IMPROVEMENT OF SUBSURFACE FLOW PREDICTABILITY USING LAND SURFACE MODEL IN THE UNSATURATED ZONE AT VARIOUS SPATIAL SCALES

A Dissertation

by

JONGGUN KIM

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Binayak P. Mohanty
Committee Members,	Gretchen R. Miller
	Hong-Bin Zhan
	Patricia K. Smith
Head of Department,	Stephen W. Searcy

August 2016

Major Subject: Biological and Agricultural Engineering

ABSTRACT

Subsurface flow in the unsaturated zone is an important component of the hydrologic cycle and plays a significant role in the water and energy balance through affecting various hydrological processes. Land surface models (LSMs) have been developed and extended during the past decades with various enhanced processes to understand and quantify the complex interaction between atmosphere and land surface systems. However, there are still critical deficiencies (e.g., simplified processes and parameterization) remaining in simulating land surface hydrology for land surface modeling. Thus, this dissertation focuses on understanding land surface processes from various land surface models and improving land surface processes and parameterization in land surface modeling in the unsaturated zone at various spatial scales.

Two main approaches (Bayesian Model Averaging (BMA) based multi-model simulation and physically based hydrologic connectivity approach) to improve the land surface modeling predictability are presented in this dissertation. The BMA-based multi-model simulation approach was developed to reflect the strengths of the models under various land surface wetness conditions and to quantify the model parameter and structural uncertainties. The physically-based hydrologic connectivity concept was proposed to characterize the subsurface flow variability based on spatially distributed patterns of wetness condition or physical controls (e.g., soil type, vegetation, topography). Hydrologic connectivity is an important concept for

understanding local processes in the context of catchment hydrology and defining flow path continuity in surface and subsurface flows. These approaches were applied in land surface modeling and tested in various hydro-climate regions and spatial scales showing significant improvement of modeling predictability. Based on the knowledge and experience gained from this dissertation, the proposed concepts will be useful to improve the hydrological model performance and better understand the subsurface flow variability in the unsaturated zone at various scales.

DEDICATION

I dedicate this dissertation to the memory of my late father (Wonbae Kim).

ACKNOWLEDGEMENTS

This dissertation would not have been successfully completed without the help of my lovely wife (Miae Hyun), family, advisors, and friends. I appreciate their encouragement, support, and love over the years.

I would like to express my sincere gratitude to Dr. Binayak P. Mohanty, my doctoral committee chair, for improving my research skills and encouraging me to successfully conduct my research. I also wish to thank my doctoral committee members Dr. Gretchen R. Miller, Hong-Bin Zhan, and Dr. Patricia Smith for their guidance and valuable comments during my research work.

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iv

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

LIST OF FIGURES viii

LIST OF TABLES xii

CHAPTER I GENERAL INTRODUCTION 1

1.1 Problem Statement 1
1.2 Motivation 2
1.3 Hypothesis 3
1.4 Research Objectives 3

CHAPTER II EFFECTIVE SOIL MOISTURE ESTIMATE AND ITS UNCERTAINTY USING MULTI-MODEL SIMULATION BASED ON BAYESIAN MODEL AVERAGING 8

2.1 Synopsis 8
2.2 Introduction 10
2.3 Methodology 13
2.3.1. Bayesian Model Averaging Based Multimodel Simulation Approach 13
2.3.2. Characteristics of the Hydrological Models 15
2.3.3 Modified-microGA (Genetic Algorithms) for Estimating Optimal Parameters and Their Uncertainty 23
2.3.4. Bayesian Model Averaging Scheme Based on the Land Surface Wetness Conditions and Model Structural Uncertainty 26
2.3.5. Statistical Analysis 29
2.3.6 Study Area and Description of Model Conditions 30
2.4 Results and Discussion 33
2.4.1 Estimation of Optimized Model Parameters and Their Uncertainties 33
2.4.2 Estimation of Effective Surface Soil Moisture and Its Uncertainty 46
2.5 Conclusions 53

CHAPTER III INFLUENCE OF LATERAL SUBSURFACE FLOW AND CONNECTIVITY ON SOIL WATER STORAGE IN LAND SURFACE MODELING 56

3.1 Synopsis 56
3.2 Introduction 57
3.3 Methodology 61
3.3.1 Study Area 61
3.3.2 Description of Model Condition and Forcing Data 63
3.3.3 Lateral Subsurface Flow Process 65
3.4. Results and Discussion 81
3.4.1 Field Scale (El-Reno Site 5) 81
3.4.2 Sub-watershed Scale (El-Reno Sub-watershed) 87
3.5. Conclusions 94

CHAPTER IV A PHYSICALLY-BASED HYDROLOGICAL CONNECTIVITY ALGORITHM FOR DESCRIBING SPATIAL PATTERNS OF SOIL MOISTURE IN THE UNSATURATED ZONE 97

4.1 Synopsis 97
4.2 Introduction 98
4.3 Methodology 102
4.3.1 Study Areas 102
4.3.2 Land Surface Model (Community Land Model) 105
4.3.3 Mixed Physical Controls in Complex Landscapes 111
4.3.4 Development of Physically-based Hydrologic Connectivity 113
4.4. Results and Discussion 117
4.4.1 Effects of Mixed Physical Controls on Soil Moisture Variability 117
4.4.2 Calibration of Soil Hydraulic Properties Based on Hydrologic Connectivity 121
4.4.3 Comparison of CLM Output Using Default and Calibrated Soil Parameters 126
4.5. Conclusions 137

CHAPTER V GENERAL CONCLUSIONS 139

REFERENCES 143

LIST OF FIGURES

Page

Figure 2. 1 Schematic diagram of the Bayesian Model Average (BMA) based multi-model simulation approach. 15

