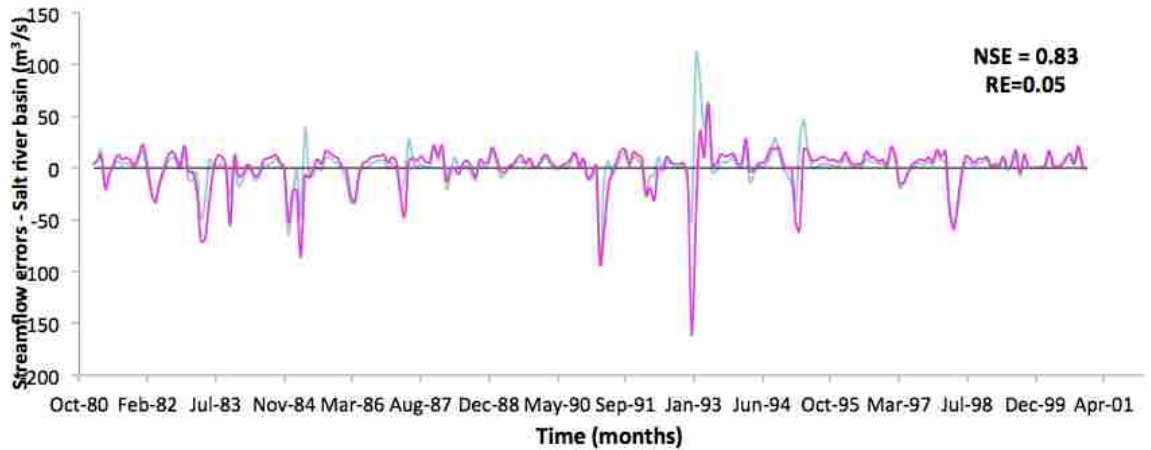
(a)



(b)

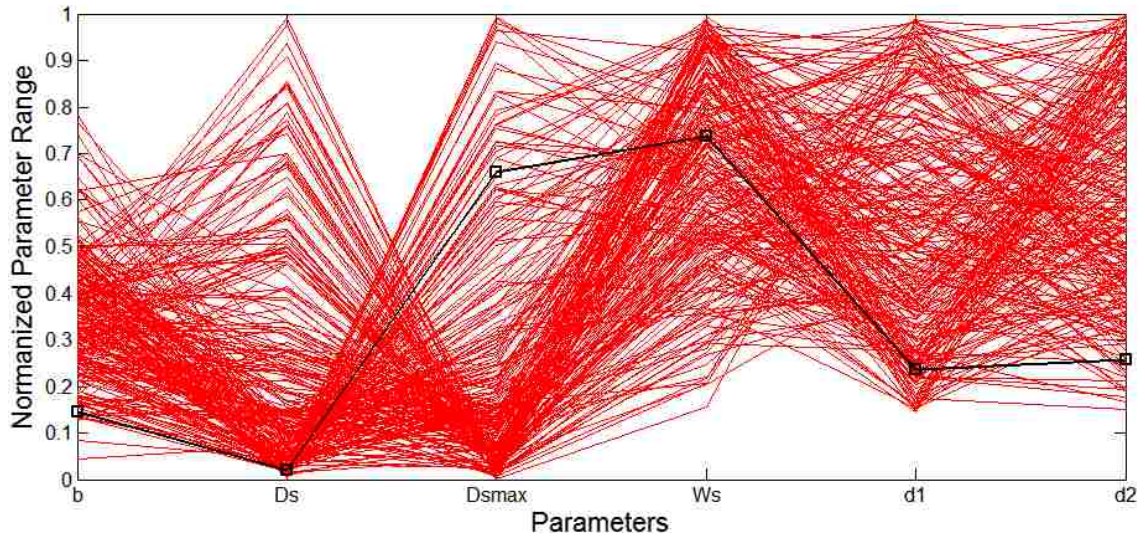


(c)

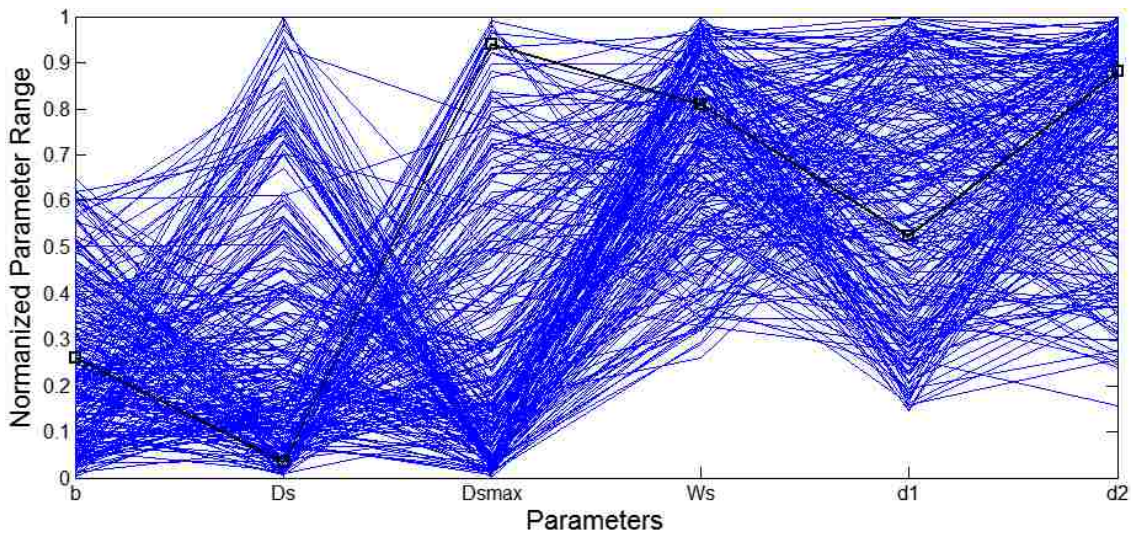


**Figure 3–1 Traditional single-criteria streamflow simulations with respect to different objective functions. (a) The black points represent observed monthly discharge under study period. The blue and pink lines represent the best simulated streamflow compared to observed time series based on NSE and RE, respectively. (b) Modeled vs. observed streamflow for both objective functions. (c) Residual errors plot.**





**Figure 3–2 Normalized parameter sets space selected by optimal value due to NSE performance. Among them, the top 5% optimal clusters are drawn (each red line represents a parameter set; the black line represents the optimal parameter set based on NSE).**



**Figure 3–3 Normalized parameter sets space selected by optimal value due to RE performance. Among them, the top 5% optimal clusters are drawn (each blue line represents a parameter set; the black line represents the optimal parameter set based on RE). Note: Each line represents one parameter set which could be given a measurement by model performance it will be meaningless with single point.**

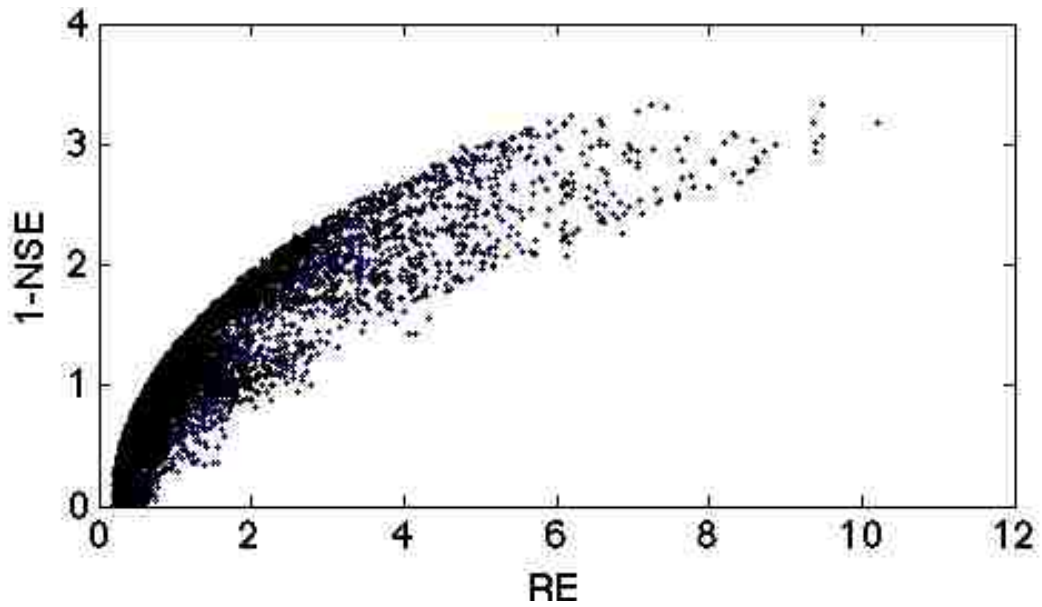


Figure 3-4 A hypothetical example trade-off between NSE and RE is shown. This trade-off is also referred to as the Pareto optimality. The performance drop into the lower left corner could satisfy both of the objective functions.