Figure 2. 2 Study sites; (a) Walnut Gulch (WG) 82 in Arizona, (b) Little Washita (LW) 13, and (c) SCAN 2023 in Oklahoma. 31

Figure 2. 3 Probability distributions and quantile box plots of the searched soil parameters of the three hydrological models using the multiple random number seeds (i.e., -1000, -950, and -750) for the LW 13 site. 35

Figure 2. 4 *In situ* and simulated surface soil moisture (0~5 cm) dynamics using the optimized soil parameters derived by the modified-*micro*GA for (a) SWAP, (b) Noah LSM, and (c) CLM at the LW 13 site during calibration and validation periods. 37

Figure 2. 5 Probability distributions and quantile box plots of the searched soil parameters of the three hydrological models using the multiple random number seeds (i.e., -1000, -950, and -750) for the WG 82 site. 40

Figure 2. 6 *In situ* and simulated surface soil moisture (0~5 cm) dynamics using the optimized soil parameters derived by the modified-*micro*GA for (a) SWAP, (b) Noah LSM, and (c) CLM at the WG 82 site during calibration and validation periods. 42

Figure 2. 7 *In situ* and simulated surface soil moisture (0~5 cm) dynamics using the optimized soil parameters derived by the modified-*micro*GA for SWAP, Noah LSM, and CLM at the SCAN 2023 site. 45

Figure 2. 8 *In situ* and simulated surface soil moisture using a single (S-BMA, dotted line) and multiple (M-BMA, black line) sets of the BMA weights and ±95 PCI at the LW 13 site. 47

Figure 2. 9 *In situ* and simulated surface soil moisture using a single (S-BMA, dotted line) and multiple (M-BMA, black line) sets of the BMA weights and ±95 PCI at the WG 82 site. 50

Figure 2. 10 *In situ* and simulated surface soil moisture using a single (S-BMA, dotted line) and multiple (M-BMA, black line) sets of the BMA weights and ±95 PCI at the SCAN 2023 site. 53

Figure 3. 1 Study sites for (a) El Reno 5 (ER 5) matching the ESTAR remote sensing footprint with multi-depth ground based soil water measurements

using truck-mounted Giddings probe (100 m spacing) and (b) El Reno sub-watershed (ER-sub) in Oklahoma. 63

Figure 3. 2 Three study cases designed for the lateral subsurface flow process. Anisotropy (α) is used to derive the saturated hydraulic conductivity in vertical and lateral directions as uniform (case 2) or connectivity-based spatially-varying (case 3) ratio. 68

Figure 3. 3 (a) *In situ* measurements at top 5 cm (pixel size: 100 × 100 m), (b) indicator maps for various thresholds of degree of soil wetness (θ/θ_s) on sampling dates, and (c) hydrologic connectivity for 5 sampling dates. Optimum threshold values for daily soil wetness were identified based on visual examination of the connectivity function vs. separation distance plots. Note that selected red boxes around indicator maps correspond to the optimum thresholds selected from the connectivity functions, representing distinct connected patterns on various sampling dates. 73

Figure 3. 4 Spatially-varying anisotropy ratio maps (pixel size: 100 × 100 m) (in 8 directions) derived from the connectivity patterns for the near surface soil layers (1st to 3rd) by combining optimum indicator maps for all sampling dates. Similar maps of the other layers (not shown) were derived from the soil moisture measured at deeper soils (up to 90cm). 75

Figure 3. 5 Dominant physical controls ((a) NDVI, (b) %Clay, (c) %Sand, and (d) Topographic Index) and (e-f) their connectivity functions for the ER-sub site. Thresholds values for different physical controls were defined based on its range and numerical analyses. Optimum threshold values for individual physical controls were identified based on visual examination of the connectivity function vs. separation distance plots. 78

Figure 3. 6 (a) Optimum indicator maps of various physical controls (NDVI, TI, %Clay, %Sand) for various soil layers (1st – 10th), (b) combined indicator maps for each soil layer, and (c) corresponding spatially-varying anisotropy ratio maps (pixel size: 100 × 100 m) at the ER-sub site. 80

Figure 3. 7 Comparison of the root zone soil moisture of (a) ground observation (pixel size: 100 × 100 m), (b) original CLM model, and modified CLM model (pixel size: 50 × 50 m) through (c) case 1, (d) case 2, and (e) case 3 at the ER 5 site, and (f) differences between the original and modified CLM model of case 3. 82

Figure 3. 8 Comparison of the simulated root zone soil moisture using the original and modified model against the observations; (a) case 1, (b) case 2, and (c) case 3 in July 6th and (d) case 3 in July 15th for the ER 5 site. 83

Figure 3. 9 Comparison of observed and simulated (case 3) near surface soil moisture dynamics (top 5 cm) using original (dotted line) and modified model (black line) at the ER 5 site..... 87

Figure 3. 10 Comparison of the root zone soil moisture (pixel size: 800 × 800 m) at various depths of (a) original model and (b) modified model (case 3), and (c) their differences for the ER-sub site..... 88

Figure 3. 11 Comparison of the simulated near surface soil moisture for (a) the average within the ER-sub site against ESTAR observations and (b) the differences of R values for each grid cell. 92

Figure 3. 12 Simulated evapotranspiration (ET), surface runoff, subsurface drainage, and water storage using the original and modified model (pixel size: 100 × 100 m) at the ER-sub site. 93

Figure 4. 1 Study sites of (a) Little Washita (LW) in Oklahoma and (b) Upper South Skunk (USS) in IOWA. The pixels represent connected and unconnected regions selected for analysis. 104

Figure 4. 2 Schematic diagram of information flow for developing connectivity index using Bayesian averaging of dominant physical controls and calibrating distributed soil hydraulic parameters. α_i is the calibrating factor for each parameter based on their standard deviation that is determined by the physically-based hydrologic connectivity index..... 110

Figure 4. 3 Dominant physical controls (soil texture, vegetation, and topography) for the (a) LW and (b) USS watersheds (Spatial resolution of 800m). 113

Figure 4. 4 Physically-based Hydrologic connectivity: (a) connectivity functions ($\tau(d)$) calculated using indicator maps ($I(y)$) of mixed physical controls, (b) indicator maps for 5 selected thresholds (s_i), and (c) physically-based hydrologic connectivity index map developed by integrating 5 indicator maps. 115

Figure 4. 5 Contributing ratios (weights, w_i) of physical controls ((a) for LW and (b) for USS) and mixed (weighted) physical controls maps ((c) for LW and (d) for USS). 118

Figure 4. 6 Comparison of spatial variations of measured saturated hydraulic conductivity (K_{sat}) and normalized mixed physical controls (%clay, %sand, NDVI, TI) according to soil types across the LW watershed. (CL: Clay Loam, L: Loam, LS: Loamy Sand, S: Sand, SiL: Silty Loam, SL: Sandy Loam) 120

Figure 4. 7 Connectivity functions for 5 representative thresholds for soil layers and connectivity index for the LW watershed. Pixels of higher index represent highly connected and higher wetness regions; lower index pixels indicate unconnected and drier regions. 122

Figure 4. 8 Connectivity functions for 5 representative thresholds for soil layers and connectivity index for the USS watershed. Pixels of higher index represent highly connected and higher wetness regions; lower index pixels indicate unconnected and drier regions. 123

Figure 4. 9 (a) Dominant physical controls, (b) default and calibrated soil hydraulic parameters, and (c) soil moisture prediction in the current (CLM) and calibrated model for a selected region (800 × 800m resolution) in the LW watershed. 125

Figure 4. 10 Comparisons of default and calibrated soil parameters according to soil texture (%sand and %clay). θ_{sat}, K_{sat}, and b are estimated based on %Sand only; ψ_{sat}, is dependent on %Clay only in pedo-transfer function of CLM. 126

Figure 4. 11 Comparison of default parameters and calibrated parameters on 5 selected pixels for the LW watershed. The bar shows the ranges of parameters for 11 soil texture classes obtained from *Cosby et al.* [1984]... 128

Figure 4. 12 (a) Comparisons of simulated and measured soil moisture dynamics on 5 pixels selected on connected and unconnected regions and (b) Correlation and RMSE with ESTAR measurements for 5 pixels in the LW watershed. 129

Figure 4. 13 Comparison of default parameters and calibrated parameters on 5 selected pixels for the USS watershed. The bar shows the ranges of parameters for 11 soil texture classes obtained from *Cosby et al.* [1984]... 131

Figure 4. 14 (a) Comparisons of simulated and measured soil moisture dynamics on 5 pixels selected on connected and unconnected regions and (b) Correlation and RMSE with ESTAR measurements for 5 pixels in the USS watershed. 132

Figure 4. 15 (a) Comparisons of measured and simulated soil moisture and (b)-(e) model outputs using default and calibrated parameters (evaporation, transpiration, surface runoff, and water storage) for the LW watershed. 135

Figure 4. 16 (a) Comparisons of measured and simulated soil moisture and (b)-(d) model outputs using default and calibrated parameters (ET, surface runoff, and water storage) for the USS watershed. 136

LIST OF TABLES

Page

Table 2. 1 Comparison of main characteristics of the three hydrological models. 18

Table 2. 2 Summary of soil hydraulic parameters and its feasible ranges used in the modified-microGA for the three hydrological models. 23

Table 2. 3 A single and multiple sets of the Bayesian Model Average (BMA) weights for the three hydrological models at the LW 13 site. 46

Table 2. 4 A single and multiple sets of the Bayesian Model Average (BMA) weights for the three hydrological models at the WG 82 site..... 49

Table 2. 5 A single and multiple sets of the Bayesian Model Average (BMA) weights for the three hydrological models at the SCAN 2023 site..... 52

Table 3. 1 Thresholds of wetness (s) of the near surface soil moisture measurements for the ER 5 site. 74

Table 3. 2 Thresholds of wetness of the root zone soil moisture measurements with depth for the ER 5 site. 76

Table 3. 3 Comparison of spatial patterns of simulated near surface soil moisture using the original and modified model with various spatial moving window sizes..... 90

Table 4. 1 Characteristics of selected pixels in the study sites. 105

Table 4. 2 Means and standard deviations for the four hydraulic parameters for various textural classes (from Table 3 in Cosby et al. [1984]). 108

Table 4. 3 Correlation coefficients and RMSE between soil moisture measurements and simulations (top 5cm) using default and calibrated soil parameters for the two study sites..... 134

CHAPTER I

GENERAL INTRODUCTION

1.1 Problem Statement

Subsurface flow in the unsaturated zone is an important hydrologic process and a key component of the water budget. Through its direct impacts on soil moisture, it can affect the water and energy fluxes at land surface and influence the regional climate and water cycle [*Gochis and Chen,* 2003; *Kumar,* 2004]. In particular, root zone soil moisture that is held in a shallow subsurface (unsaturated zone) plays a significant role in partitioning of precipitation to surface runoff versus infiltration and subsequently the infiltrated water to evapotranspiration versus recharge [*Zhu and Mohanty,* 2006]. Soil moisture information about the shallow subsurface can be obtained at various scales ranging from a local scale (*in situ* sampling) to a remote sensing scale (e.g., AMSR-E, SMOS, SMAP). Field observations (*in situ*) provide soil moisture profiles at a fine scale resolution but these are very sparse, while remotely sensed soil moisture products provide only near-surface soil moisture (1~5 cm) but enable us to estimate soil moisture of large areas. Various up- or down-scaling approaches have been developed to account for the discrepancy between spatial and temporal scales through hydrological modeling. Many hydrological models have been developed and widely used for root zone soil moisture predictions at various scales. However, these models still have limitations to accurately estimate root zone soil moisture due to uncertainties in their inherent model parameterization and structures. Furthermore, current hydrological models based on