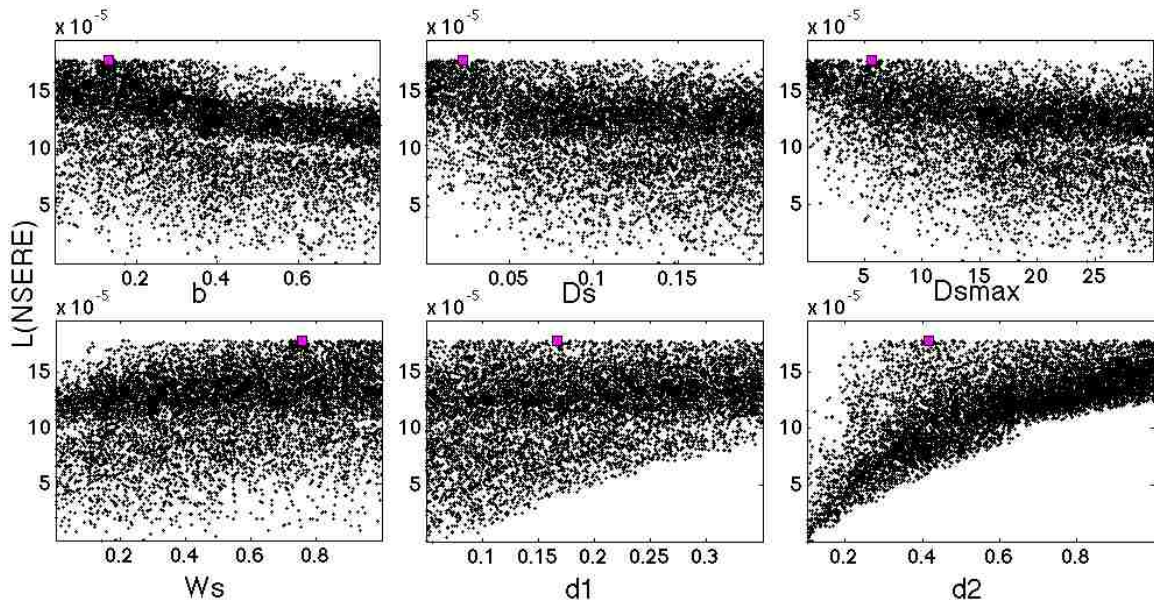
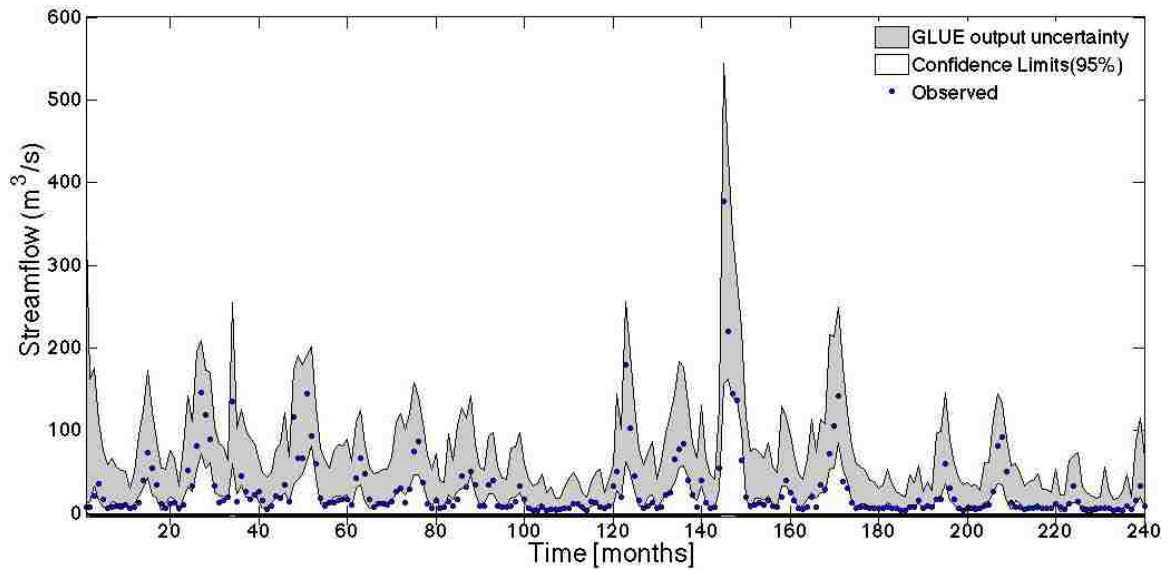
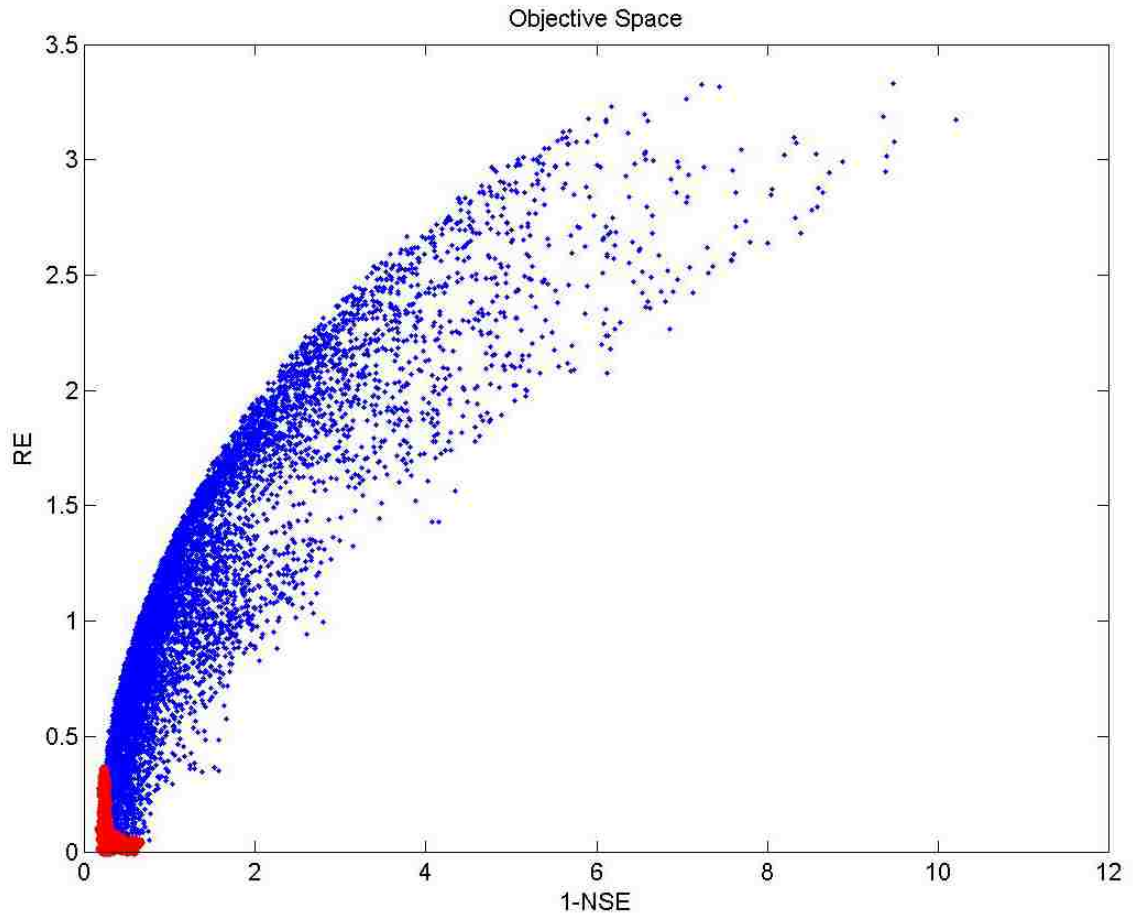