simplified processes for subsurface flow cannot effectively account for catchment hydrologic characteristics and identify the spatially distributed soil moisture in complex heterogeneous landscapes. Hence, alternate methods are necessary to reduce errors due to the model parameterization and structural uncertainties in estimating effective root zone soil moisture. The simplified processes and parameterization in hydrological models also need to be improved to better understand spatially distributed patterns of subsurface flow in the unsaturated zone.

1.2 Motivation

Land surface hydrologic processes are important components in land surface modeling to understand and quantify the complex interaction between atmospheric and land surface systems. Land surface models (LSMs) have been developed and extended over past decades with various enhanced processes. However, there are still critical deficiencies remaining in simulating land surface hydrology for land surface modeling. It is necessary to explore several hydrological models widely used in land surface modeling to compare their characteristics in land surface processes, and then strengths of the models in estimating root zone soil moisture can be reflected through multi-model simulations.

In addition, most of the land surface models have simplified processes for subsurface flow in the unsaturated zone. Particularly, they are neglecting the horizontal exchanges of water (lateral subsurface flow) at the grid or sub-grid scales, focusing only on the vertical exchanges of entities including water. A one-dimensional process cannot

properly capture the spatially distributed patterns of surface or subsurface flow in heterogeneous landscapes. Lateral subsurface flow should be considered in land surface modeling to successfully describe the spatial variability of subsurface flow in the unsaturated zone. Hydrologic connectivity is a promising concept for understanding local processes in the context of catchment hydrology. It is defined as flow path continuity in surface and subsurface flows and connectivity metrics for spatial distributed wetness patterns in complex landscapes. The concept of hydrologic connectivity can be used to understand spatial variability of subsurface flow and improve current subsurface flow processes and parameterization schemes in land surface modeling.

1.3 Hypothesis

The hypothesis of this study is that hydrological models have their own strengths for particular soil wetness conditions (e.g., wet, moderately wet, and dry conditions) in estimating effective root zone soil moisture dynamics. Hydrologic connectivity derived from soil wetness conditions and dominant physical controls (e.g., soil properties, topography, and vegetation) is useful to describe the spatial variability of subsurface flow in the unsaturated zone.

1.4 Research Objectives

The primary objective of this research is to improve subsurface flow predictability using land surface models in the unsaturated zone. Based on the above

hypothesis, modeling capabilities of various hydrological models were explored to identify their strengths under different wetness conditions. In addition, hydrologic connectivity based on surface wetness conditions and dominant physical controls was included in land surface modeling to enable better prediction of subsurface flow in the unsaturated zone.

The research objectives are:

- To develop a BMA(Bayesian Model Averaging)-based multi-model simulation approach based on the land surface wetness conditions, estimating effective soil moisture dynamics with a genetic algorithm scheme for soil hydraulic parameter optimization, and to evaluate various model parameters and structural uncertainties under various hydro-climatic conditions,
- To develop better hydrologic understanding and modeling capability in complex landscapes using a connectivity-based lateral subsurface flow algorithm, and to demonstrate subsurface flow variability effectively using spatially distributed patterns of root zone wetness conditions and physical controls at field and sub-watershed scales,
- To investigate the effects of mixed physical controls on soil moisture spatial variability in two different hydro-climate regions, to develop a physically-based hydrologic connectivity algorithm to better understand catchment hydrologic characteristics and identify spatial patterns in soil moisture, and to improve parameterization of soil hydraulic properties using physically-based

hydrologic connectivity to better describe the soil moisture spatial variability in land surface modeling.

In Chapter II, three hydrological models (i.e., Noah Land Surface Model, Noah LSM; Soil-Water-Atmosphere-Plant, SWAP; and Community Land Model, CLM) which have been widely used in modeling of land surface water and energy fluxes at various spatial scales were chosen to explore strengths and weaknesses of their model parameterization and structures in estimating near surface soil moisture dynamics. Genetic Algorithm (GA) and other evolutionary algorithms were used for parameter estimation in the hydrological models using inverse modeling, and a Bayesian technique (Bayesian Model Averaging (BMA) with condition-specific weights) was employed for evaluating the uncertainties in hydrologic predictions related to soil moisture and hydrologic fluxes. The BMA-based multi-model simulation approach based on various soil wetness conditions was developed to effectively reflect strengths of the models and to quantify model parameter and structural uncertainties. This approach was tested in two different hydro-climatic conditions (Little Washita (LW 13) site in Oklahoma (sub-humid) and Walnut Gulch (WG 82) site in Arizona (semiarid)). Model simulations using the approach were validated with the *in situ* soil moisture measurements (0–5 cm) during the Southern Great Plains experiment 1997 (SGP97, day of year (DOY): 170–197) for the LW 13 site and Soil Moisture Experiment 2004 (SMEX04, DOY: 216–238) for the WG 82 site.

In Chapter III, the subsurface flow processes in land surface modeling (in CLM) was modified with a three-dimensional flow process based on Richards' equation in the unsaturated zone to consider lateral flow process in the unsaturated zone. To investigate the impact of subsurface lateral flow and its connectivity on soil water storage three cases were designed including (1) complex surface topography only, (2) complex surface topography in upper soil layers and soil hydraulic properties with uniform anisotropy, and (3) complex surface topography and soil hydraulic properties with spatially varying anisotropy. The concept of hydrologic connectivity was employed to derive the spatially varying anisotropy which was used for estimating lateral hydraulic conductivity. Hydrologic connectivity was developed using spatial patterns of soil wetness conditions and physical controls (e.g., soil type, vegetation, topography) assuming that the variables have equal effects on the hydrological processes. The assumption may not be applicable into other landscapes, but it is difficult to identify their contributions to redistributing subsurface soil moisture because it can vary with complex landscape characteristics. In addition, the assumption was validated in this study. These cases were tested at two different scales (El-Reno site 5 (ER 5: field scale) and El-Reno sub-watershed (ER-sub: sub-watershed scale) located in the North Canadian River basin in Oklahoma). The proposed approach was validated with the daily *in situ* soil moisture (49 sampling points) measured in top 5 cm soil (18 June to 17 July) and in depths of 0–15, 15–30, 30–45, 45–60, and 60–90 cm (6–15 July) during the Southern Great Plains experiment 1997 (SGP97) for the ER 5 site, and with Electronically Scanning Thin Array Radiometer (ESTAR) pixel-based (800 × 800 m)

near surface soil moisture products obtained during SGP97 (18 June to 17 July) for the ER-sub site.

In Chapter IV, the hydrologic connectivity methodology developed in Chapter III was extended to consider the impact of mixed dominant physical controls based on their contribution ratios (no equal contribution) on soil moisture variability. The Bayesian averaging scheme was used to derive the contributing ratios of physical variables. The physically-based hydrologic connectivity algorithm was developed using indicator maps of mixed physical controls (based on the contributing ratios) and hydrologic connectivity functions for various thresholds. Soil hydraulic properties in land surface modeling were calibrated based on hydrologic connectivity to effectively reflect the spatial variability of subsurface flow. This approach was tested in two different hydro-climatic conditions (Little Washita (LW) watershed in Oklahoma (sub-humid) and Upper South Skunk (USS) watershed in IOWA (humid)). The proposed approach was validated with an Electronically Scanning Thin Array Radiometer (ESTAR) pixel-based (800 × 800 m) near surface soil moisture products obtained during SGP97 (June 18th – July 17th, 1997) for the LW watershed and Aircraft Polarimetric Scanning Radiometer (PSR) observed during SMEX02 (June 25th – July 12th, 2002) for the USS watershed.

Chapter V summarizes the general conclusions obtained from the proposed approach in this research work.

CHAPTER II

EFFECTIVE SOIL MOISTURE ESTIMATE AND ITS UNCERTAINTY USING MULTI-MODEL SIMULATION BASED ON BAYESIAN MODEL AVERAGING[1]

2.1 Synopsis

Various hydrological models have been developed for estimating root zone soil moisture dynamics. These models, however, incorporated their own parameterization approaches indicating that the output from the different model inherent structures might include uncertainties because we do not know which model structure is correct for describing the real system. More recently, multimodel approaches using a Bayesian Model Averaging (BMA) scheme have improved the overall predictive skill while individual models retain their own uncertainties for simulating soil moisture based on a single set of weights in modeling under different land surface wetness conditions (e.g., wet, moderately wet, and dry conditions). In order to overcome their limitations, we developed a BMA-based multimodel simulation approach using various soil wetness conditions for estimating effective surface soil moisture dynamics (0–5 cm) and quantifying uncertainties efficiently based on the land surface wetness conditions. The newly developed approach adapts three different hydrological models (i.e., Noah Land

[1] Reprinted with permission from "Effective Soil Moisture Estimate and its Uncertainty using Multi-Model Simulation based on Bayesian Model Averaging" by Jonggun Kim, Binayak P. Mohanty, and Yongchl Shin (2015), *J. Geophys. Res. Atmos., 120*, doi:10.1002/2014JD022905, Copyright 2015 American Geophysical Union.

Surface Model, Noah LSM; Soil-Water-Atmosphere-Plant, SWAP; and Community Land Model, CLM) for simulating soil moisture. These models were integrated with a modified-microGA (advanced version of original Genetic Algorithm (GA)) to search for optimized soil parameters for each model. Soil moisture was simulated from the estimated soil parameters using the hydrological models in a forward mode. It was found that SWAP performed better than others during wet conditions, while Noah LSM and CLM showed good agreement with measurements during dry conditions. Thus, we inferred that performance of individual models with different model structures can be different with varying land surface wetness. Taking into account the effects of soil wetness on different model performances, we categorized soil moisture measurements and estimated different weights for each category using the BMA scheme. Effective surface soil moisture dynamics were obtained by aggregating multiple weighted soil moisture. Our findings demonstrated that the effective soil moisture estimates derived by this approach showed a better match with the measurements compared to the original models and single-weighted outputs. A multimodel simulation approach based on land surface wetness enhances the ability to predict reliable soil moisture dynamics and reflects the strengths of different hydrological models under various soil wetness conditions.