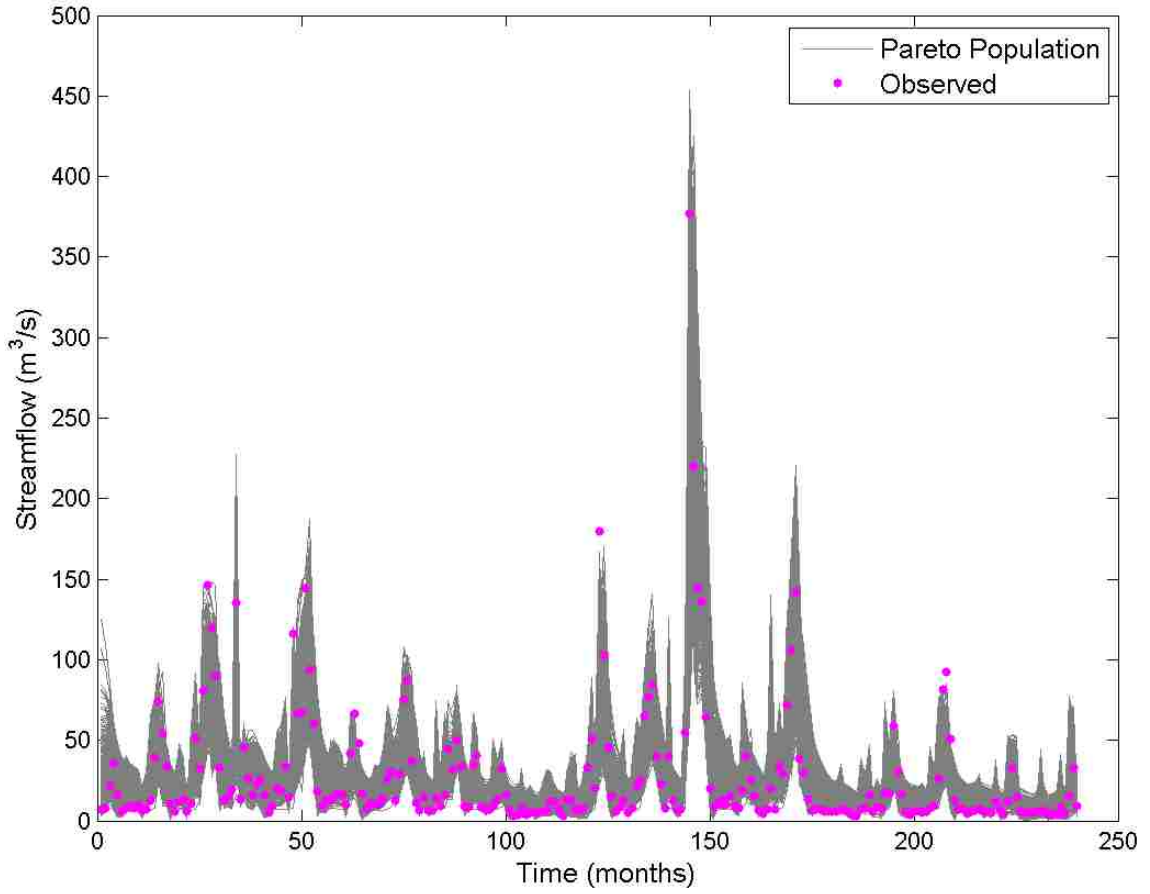
Figure 3-5 GLUE likelihood measures with multi-criteria NSE&RE tradeoffs for Salt River basin. The application of GLUE in this research is subjectively selected the likelihood function as terminal criteria (defined in equation 3-4 and 3-5). Scatter plots also represent model performance with higher likelihood values by better simulations. All of those likelihood values spread through the parameter space should be added up to one.



**Figure 3–6 GLUE output associated confidence limits. Hydrograph of 95% percentile confidence prediction bounds estimated by GLUE with multi-criteria tradeoffs. The blue point indicates observed streamflow and the grey shaded area represents the prediction uncertainty resulted from GLUE estimation. The black line demonstrates the uncertainty bounds with specific (95%) confidence level.**



**Figure 3–7 Pareto front selected based on trade-offs between NSE and RE. The blue points represent model performances from all random simulations with respect to the adopted two objective functions. The red points are the optimal sets filtered by pareto ranking algorithm. The cluster simulations related to these red points are remarked as pareto uncertainty bounds.**



**Figure 3–8 Pareto output uncertainty. The pink points are observed monthly discharge time series. The grey areas are extracted from pareto optimal space which represents the red points simulation cluster as shown in the last figure.**

## Chapter 4 Conclusions and Future Works

### 4.1 Conclusions

With the development of highly complex hydrologic models using a wide range of parameterization methods to representing hydrologic processes, several challenges arise with respect to parameter estimation and quantification of model uncertainty. A novel approach of estimation-uncertainty investigations was proposed and demonstrated. The procedure also considers temporal variability of parameters through time-series. The issues of identifiability and overparameterization with respect to parameter uncertainty and qualitative information were explored. The techniques used in this research are programmed within MATLAB using the Linux platform and the VIC model. These scripts will be packaged and available to the community (similar to SWAT-CUP designed to integrate calibration/uncertainty analysis) (Abbaspour et al., 2007; Singh et al., 2013).

#### 4.1.1 Building a Framework to Effectively and Efficiently Demonstrate Estimation Scheme

To achieve an efficient and effective simulation with respect to specified hydrological aspects, it is necessary to recognize the required degree of complexity by the

application of complex distributed hydrologic model with proper parameterization. The characteristics of the framework are presented in Figure 4-1. The tasks of the new framework are to recognize parameter sensitivity and identifiability within a spatially distributed and temporally varying procedure for large-scale hydrologic models. This framework was demonstrated through application of the VIC model within the Gila River basin. The strategy includes the projection of parameter space through a population-searching algorithm such as the Monte Carlo Uniform Random Sampling (MCURS) approach based on climatic conditions; a regional sensitivity and identifiability analysis (RSA) by categorizing free parameters as inactive, lumped, and distributed properties; implementation of a dynamic approach (DYNIA) to identify parameters' transient heterogeneity varying through time-series. The overall approach allows the modeler to discriminate overparameterization conditions by recognizing them as inactive, lumped, distributed, improper representation of model aspects, to reduce parameter dimensionality and finally to address an efficient and effective simulation of hydrological processes.

The framework takes into consideration both basin characteristics (climatological condition) and temporal variations while coupling with multi-parameter interaction and correlation. It represents an enhanced evaluation of parameters with basin-wide variability, cross-subbasin variability, and temporally based dynamic variability. The approach also provides a flexible approach for improving hydrologic estimation effectively and efficiently. The case results reveal implications for hydrologic aspects process studies for future uncertainty research (Chapter 3).

#### **4.1.2 Quantifying Hydrological Responses through Integrated Multi-Criteria Simulation**

The improved effectiveness and efficiency of model simulation (Chapter 2) through adjusting to a moderate level of complexity and parameterization still comes with parameter uncertainty. By this recognition, it becomes clear that the classical optimization approach as deterministic single value optimal forecast is flawed (Beven et al., 2012). Instead, the improved estimation procedure presented in this research with respect to the equifinality phenomena results in an improved representation of variability. The development and application of a novel framework with qualitative information, allowed us to investigate two important questions: How can we discriminate between behavioral parameter sets as optimal conditions from random parameter population? Also, how can we reveal the extent of estimation by aggregating those optimums (GLUE & Pareto Optimality)?

A framework measurement of trade-offs was explored within a multi-criterion estimation routine using different objective functions (Nash-Sutcliffe Efficiency and Relative Error) based on simulated peak streamflows and water volumes. The quantification was investigated by GLUE with subjective selection of likelihood with multi-criteria function and Pareto optimality approach. The results were shown in specific confidence percentile and Pareto space. The parameter uncertainty was shown as a range of possible time-series simulations with upper and lower limits (meaningful representation of uncertainty bounds) with most probable indicators, such as average, maximum & minimum, confidence interval and variance. Within the improved context, the single estimation based on global optimal simulations or the combination with typical



parameter forecasts with likelihood weighting on each time step could be highlight. The subjective output as a user-preferred solution can be extracted as a deterministic format, which is reliable and more understandable within a decision-making process as applied in water resources engineering.

#### **4.1.3 Integration of a Post-Processing Tool for Estimation/Uncertainty Analysis**

The practical output for this research is a series of MATLAB scripts produced within the Linux platform and applied to the VIC model. These scripts include optimization methods, uncertainty analysis techniques, data formatting pre-process (grid, binary, ASCII, etc), output probabilistic post-process and script connectors between each other with various platforms by coupling different computer languages (C++, Fortran, Matlab, Linux operation). These scripts will be bundled into a package and made available to the water resources community in a format similar to SWAT-CUP. SWAT-CUP was designed to integrate calibration/uncertainty analysis as a post-process tool for the Soil Water Assessment Tool (SWAT) (Abbaspour et al., 2007). The modeling package will be a helpful and user-friendly tool to support large-scale hydrologic modeling.

## **4.2 Recommendations for Future Research**

With the recognition of inherent uncertainty properties in hydrologic sciences, the enhanced pareto estimation framework should be further explored beyond the usual classical optimal “best” deterministic scheme through coupling with qualitative information. The current progress of this research mainly focuses on parameter uncertainty analysis. The factors contributing to uncertainty of hydrologic behavior

simulations (Hall et al., 2008) include incomplete knowledge of the system, variability in system properties, randomness in systems, measurement and sampling errors, and actual scales of the system (Georgakakos et al., 2004; Meyer et al., 2007). It is important and necessary to comprehensively and quantitatively describe the influence of uncertainty on the entire simulation process (Ajami et al., 2007).

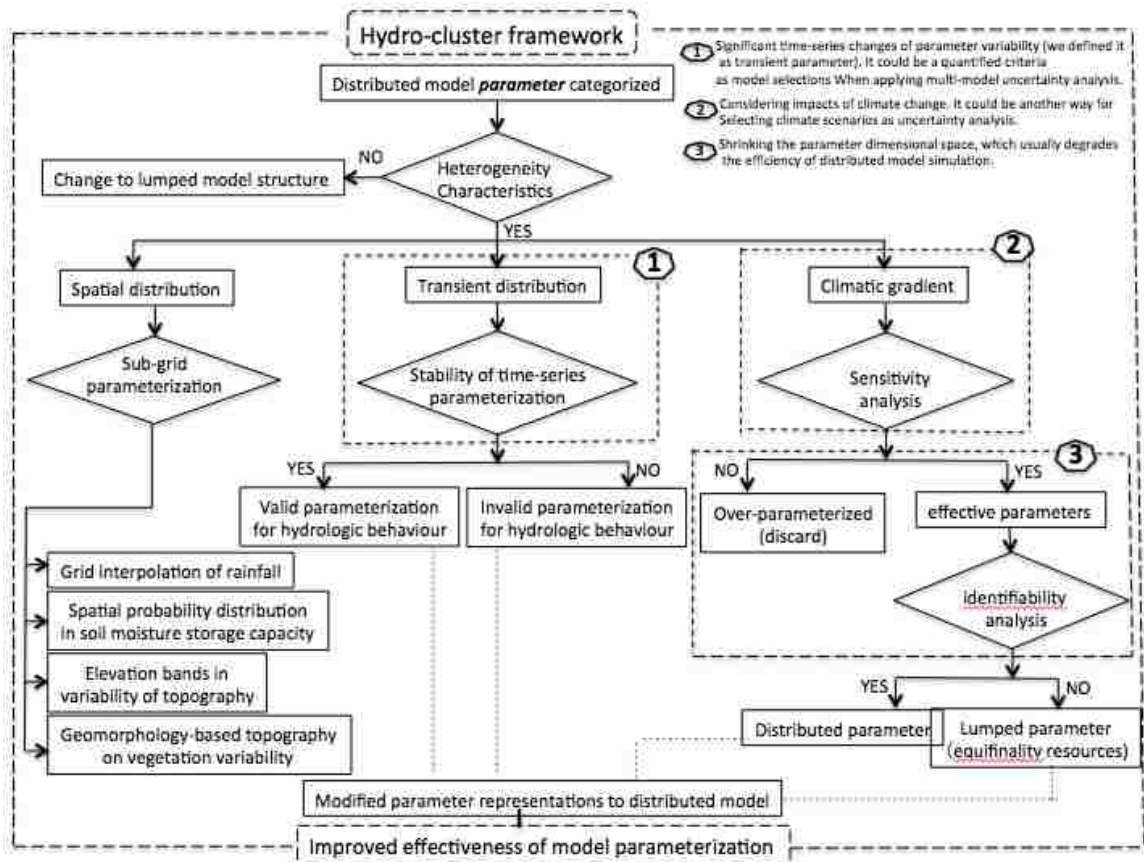
The various sources of uncertainty can be divided into three categories: (1) parameter uncertainty; (2) concept model uncertainty; and (3) scenario uncertainty. Parameter uncertainty includes the unknown distribution of model parameters. Concept model uncertainty is due to assumptions in the underlying physical description in the model. Scenario uncertainty is due to unknowns regarding future conditions. The long-term aim of the framework proposed in this study is to demonstrate a systematic uncertainty analysis that can be applied to climate change impact studies (Hamlet and Lettenmaier, 1999; Engeland et al., 2005; Hamlet et al., 2010) in the future by the following objectives: (1) Develop and demonstrate a framework for evaluating uncertainty in a study to improve management of water resources; (2) Quantify uncertainty (including parameter, model, and scenario based uncertainties in a climate change assessment for water resources; and (3) Combine and evaluate the statistical confidence associated with each uncertainty component to overall impacts and to assess where uncertainties can be reduced through future research.

Through my dissertation research, I have conducted a systematic evaluation of an effective and efficient framework with a large scale distributed hydrologic model by parameter estimation. Additionally, the building of the current framework advanced understanding of the influences of scales, climatic conditions, heterogeneity (spatial and

temporal with dynamic), which will help to address the future model and scenario components.

The scripting products developed through this study are ideally suited for investigating uncertainty associated with model structure and climate change scenario uncertainty. The resulting tool can comprehensively and reliably address these issues to improve decision-making.

The implementation of such a framework will facilitate an improved approach for addressing climate change impacts on natural systems. Further, a study of this nature will advance understanding of how climate change is likely to impact water resources including water supplies. Finally, the approach can help modelers and decision makers better understand the relative contributions to uncertainty in order to inform future investments to reduce uncertainties.



**Figure 4–1 Characteristics and flow chart for the framework of estimation-quantification uncertainty to large scale distributed model building in this research.**

## Chapter 5 References

- [1]. Ajami, N.K., Duan, Q. and Sorooshian, S., 2007. An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structure uncertainty in hydrologic prediction. *Water Resour. Res.* 43: W01403, 1-19
- [2]. Ajami, N.K., Gupta, H.V., Wagener, T. and Sorooshian, S. 2004. Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. *Journal of Hydrology*, 298(1-4), 112-135.
- [3]. Abbaspour, K.C., Johnson, A., Genuchten, M.Th. van, 2004. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone Journal*, 3: 1340-1352
- [4]. Abdulla, F., and D. Lettenmaier 1997. Development of regional parameter estimation equations for a macroscale hydrologic model. *J. Hydrol.*, 197, 230-257
- [5]. Andreadis, K.M. and Lettenmaier, D.P., 2006. Trends in 20<sup>th</sup> century drought over the continental United States. *Geophysical Research Letters*, 33 (10): L10403, 10.1029/2006GL025711
- [6]. Abdulla, F.A., Lettenmaier, D.P., Wood, E.F. and Smith, J.A., 1996. Application of a macroscale hydrologic model to estimate the water balance of the Arkansas-Red river basin. *Journal of Geophysical Research*, 101 (D3), pp. 7449-7459

- [7]. Ang, A.H.S. and Tang, W.H., 1975. Probability concepts in engineering planning and design: Volume 1-basic principles. John Wiley & Sons, New York, NY, US
- [8]. Arnold, J.G., Srinivasan, R., Muttiah, R.S., 1998. Large area hydrologic modeling and assessment – Part 1: Model development. Journal of the American Water Resources Association, 34 (1): 73-89
- [9]. Atkinson, S.E., Woods, R.A. and Sivapalan, M., 2002. “Climate and landscape controls on water balance model complexity over changing time-scales”. Water Resources Research, Vol. 38 (12), 1314, doi: 10.1029/2002WR001487
- [10]. Abbaspour, K.C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J. and Srinivasan R., 2007. Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. Journal of Hydrology, 333: 413-430
- [11]. Ajami, N.K., Q. Duan, and S. Sorooshian. 2007. An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structure uncertainty in hydrologic prediction. Water Resour. Res. 43: W01403, 1-19, doi: 10.1029/2005WR004745
- [12]. Bastidas, L.A., Gupta, H.V., Sorooshian, S., Shuttleworth, W.J. and Yang, Z.L., (1999). Sensitivity analysis of a land surface scheme using multicriteria methods. Journal of Geophysical Research, Vol. 104, No. D 16, P19481-19490
- [13]. Bilmes J A., 1998. A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models.

Department of Electrical Engineering and Computer Science U.C. Berkeley,  
Technial Report, pp. 1-13

- [14]. Beven, K.J. and Krikby, M., 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Science Bulletin*, 24(1), 43-69
- [15]. Beven K.J., 1989. Changing ideas in hydrology-the case of physically-based models. *Journal of hydrology*, Vol.105, issue 1-2, P157-172
- [16]. Beven, K.J. and Binley, A.M., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Processes*, 6 (3): 279-298
- [17]. Beven, K.J. and Moore, I.D., 1992. Terrain analysis and distributed modeling in hydrology. John Wiley & Sons, 1992: 249
- [18]. Beven, K.J., (1998). Generalised Likelihood Uncertainty Estimation (GLUE) package v. 1.0 – User manual. University of Lancaster, UK
- [19]. Beven, K.J., 1993. Prophecy, reality and uncertainty in distributed hydrological modeling. *Adv. Water Resour.*, 16: 41-51
- [20]. Beven K. J. 1996. A discussion of distributed hydrological modeling. *Water science and technology library*, Vol. 22, 255-278
- [21]. Beven, K. J., 2000. Uniqueness of place and process representations in hydrological modeling. *Hydrology and Earth System Sciences*, vol.4 (2),203-213

- [22]. Beven, K., 2001. How far can we go in distributed hydrological modeling? *Hydrology and Earth System Sciences*, 5(1), 1-12
- [23]. Beven, K.J., 2000. Uniqueness of place and process representations in hydrological modeling. *Hydrol. Earth Syst. Sci.*, 4, 203-214
- [24]. Beven, K.J., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* 249, 11-29
- [25]. Beven, K.J., 2006. A manifesto for the equifinality thesis, *Journal of Hydrology*, 320 (1), 18-36
- [26]. Beven, K.J., 2006. Linking parameters across scales: Subrid parameterizations and scale dependent hydrological models. *Hydrological Processes*, vol. 9 (5-6), doi: 10.1002/hyp.3360090504
- [27]. Beven, K.J., Smith, P.J. and Freer, J.E., 2007. Comment on “Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology” by Pietro Mantovan and Ezio Todini. *J. Hydrol.*, 338: 315-318
- [28]. Beven, K., P. Smith, and A. wood (2011). On the colour and spin of epistemic error (and what we might do about it). *Hydrol. Earth Syst. Sci. Discuss.*, 8, 5355-5368
- [29]. Beven, K.J. and Alcock, R.E., 2012. Modelling everything everywhere: a new approach to decision-making for water management under uncertainty. *Freshwater Biology*, 57 (Suppl. 1), 124-132, doi: 10.1111/j.1365-2427.2011.02592.x