2.2 Introduction

Soil moisture plays a key role in hydrologic processes such as soil water retention, infiltration, evapotranspiration, and groundwater recharge, which control water balance and land surface energy balance [*Zhu and Mohanty,* 2006; *Brocca et al.,* 2010; *Leung et al.,* 2011]. Several hydrological models have been developed and used widely for soil moisture predictions including Noah Land Surface Model (Noah LSM) [*Ek et al.,* 2003], Soil-Water-Atmosphere-Plant (SWAP) [*Van Dam et al.,* 1997], Community Land Model (CLM) [*Oleson et al.,* 2010], Variable Infiltration Capacity [*Liang et al.,* 1994], and Mosaic Land Surface Model (Mosaic LSM) [*Koster and Suarez,* 1996], among others. The Global Land Data Assimilation Systems use these hydrological models for validating pixel-scale soil moisture from satellite platforms and evaluating water/energy cycle and fluxes near the land surface [*Liu et al.,* 2009]. The North American Land Data Assimilation System has monitored and predicted hydrological drought conditions using state variables (e.g., soil moisture dynamics, runoff, evaporation, etc.) estimated from various hydrological models [*Ek et al.,* 2011]. However, these models incorporated with their own parameterization schemes and simplified processes might not consider adequately real-world conditions, indicating that each model has its own strengths and drawbacks for certain processes [*Hsu et al.,* 2009]. Thus, inherent model structures might produce different model outputs and cause uncertainties due to different model structures and input parameters (i.e., atmospheric forcings, soil textures, vegetation covers, initial and bottom boundary conditions, etc.). Many stochastic techniques and methods have been developed and extended to

overcome the limitations of modeling. Genetic algorithms (GAs) [*Holland,* 1975], Shuffled Complex Evolution-University of Arizona (SCE-UA) [*Duan et al.,* 1992], and Particle Swarm Optimization (PSO) [*Kennedy and Eberhart*, 2001] have been applied in estimating effective model parameters. Bayesian Model Averaging (BMA) [*Hoetting et al.,* 1999], Hydrological Uncertainty Processor [*Krzysztofowicz,* 1999; *Krzysztofowicz and Kelly*, 2000], Ensemble Model Output Statistics (E-MOS) [*Gneiting et al.,* 2005], and Model Conditional Processor [*Todini,* 2008; *Coccia and Todini*, 2011] have been used to account for the model structural uncertainties. GAs have been used to minimize errors in searching optimized model parameters based on inversion model [*Reed et al.,* 2000; *Ines and Mohanty*, 2008, 2009; *Zhang et al.,* 2009; *Shin et al.,* 2012; *Shin and Mohanty*, 2013; *Shin et al.,* 2013]. *Zhang et al.* [2008] integrated several global optimization algorithms (i.e., GA, SCE-UA, PSO, etc.) with Soil and Water Assessment Tool and compared their performances in calibrating model input parameters. They showed that GA found better optimized model parameters than others, although a large number of computational resources were required. Further, the near-surface [*Ines and Mohanty,* 2008] and layer-specific data assimilation [*Shin et al.,* 2012] approaches using GA coupled with SWAP based on inversion model were developed for quantifying effective soil hydraulic properties in the homogeneous and heterogeneous soil profiles. Their findings indicated that the estimated effective soil parameters at the near-surface and subsurface layers can be adequately conditioned by GA. However, although model parameter uncertainties for a single model can be minimized by simulation-optimization schemes (e.g., GA-SWAP, etc.), bias due to different model structures still remain

(considerably) in model outputs [*Hoetting et al.,* 1999; *Georgakakos et al.,* 2004; *Ajami et al.,* 2007]. A BMA scheme has been proposed to account for model structural uncertainties efficiently and improve predictive capabilities of different models through a weighted average of probability density functions (PDFs) of hydrological models [*Hoetting et al.,* 1999]. Currently, the technique has been applied to multiple hydrological model simulations as averaging schemes and weather prediction models to create forecast ensembles [*Raftery et al.,* 2005; *Wöhling and Vrugt,* 2008; *Duan and Phillips,* 2010; *Wu et al.,* 2012]. BMA usually estimates a representative weight (a single set of weights) for individual PDF of each model over the training period and then in turn aggregates different model predictions based on the estimated weights indicating how each model contributes to the predictive skill [*Ajami et al.,* 2007; *Rojas et al.,* 2008; *Tsai and Li,* 2008; *Zhang et al.,* 2009]. However, the weighted values can vary in the model performances during the training period because some hydrological models predict better outputs during the rainy period, while others perform better under the (relatively) dry condition [*Radell and Rowe,* 2008; *Hsu et al.,* 2009]. In order to overcome these limitations, recent studies adopted the sliding window algorithm to obtain the weights of individual models optimally [*Raftery et al.,* 2005; *Vrugt and Robinson,* 2007]. This approach assigns different weights to the models as the window slides over the training period. However, the strengths of hydrological models may not be adequately reflected by assigning different weights to the models during the training period across time. Thus, *Duan et al.* [2007] improved the BMA scheme for stream flow predictions using an alternative way that adopts multiple sets of weights to consider

different portions of the hydrograph instead of time-based weighting schemes. None of the previous studies, however, considered an approach of soil-wetness-based weighting scheme. Such a scheme may be more suitable for identifying soil moisture variability because soil moisture predictions from different hydrological models vary based on antecedent land surface wetness conditions (i.e., wet, moderately wet, and dry conditions). In this study, we explored a multiple-model simulation approach for estimating effective surface soil moisture dynamics (0–5 cm) and quantifying uncertainties due to different model parameters and structures. The objectives of this study are twofold: (1) to develop a BMA-based multimodel simulation approach based on the land surface wetness conditions in estimating effective soil moisture dynamics with a modified-microGA (Genetic Algorithm) for soil hydraulic parameter optimization and (2) to evaluate different model parameters and structural uncertainties under different hydroclimatic conditions.