- [30]. Beven, K. J., 2012. Causal models as multiple working hypotheses about environmental processes. *Comptes Rendus Geoscience*, 344, 77-88
- [31]. Binley, A.M. and Beven, K.J. 1991. “Physically-based modeling of catchment hydrology: a likelihood approach to reducing predictive uncertainty”, in Farmer, D.G. and Rycroft, M.J. (Eds). *Computer modeling in the Environmental sciences*, Clarendon Press, Oxford, 75-88
- [32]. Bell, V.A. and Moore, R.J., 2000. The sensitivity of catchment runoff models to rainfall data at different spatial scale. *Hydrology and Earth System Sciences*, 4 (4), 653-667
- [33]. Blasone, R.S., Madsen, H., Rosbjerg, D., (2008). Uncertainty assessment of integrated distributed hydrological models using GLUE with Markov chain Monte Carlo sampling. *Journal of hydrology* 353, 18-32
- [34]. Blasone, R., Vrugt, J., Madsen, H., Rosbjerg, D., Robinson, B. and Zyvoloski, G., 2008. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. *Advances in Water Resources*, 31, 630-648
- [35]. Blondin, C., 1991. Parameterization of land-surface processes in numerical weather prediction, in *Land Surface Evaporation: Measurements and Parameterization*, edited by Schmugge, T.J. and Andre, J.C., pp. 31-54, Springer-Verlag, New York
- [36]. Boyle, D., Gupta, H., and Sorooshian, S., (2000). Toward improved calibration of hydrologic models: Combing the strengths of manual and automatic methods. *Water Resour. Res.*, 36, 3663-3674

- [37]. Boyle, D.P., Gupta, H.V., and Sorooshian, S., 2000. Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resour. Res.*, 34, 3663-3674
- [38]. Boyle D., Gupta, H., Sorooshian S., Koren V., Zhang Z., and Smith M., (2001). Towards improved streamflow forecasts: The value of semi-distributed modeling. *Water Resour. Res.*, 37 (11), 2739-2759
- [39]. Blöschl, G., 2006. Hydrologic synthesis: Across processes, places, and scales. *Water Resour. Res.*, 42, W03S02, doi: 10.1029/2005WR004319
- [40]. Cui, Z., Vieux, B.E., Neeman, H. and Moreda, F., 2005. Parallelisation of a distributed hydrologic model. *International Journal of Computer Applications in Technology*, 22: 42-52
- [41]. Christensen, N.S., Wood, A.W., Voisin, N., Lettenmaier, D.P. and Palmer, R.N., 2004. Effects of climate change on the hydrology and water resources of the Colorado River basin. *Climatic Change*, 62 (1-3), 337-363
- [42]. Christensen, N.S. and Lettenmaier, D.P., 2006. A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River Basin. *Hydrol. Earth Syst. Sci.*, 11: 1417-1434
- [43]. Cibin, R., Sudheer, K.P. and Chaubey, I., 2010. Sensitivity and identifiability of stream flow generation parameters of the SWAT model. *Hydrol. Process.*, 24: 1133-1148, doi: 10.1002/hyp.7568

- [44]. Das, I., 1999. A preference ordering among various Pareto optimal alternatives. *Structural Optimization*, 18 (1), 30-35
- [45]. Duband, D., et al., 1993. Unit-hydrograph revised – an alternate iterative approach to UH and effective precipitation identification. *Journal of Hydrology*, 150 (1), 115-149
- [46]. Demaria, E.M., Nijssen, B. and Wagener, T., 2007. Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model. *Journal of Geophysical Research*, vol. 112, D11113, doi: 10.1029/2006JD007534
- [47]. Decharme, B. and Douville, H., 2005. Introduction of a sub-grid hydrology in the ISBA land surface model. *Climate Dynamics*, doi: 10.1007/s00382-005-0059-7
- [48]. Duan, Q., Sorooshian, S., Gupta, H.V., Rousseau, A.N. and Turcotte, R., 2003. *Advances in calibration of watershed models*. AGU, Washington, DC
- [49]. Gupta, H.V., Bastidas, L.A., Vrugt, J.A. and Sorooshian, S., 2013. Multiple Criteria Global Optimization for Watershed Model Calibration, in *Calibration of Watershed Models* (eds Duan, Q., Gupta, H., Sorooshian, S., Rousseau, A., and Turcotte, R.). American Geophysical Union, Washington, D.C., doi: 10.1029/WS006p0125
- [50]. Efstratiadis, A., and D. Koutsoyiannis, (2010). One decade of multi-objective calibration approaches in hydrological modeling: A review. *Hydrol. Sci. J.*, 55(1), 58-78
- [51]. Elsner, M.M. and Hamlet, A., 2010. Variable Infiltration Capacity model implementation. Climate Impact Group (CIG) report

- [52]. Engeland, K., Xu, C.Y. and Gottschalk, L., 2005. Assessing uncertainties in a conceptual water balance model using Bayesian methodology. *Hydrological Sciences Journal*, 50 (1), 45-63
- [53]. Euser, T., Winsemius, H., Hrachowitz, M. and Savenije, H., 2014. Effect of spatial forcing data and landscape heterogeneity on performance and consistency of model structures. *Geophysical Research Abstracts*, vol. 16, EGU2014-5025
- [54]. Franks S.W., Gineste P., Beven K.J. and Merot P. (1998). On constraining the predictions of a distributed model: the incorporation of fuzzy estimates of saturated areas into the calibration process. *Water Resources Research*, vol 34 (4), 787-797
- [55]. Franks S.W., Beven K. J. and Gash J.H.C. (1999). Multi-objective conditioning of a simple SVAT model. *Hydrology and Earth System Sciences*, 3(4), 477-489
- [56]. Freer, J., Beven, K.J. and Ambroise, B., (1996). Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resour. Res.*, 32(7), 2161-2173
- [57]. Feyen, L., Kalas, M. and Vrugt, J.A., 2008. Semi-distributed parameter optimization and uncertainty assessment for large-scale streamflow simulation using global optimization. *Hydrological Science*, 53 (2)
- [58]. Franchini, M. and Pacciani, M., 1991. Comparative-analysis of several conceptual rainfall runoff models. *Journal of Hydrology*, 122 (1-4), 161-219
- [59]. Gassman, P.W., Reyes, P.W., Green, M.R. and Arnold, J.G., 2007. The soil and water assessment tool: Historical development, applications, and future research

directions. American Society of Agricultural and Biological Engineers, 50 (4): 1211-1250

- [60]. Gelhar, L.W., 1986. Stochastic subsurface hydrology from theory to applications. Water Resour. Research, vol. 22 (9): 135s-145s
- [61]. Grayson, R. and Blöschl, G., 2001. Spatial pattern in catchment hydrology: observations and modeling. The university of Cambridge press, UK.
- [62]. Gharari S., Hrachowitz, M., Fenicia, F. and Savenije, H.H.G., 2013. An approach to identify time consistent model parameters: sub-period calibration. Hydrol. Earth Syst. Sci., 17, 149-161
- [63]. Georgakakos, K.P., Seo, D.J., Gupta, h., Schaake, J. and Butts, M.B., 2004. Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. Journal of Hydrology, 298: 222-241
- [64]. Gupta, H.V., Sorooshian, S. & Yapo, P.O. (1998). Toward improved calibration of hydrologic models: multiple and non-commensurable measures of information. Water Resour. Res. 34 (4), 751-763
- [65]. Harr, M.E., 1989. Probabilistic estimates for multivariate analyses. Applied Mathematical Modelling, 13 (5): 313-318
- [66]. Hamill, T.M., Bates, G.T., Whitaker, J.S., Murray, D.R., Fiorino, M., Galarneau, T.J., Zhu, Y. and Lapenta, W., 2013. NOAA's Second-Generation Global Medium-Range Ensemble Reforecast Dataset. Bull. Amer. Meteor. Soc., 94: 1553-1565, doi: <http://dx.doi.org/10.1175/BAMS-D-12-00014.1>

- [67]. Hamlet, A.F., Lettenmaier, D.P., 1999. Effects of climate change on hydrology and water resources of the Columbia River Basin, *J. Am Water Resour. Assoc.*, 35:1597-1624
- [68]. Hamlet, A.F., Lee, S.Y., Mickelson, K.B., Elsner, M.M., 2010. Effects of projected climate change on energy supply and demand in the Pacific Northwest and Washington State, *Climatic Change*, doi: 10.1007/s10584-010-9857-y
- [69]. Hundecha, Y. and Bardossy, A., 2004. Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. *J. of Hydro.*, 292: 281-295
- [70]. Hall, J. and Solomatine, D., 2008. A framework for uncertainty analysis in flood risk management decisions. *Intl. J. River Basin Management*, 6 (2), 85-98
- [71]. Kavetski, D., Kuczera, G. and Franks, S.W., 2006. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. *Water Resources Research*, 42, W03407, doi: 10.1029/2005WR004368
- [72]. Koster, R.D., Suarez, M.J., Ducharne, A., 2000. A catchment-based approach to modeling land surface processes in a general circulation model 1. Model structure. *Journal of Geophysical Research-Atmospheres*, 105 (d20): 24809-24822
- [73]. Kochendorfer, J.P. and Ramirez, J.A., 2005. The impact of land-atmosphere interactions on the temporal variability of soil moisture at the regional scale. *Journal of Hydrometeorology*, 6 (1): 53-67

- [74]. Kollat, J.B., Reed, P.M. and Wagener, T., (2012). When are multiobjective calibration trade-offs in hydrologic models meaningful?. *Water Resour. Res.*, 48, W03520, doi:10.1029/2011WR011534
- [75]. Liang, X., Wood, E.F., Lettenmaier, D.P., 1996. Surface soil moisture parameterization of the VIC-2L model: Evaluation and modifications. *Global Planet Change*, 13: 195-206
- [76]. Liang, X., Lettenmaier, D.P., Wood, E.F. and Burges, S.J., 1994. "A simple hydrologically based model of land surface water and energy fluxes for General Circulation Models". *Journal of Geophysical Research*, Vol.99, pp. 14415-14428
- [77]. Liang, X., Xie, Z.H. and Huang, M.Y., 2003. A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model. *Journal of Geophysical Research-Atmospheres*, 108 (D16): Art. No. 8613
- [78]. Liang, X., Guo, J.Z. and Leung, L.R., 2004. Assessment of the effects of spatial resolutions on daily water flux simulations. *Journal of Hydrology*, 298 (1-4): 287-310
- [79]. Lohmann, D., Nolte-Holube, R. and Raschke, E., 1996. A large scale horizontal routing model to be coupled to land surface parameterization schemes. *Tellus* 48 A, pp. 708-712
- [80]. Lovett, G.M., Jones, C.G., Turner, M.G., and Weathers, K.C., 2006. Ecosystem function in heterogeneous landscapes. Springer Science & Business Media, Inc.

- [81]. Meins, F.M., 2013. Evaluation of spatial scale alternatives for hydrological modeling of the Lake Naivasha basin, Kenya. Thesis, University of Twente
- [82]. Maurer, E.P., Wood, A.W., Adam, J.C., Lettenmaier, D.P. and Nijssen, B., 2002. A long-term hydrologically-based data set of land surface fluxes and states for the conterminous United States. *J. Climate* 15, 3237-3251
- [83]. Mattikalli, N.M., Devereus, B.J, and Richards, K.S., 1996. Prediction of river discharge and surface water quality using an integrated geographical information system approach. *International Journal of Remote Sensing*, 17 (4): 683-701
- [84]. Matheussen, B., Kirschbaum, R.L., Goodman, I.A., O'Donnell, G.M. and Lettenmaier, D.P., 2000. Effects of land cover change on streamflow in the interior Columbia Basin. *Hydrol. Proc.*, 14 (5): 867-885
- [85]. McCloskey, G.L., Ellis, R.J., Waters, D.K. and Stewart, J., 2011. PEST hydrology calibration process for source catchments – applied to the Great Barrier Reef, Queensland. 19<sup>th</sup> International Congress on Modelling and Simulation, Perth, Australia
- [86]. McDonnell, J.J., Sivapalan, M., Vache, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M.L., Selker, J., and Weiler, M., 2007. Moving beyond heterogeneity and process complexity: a new vision for watershed hydrology. *Water Resources Research*, vol. 43, W07301, doi: 10.1029/2006WR 005467
- [87]. McDonnell, J.J. and Woods, R.A., 2004. On the need for catchment classification. *J. Hydrol.*, 299, 2-3



- [88]. Meyer, P.D., Ye, M., Rockhold, M.L., Nueman, S.P. and Cantrell, K.J., 2007. Combined estimation of hydrogeologic conceptual model, parameter, and scenario uncertainty with application to Uranium transport at the Hanford site 300 Area. NUREG/CR-6940, prepared for US Nuclear Regulatory commission, Washington, DC
- [89]. Miller, S.N., Semmens, D.J., Goodrich, D.C., Hernandez, M., Miller, R.C., Kepner, W.G. and Guertin, D.P., 2007. The automated geospatial watershed assessment tool. *Environmental Modelling & Software*, 22: 365-377
- [90]. Mo, K.C., Paegle, J.N. and Higgins, R.W., 1997. Atmospheric processes associated with summer floods and droughts in the central United States. *Journal of Climate*, 10 (12): 3028-3046
- [91]. Muleta, M.K., Nicklow, J.W. and Bekele, E.G., 2007. Sensitivity of a distributed watershed simulation model to spatial scale. *J. Hydrol. Eng.*, 12 (2), 1630172
- [92]. Nash, J. E. and Sutcliffe, J. V., 1970. River flow forecasting through conceptual models. *Journal of Hydrology*, 10:282-290.
- [93]. Nijssen, B., Lettenmaier, D.P., Liang, X., Wetzel, S.W. and Wood, E.F., 1997. Streamflow simulation for continental-scale river basins. *Water Resour. Res.*, 33: 711-724
- [94]. Nijssen, B.N., O'Donnell, G.M., Lettenmaier, D.P. and Wood, E.F., 2001a. Predicting the discharge of global rivers, *Journal of Climate*, 14: 3307-3323
- [95]. Nijssen, B.N., O'Donnell, G.M. and Hamlet, A.F, 2001b. Hydrologic sensitivity of global rivers to climate change. *Climatic Change*, 50 (1-2): 143-175

- [96]. Nijssen, B.N., Schnur, R. and Lettenmaier, D.P., 2001c. Global retrospective estimation of soil moisture using the variable infiltration capacity land surface model, 1980-93, *Journal of Climate*, 14 (8): 1790-1808
- [97]. O'Connell, A.M., 1990. Microbial decomposition (respiration) of litter in eucalypt forest of south-western Australia: An empirical model based on laboratory incubations. *Soil Biology and Biochemistry*, 22, 153-160
- [98]. O'Connell, A.M., 1998. The overvaluation of purchasing power parity. *Journal of International Economics*, 44: 1-19
- [99]. Panday, S. and Huyakorn, P.S., 2004. A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. *Advances in Water Resources*, 27: 361-382
- [100]. Pareto, V., 1896. *The new theories of Economics*. Journal of Political Economy, The University of Chicago Press, 485-502
- [101]. Pokhrel, P., Gupta, H.V. and Wagener, T., 2008. A spatial regularization approach to parameter estimation for a distributed watershed model. *Water Resources Research*, vol. 44, W12419, doi: 10.1029/2007WR006615
- [102]. Payne, J.T., Wood, A.W., Hamlet, A.F., Palmer, R.N. and Lettenmaier, D.P., 2004. Mitigating the effects of climate change on the water resources of the Columbia River basin, *Climate Change*, 62: 233-256
- [103]. Ray, N., Giuliani, G., Gorgan, D. and Lehmann, A., 2012. Distributed geocomputation for modeling the hydrology of the black sea watershed.

Environmental Security in Watersheds: The sea of Azov NATO Science for Peace and Security Series C: Environmental Security, pp. 141-157

- [104]. Refsgaard, J.C. and Knudsen, J., 1996. Operational validation and intercomparison of different types of hydrological models. *Water Resources Research*, 32 (7): 2189-2202
- [105]. Refsgaard, J.C. and Storm, B., 1996. Construction calibration and validation of hydrological models. In: Abbott, M.B., Refsgaard, J.C. (Eds.), *Distributed Hydrological Modelling*, Water Science and Technology Library, vol. 22, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 41-54
- [106]. Refsgaard, J.C., 1997. Parameterisation, calibration and validation of distributed hydrological models. *Journal of Hydrology*, 198 (1-4): 9-976
- [107]. Rose, K., Smith, E.P., Gardner, R.H., and Bartell, S.M. 1991. "Parameter sensitivities, Monte Carlo filtering, and model forecasting under uncertainty", *J. Forecasting*, 10, 117-133.
- [108]. Richard, W.K., 2002. Techniques for estimating uncertainty in climate change scenarios and impact studies. *Climate Research*, vol. 20: 167-185
- [109]. Richard, S., Bourdillon, R., Roussel, D. and Turcotte, R., 2013. Global calibration of distributed hydrological models for large-scale applications. *Journal of Hydrologic Engineering*, 18: 719-721
- [110]. Schwarz, G., 1978. Estimating the dimension of a model. *The annals of Statistics*, 6, 2, 461-464

- [111]. Shuttleworth, W.J., 1993. Evaporation, in Handbook of hydrology, edited by Maidment, D.R., pp. 4.1-4.53, McGraw-Hill, Inc., New York
- [112]. Singh, S.K., Bardossy, A. and Wagener, T., 2008. Robust Dynamic parameter estimation for hydrological models. American Geophysical Union, fall meeting, abstract # H43D-1031
- [113]. Singh, V., Bankar, N., Salunkhe, S.S., Bera, A.K. and Sharma, J.R., 2013. Hydrological stream flow modeling on Tungabhadra catchment: parameterization and uncertainty analysis using SWAT CUP. Current Science, vol. 104 (9)
- [114]. Seibert, J., 2000. Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. Hydrol. Earth System Sci, 4 (2), 215-224
- [115]. Sorooshian, S. and Gupta, V.K., 1985. The analysis of structural identifiability: Theory and application to conceptual rainfall-runoff models. Water Resour. Res., 21 (4), 487-495
- [116]. Schoups, G., van de Giesen, N.C. and Savenije, H.G., 2008. Model complexity control for hydrologic prediction. Water Resources Research, vol. 44, W00B03, doi: 10.1029/2008WR006836
- [117]. Singh, V.P. and Frevert, D.K., 2002. Mathematical models of large watershed hydrology. Water Resour. Publ., 891, Highlands Ranch, Colo.
- [118]. Shrestha, R., Tachikawa, Y. and Takara, K., 2006. Input data resolution analysis for distributed hydrological modeling. Journal of Hydrology, vol. 319: 36-50