2.3 Methodology

2.3.1. Bayesian Model Averaging Based Multimodel Simulation Approach

We developed a multimodel simulation approach adapting various hydrological models based on a Bayesian Model Averaging (BMA) scheme for estimating effective surface (0–5 cm) soil moisture dynamics and quantifying uncertainties due to different model parameterizations and structures. Figure 2.1 shows the schematic diagram of our proposed approach. In this study, we adapted three different hydrological models (i.e., Noah Land Surface Model, Noah LSM; Community Land Model, CLM; and Soil-Water-

Atmosphere-Plant, SWAP) for estimating surface soil moisture dynamics reflecting their inherent strengths. Noah LSM and CLM have been used extensively in evaluating water/energy cycles and fluxes including soil moisture prediction near the land surface as the dynamic land surface component of global climate modeling (e.g., Community Earth System Model and Weather Research and Forecasting), and SWAP also has been verified and used widely for predicting crop yields and soil moisture status in various studies [*Oleson et al.,* 2008; *Hong et al.,* 2009; *Shin et al.*, 2012]. A modified-microGA was integrated with these models for searching optimized parameters of each hydrological model and quantifying the model parameter uncertainty. To quantify the model structural uncertainty, we employed the BMA scheme calculating different weights of simulated results based on output fitness values of individual models. The multimodel simulation approach based on the BMA scheme was evaluated under two different hydroclimatic conditions.

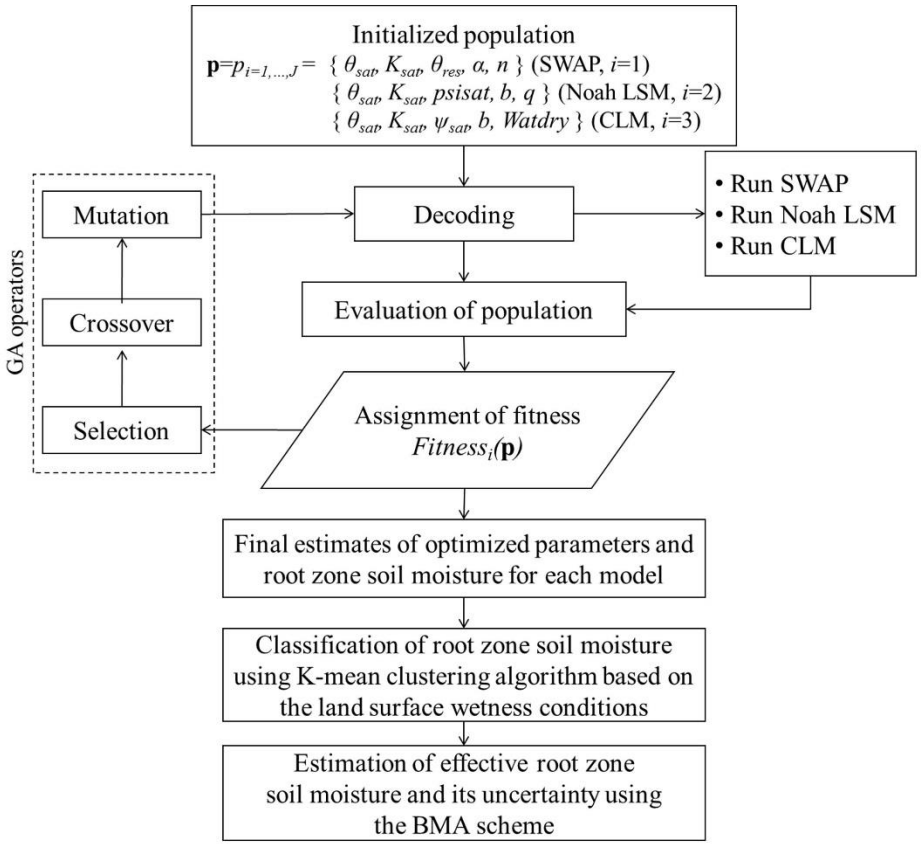

Figure 2. 1 Schematic diagram of the Bayesian Model Average (BMA) based multi-model simulation approach.

2.3.2. Characteristics of the Hydrological Models

2.3.2.1. Noah Land Surface Model

The original Noah Land Surface Model (Noah LSM v2.7.1) is an advanced version of the Oregon State University land model [*Ek et al.,* 2003]. This model has been widely used in both coupled (integrated with other models) and uncoupled (stand-alone) modes for simulating water and energy fluxes at various spatial scales. In this study, we adapted the uncoupled mode as a one-dimensional (1-D), physically based model for

estimating the soil moisture dynamics at field scales. Noah LSM calculates the total evapotranspiration by summing the direct evaporation from top soil layer, canopy evaporation, and potential Penman-Monteith transpiration [*Rosero et al.,* 2010]. The model has typically four soil layers with the thicknesses of 10, 30, 60, and 100 cm (total soil depth of 200 cm), but we changed top soil layer depth to 5 cm (while maintaining the same total root zone depth) to be compared to the soil moisture observation (top 5 cm) in this study. It adapts a diffusion form of the Richards' equation (equation (2.1)) for soil moisture estimation. Soil water retention and hydraulic conductivity and are calculated based on the *Clapp and Hornberger* [1978] equations (equations (2.2) and (2.3)),

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}(D(\theta)\frac{\partial \theta}{\partial z}) + \frac{\partial K(\theta)}{\partial z} + Q \tag{2.1}$$

$$\psi = \psi_{sat}\left(\frac{\theta}{\theta_{sat}}\right)^{-b} \tag{2.2}$$

$$K(\theta) = K_{sat}\left(\frac{\theta}{\theta_{sat}}\right)^{2b+3} \tag{2.3}$$

where θ is the volumetric soil water content ($cm^3\ cm^{-3}$), z is the soil depth (cm) taken positive upward, $D(\theta)$ is the soil water diffusivity ($cm^2\ d^{-1}$) ($K(\theta)\frac{\partial \psi}{\partial \theta}$), $K(\theta)$ is the unsaturated hydraulic conductivity ($cm\ d^{-1}$), Q is a soil moisture sink term, which is the root water extraction rate by plants, ψ and ψ_{sat} are the soil matric potential and saturated

soil matric potential (cm), *b* is the curve fitting parameter related to the pore size distribution (-), and θ_{sat} and K_{sat} are the saturated soil moisture content ($cm^3\ cm^{-3}$) and saturated hydraulic conductivity ($cm\ d^{-1}$), respectively.