- [119]. Schulz, K., Seppelt, R., Zehe, E., Vogel, H.J. and Attinger, S., 2006. Importance of spatial structures in advancing hydrological sciences. *Water Resour. Res.*, 42, W03S03, doi: 10.1029/2005WR004301
- [120]. Sivapalan, M., 2005. Pattern, process and function: Elements of a new unified hydrologic theory at the catchment scale. In *Encyclopaedia of Hydrologic Sciences*, vol. 1 (part 1), edited by Anderson, M.G., chap. 13: 193-219, John Wiley, Hoboken, H.J.
- [121]. Spear, R.C. and Hornberger, G.M., 1980. Eutrophication in Peel Intel – II. Identificaiton of critical uncertainties via generalized sensitivity analysis. *Water Research*, 14: 43-49
- [122]. Schoups, G., van de Giesen, N.C. and Savenije, H.G., 2008. Model complexity control for hydrologic prediction. *Water Resources Research*, vol. 44, W00B03, doi: 10.1029/2008WR006836
- [123]. Thienann, M., Trosset, M.W., Gupta, H.V. and Sorooshian, S., 2001. Recursive Bayesian inference of hydrologic models. *Water Resour. Res.*, 37 (10): 2521-2535
- [124]. Torch, P.A., Carrillo, G.A., Heidbuchel, I., Rajagopal, S., Switanek, M., Volkmann, T.H.M. and Yaeger, M., 2009. Dealing with landscape heterogeneity in watershed hydrology: a review of recent progress toward new hydrological theory. *Geography Compass* 3(1), 375-392, 10.1111/j.1749-8198.2008.00186
- [125]. Uhlenbrook, S., Seibert, J., Leibundgut, C. and Rohde, A., 1999. Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structures. *Hydrological Sciences Bulletin*, 44 (5): 779-797

- [126]. Vano, J.A., Scott, M., Voisin, N., Stockle, C.O., Hamlet, A.F., Mickelson, K.E.B., Elsner, M.M. and Lettenmaier, D.P., 2010. Climate change impacts on water management and irrigated agriculture in the Yakima River basin, Washington, US. Washington State Climatic Change, doi: 10.1007/s10584-010-9856-z
- [127]. Vrugt, J.A., Schoups, G., Hopmans, J.W., Young, C., Wallender, W.W., and Harter, T., 2005. Inverse modeling of large-scale spatially distributed vadose zone properties using global optimization. Water Resour. Res., 41, W06003, doi: 10.1029/2004WR003698
- [128]. VIC manual, Gao, H., Tang, Q., Shi, X., Zhu, C., Bohn, T.J., Su, F., Sheffield, J., Pan, M., Lettenmaier, D.P., and Wood, E.F., 2010. Water Budget record from Variable Infiltration Capacity (VIC) model. Algorithm Theoretical Basis document for Terrestrial Water Cycle Data Records. <http://www.hydro.washington.edu/Lettenmaier/Models/VIC/Overview/ModelOverview.shtml>
- [129]. Wagener, T., 1998. Developing a prototype knowledge based system to support rainfall-runoff modeling. M.Sc. Thesis, Delft University of Technology, Delft, NL
- [130]. Wheeler, H.S., Jolley, T.J., Onof, C., Mackay, N. and Chandler, R.E., 1999. Analysis of aggregation and disaggregation effects for grid-based hydrological models and the development of improved precipitation disaggregation procedures for GCMs. Hydrol. Earth Syst. Sci., 3: 95-108

- [131]. Wagener, T., Boyle, D.P., Lees, M.J., Wheater, H.S., Gupta, H.V. and Sorooshian, s., 2001. A framework for the development and application of hydrological models. *Hydrology and Earth System Sciences*, 5 (1): 13-26
- [132]. Wagener, T. and Wheater, H.S. 2006. Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. *Journal of Hydrology*, 320(1-2), 132-154.
- [133]. Wagener, T., McIntyre, N., Lees, M.J., Wheater, H.S. and Gupta, H.V., 2003a. Towards reduced uncertainty in conceptual rainfall-runoff modeling: Dynamic identifiability analysis. *Hydrological Processes*, 17: 455-476
- [134]. Wagener, T. and Gupta, H.V. 2005. Model identification for hydrological forecasting under uncertainty. *Stochastic Environmental Research and Risk Assessment*. DOI 10.1007/s00477-005-0006-5.
- [135]. Wagener, T., Sivapalan, M., Torch, P.A. and Woods, R.A., 2007. Catchment classification and hydrologic similarity, *Geogy. Compass*, 1, doi: 10.1111/j.1749-8198.2007.00039x
- [136]. Wagener, T., Sivapalan, M., Torch, P.A., McGlynn, B.L., Harman, C.J., Gupta, H.V., Kuman, P., Rao, P.S.C., Basu, N.B. and Wilson, J.S., 2010. The future of hydrology: An evolving science for a changing world. *Water Resources Research*, vol. 46, doi: 10.1029/2009WR008906
- [137]. Wagener, T., Camacho, L.A. and Wheater, H.S., 2002b. Dynamic parameter identifiability of a solute transport model. *J. Hydroinformatics*, 04 (3): 199-211

- [138]. Whittaker, G., Confesor Jr, R., Di Luzio, M. and Arnold, J.G., 2010. Detection of overparameterization and overfitting in an automatic calibration of SWAT. Transactions of the ASABE, 53 (5), pp. 1487-1499, doi: 10.13031/2013.34909
- [139]. Wood, E.F., 1991. Global scale hydrology – advances in land surface modeling. Reviews of Geophysics 29: 193-201 Part 1 Supl. S.
- [140]. Wood, E.F., Lettenmaier, D.P. and Zartarian, V.G., 1992. A land-surface hydrology parameterization with subgrid variability for general-circulation models. Journal of Geophysical Research-Atmospheres, 97 (D3): 2717-2728
- [141]. Wood, E.F., Lettenmaier, D.P., Liang, X., Nijssen, B. and Wetzel, S.W., 1997. Hydrological modeling of continental-scale basins. Annual. Rev. Earth Planet Science, 25: 279-300
- [142]. Wooldridge, S.A., Kalma, J.D. and Walker, J. P., 2003. Importance of soil moisture measurements for inferring parameters in hydrologic models of low-yielding ephemeral catchments. Environmental modeling & software, Vol. 18 (1): 35-48
- [143]. Yapo, P., Gupta, H., and Sorooshian, S., (1998). Multi-objective global optimization for hydrologic models. Journal of Hydrology, Vol. 204, 1-4, pp. 83-97
- [144]. Yao, H., Hashino, M., 2001. A completely-formulated distributed rainfall-runoff model for the catchment scale. IAHS publication: 183-190
- [145]. Zhao, R.J., et al., 1980. The Xinanjiang model. Hydrological Forecasting Proceedings Oxford Symposium, IASH 129, 351-356



- [146]. Zeug, S.C. and Winemiller, K.O., 2007. Relationship between hydrology, spatial heterogeneity, and fish recruitment dynamics in a temperate floodplain river. *River Research and Applications*, vol. 24 (1), doi: 10.1002/rra.1061

## APPENDICES

## Appendix A An Approach to Measure Parameter

### Identifiability

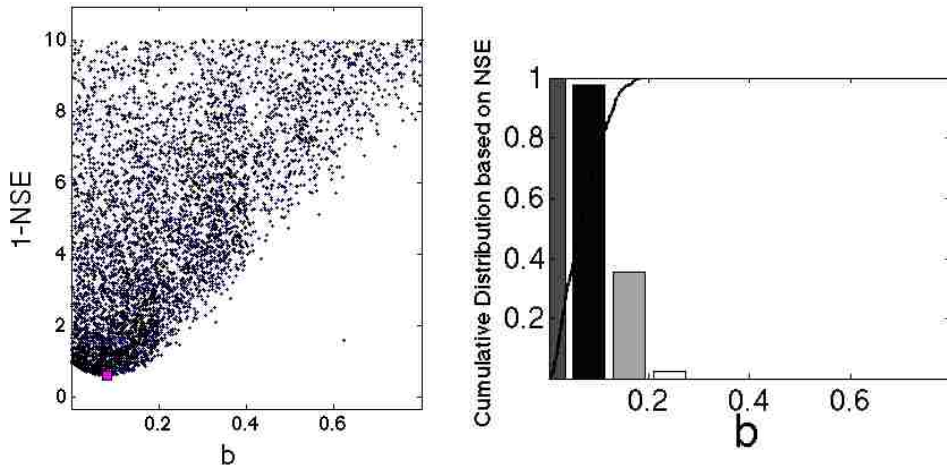
The measurement of identifiability is based on a projection of parameter space using Monte Carlo Uniform Randomly Simulation (MCURS). Choosing a parameter prior distribution with an upper and lower bound limits as feasible parameter space, then divide the Monte Carlo sampling into equally ten groups based on the rank of model performance with respect to a selected objective function. Then calculating the cumulative distribution of the best performing in each group. For parameters' prior distribution as uniformly, accordingly, the cumulative distribution should be a straight line with non-identified properties. So the gradient of the cumulative distribution line could be seen as an indicator to measure how well or poor identifiable for a parameter (Figure A-1). Take the Salt River basin and San Pedro River basin as an example with parameter  $b$  and  $d_2$ .

Figure A-2 shows the same parameter  $b$  with different climatic gradient region. From the scatter plot, it could be recognized from the identifiability definition, the  $b$  in San Pedro River basin should be highly identifiable than in Salt River basin, in correspondence to the gradient diagram, the cumulative distribution line with a steeper gradient in San Pedro than in Salt River basin.

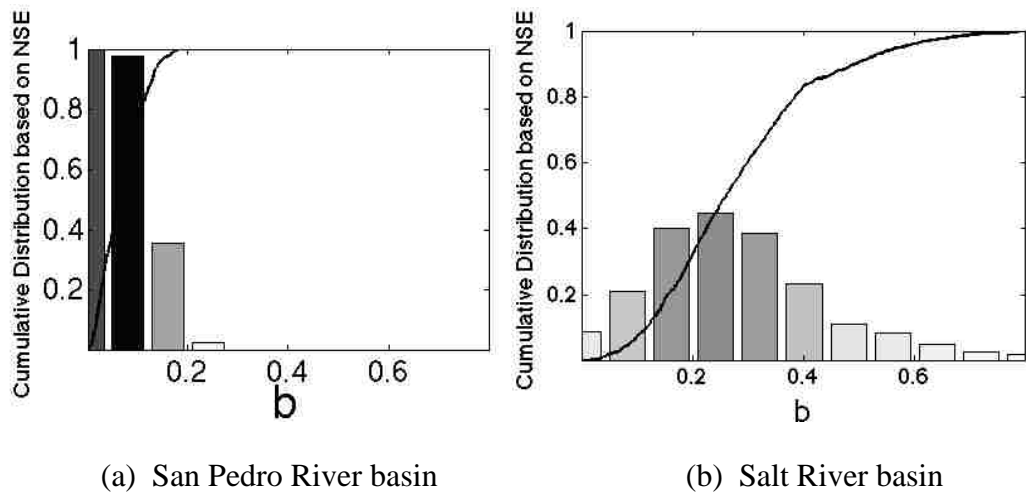
Figure A-3 indicates the measurement of identifiability with different parameter  $b$  and  $d_1$  on San Pedro River basin. Obviously, the parameter  $b$  is much identified than  $d_1$ .

This approach could be a standard of choosing a fit degree of model complexity with respect to specific hydrologic behavior, climatic gradient or hydrologic impacts analysis.

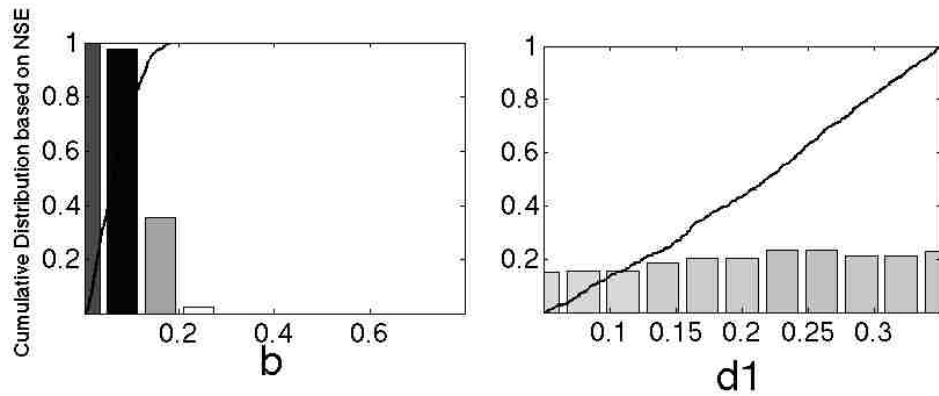
If the gradient was examined in every time steps, it will be extended to the approach to parameter temporal dynamic heterogeneity analysis, the details of the application are listed in Appendix B. How gradients for a well identified and poorly identified parameter are recognized as a function of time step.



**Figure A-1 Shows the definition and calculation to measure identifiability. Take parameter b on San Pedro River basin as an example.**



**Figure A-2 Shows the identifiability to same parameter b with different climatic gradient region. The parameter b has much higher identifiability in San Pedro River basin than Salt River basin**



**Figure A–3 Measurement of identifiability with parameter  $b$  and  $d_1$  on San Pedro River basin. It demonstrates the parameter  $b$  has much higher identifiability than parameter  $d_1$ . The cumulative distribution based on NSE with parameter  $d_1$  is almost straight line which means non-identified in this region.**

## **Appendix B Application of the Approach to Measure**

### **Identifiability - DYNIA**

The Dynamic Identifiability Analysis (DYNIA) is a new method to locating periods of high identifiability for individual parameters through time series and could detect failures of model components (Wagener et al., 2003a). It is appropriate to improve the amount of information that can be extracted from simulations under the context of “cluster” optimization scheme. The elements of this method include employed by Regional Sensitivity Analysis (RSA) and the Generalized Likelihood Uncertainty Estimation (GLUE).

The steps of this approach can flow through Figure B-1. Firstly, examining the feasible parameter space based on the population-searching algorithm using Monte Carlo method with a uniform prior distribution to project a random prior parameter space; Secondly, calculating the identifiability measurement (Appendices A) for the best performing parameter values (e.g., the top 10% sections); Finally, segmenting the range of each parameter and calculating the identifiability measurement (gradient) in each container. The procedures are employed in each time step. So the plot results could be shown as a function of time with identifiability measurement. The degree of identifiability was recognized as color grids, with well identifiability of the parameter in dark grid, in versus, sign a light color grid when is poor identified.

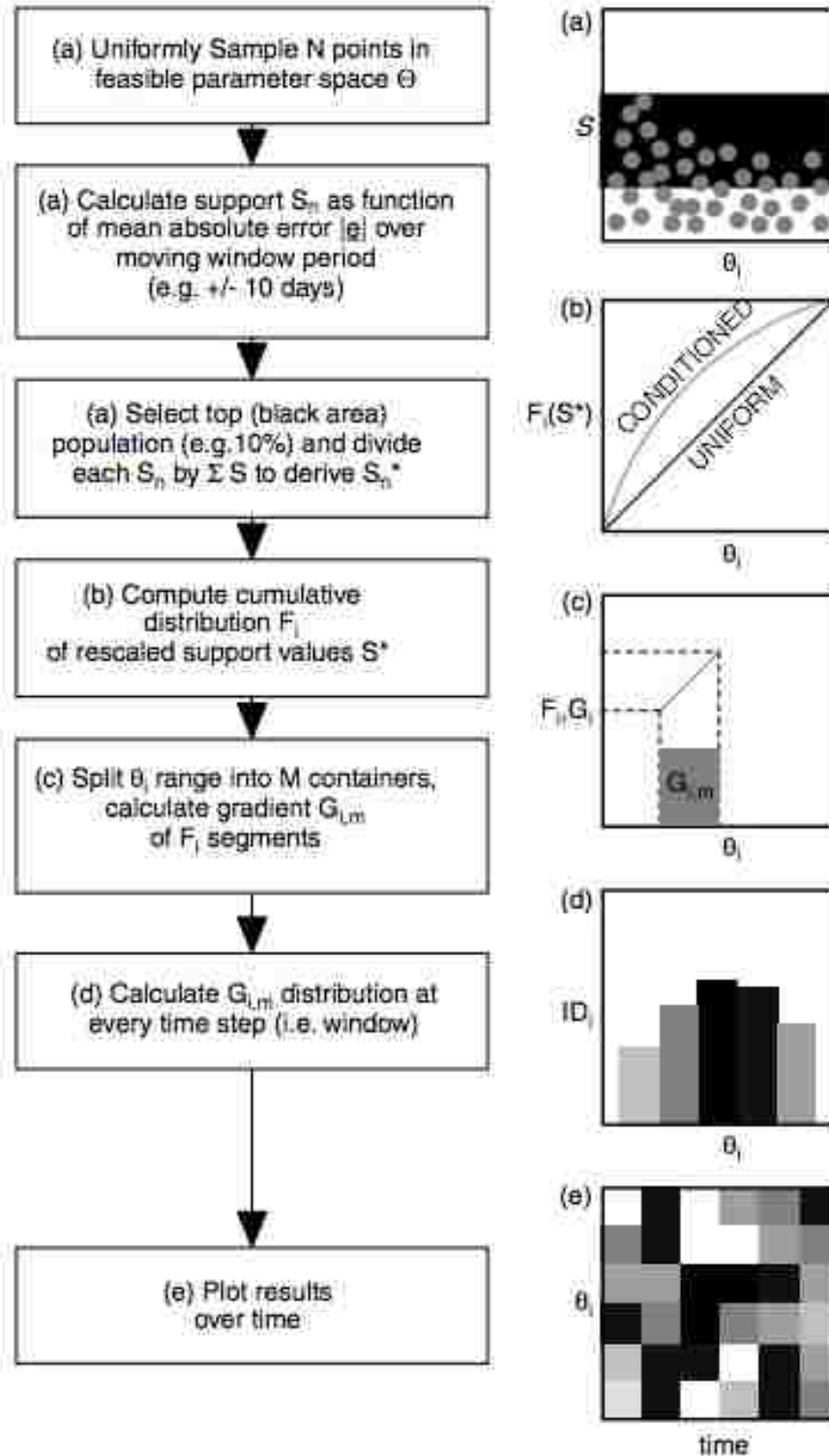


Figure B-1 Procedure to DYNIA (Wagener et al., 2003a)