Noah LSM has been enhanced to achieve better performance by incorporating complex canopy resistance, bare soil evaporation, surface runoff, and higher-order time integration schemes. Additional model processes and assumptions are provided in Table 2.1 and by *Ek et al.* [2003]. The model has been tested and showed good performance in humid and temperate hydroclimate regions [*Koren et al.,* 1999; *Sridhar et al.,* 2002; *Ek et al.,* 2003]. However, it still has limitations when applied to arid hydroclimate regions. The limitations might be caused by the assumption that latent heat flux associated strongly with evaporation and the distribution of soil moisture content is negligible in the top soil layer when the water content is lower than the wilting point (drying season) [*Katata et al.,* 2007]. Also, the thickness of top soil layer (10 cm as a default) is thicker than those of other models (i.e., SWAP and CLM), which can lead to overestimations of soil moisture [*Sahoo et al.,* 2008].

Table 2. 1 Comparison of main characteristics of the three hydrological models.

	Noah LSM	SWAP	CLM
Default thickness of top soil layer	10 cm (Total 4 layers)	1 cm (Total 10 layers with small compartments)	1.75 cm (Total 10 layers)
Runoff scheme	Simple Water Balance (SWB) model	Horton and Dunne Overland flow	TOPMODEL
Soil hydraulic properties	Clapp and Hornberger	van Genuchten and Mualem	Clapp and Hornberger
Surface evaporation	Penman potential evaporation	Penman-Monteith	Philip and De Vries diffusion model and BATS model
Plant system	Canopy resistance function	Linear production function and WOFOST model	Dynamic global vegetation model
Bottom boundary condition	Free drainage	Free drainage	Dynamic groundwater table (SIMGM)

2.3.2.2 Soil-Water-Atmosphere-Plant (SWAP) Model

Soil-Water-Atmosphere-Plant (SWAP) [*Van Dam et al.,* 1997] has been used for simulating soil water flow between the soil, water, atmosphere, and plant systems [*Agnese et al.*, 2007; *Ying et al.*, 2011]. This model contains physical processes for soil water flow, potential and actual evapotranspiration, crop growth, and irrigation. Daily potential evapotranspiration is estimated using the Penman-Monteith method with daily meteorological data or crop factors (i.e., minimum resistance, leaf area index, and crop height), and the actual evapotranspiration rate can be calculated using the root water uptake reduction and maximum soil evaporation flux [*Van Dam et al.,* 1997] (Table 2.1). This model simulates soil moisture dynamics in the soil profile using the mixed form Richards′ equation (equation (2.4)) and the soil hydraulic properties represented by the

analytical expression of Mualem and van Genuchten (equations (2.5) and (2.6)) [*Mualem*, 1976; *Van Genuchten*, 1980],

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K(\psi)(\frac{\partial \psi}{\partial z} + 1) \right] - Q \tag{2.4}$$

$$S_e = \frac{\theta(\psi) - \theta_{res}}{\theta_{sat} - \theta_{res}} = \left[\frac{1}{1 + |\alpha \psi|^n} \right]^m \tag{2.5}$$

$$K(\psi) = K_{sat} S_e^{\lambda} \left[1 - (1 - S_e^{1/m})^m \right]^2 \tag{2.6}$$

where n (-), m (-), λ (-), and α (cm^{-1}) are the empirical shape factors of the retention and conductivity functions, m=1-1/n, S_e is the relative saturation (-), θ_{res} is the residual water content ($cm^3\ cm^{-3}$), and $\theta(\psi)$ and $K(\psi)$ are the water content ($cm^3\ cm^{-3}$) and hydraulic conductivity ($cm\ d^{-1}$) at matric potential ψ, respectively.

SWAP simulates water flow, solute transport, heat flow, and crop growth simultaneously at field scales. In order to better simulate infiltration and evaporation fluxes in the vertical soil column, the soil profile (0–200 cm) was discretized in this study with finer intervals (1, 5, and 10 cm for the 1^{st}–10^{th}, 11^{th}–20^{th}, and 21^{st}–32^{nd} layers, respectively, except of 20 cm for the 33rd layer), especially for the soil surface (0–10 cm) where water content and pressure head gradients change sharply [*Van Dam et al.*, 1997]. However, a key limitation of the SWAP model is that it does not consider the

regional groundwater hydrology and seasonal variation of boundary fluxes at the lower boundary [*Kroes et al.,* 1998]. For the detailed information about SWAP readers can refer to *Van Dam et al.* [1997].

2.3.2.3 Community Land Model

Community Land Model (CLM) [*Oleson et al.,* 2010] is a land surface model that provides land surface forcing as the physical boundary for the atmospheric model in the Community Climate System Model. This model estimates bare soil evaporation based on the *Philip and de Vries* [1957] diffusion model and calculates transpiration using an aerodynamic approach of the Biosphere Atmosphere Transfer Scheme (BATS) model [*Dickinson et al.,* 1993]. Other model processes are provided in Table 2.1. CLM has a 10 layered soil column with the fixed thickness of 1.75, 2.76, 4.55, 7.5, 12.36, 20.38, 33.60, 55.39, 91.33, and 113.7 cm (total depth of 343 cm), and in this study averaged soil water content of the first two soil layers are used for comparison with the observations. The vertical soil water flow is solved by the modified Richards′ equation (equation (2.7)) [*Zeng and Decker,* 2009]. This equation is derived by subtracting the hydrostatic equilibrium soil moisture distribution from the original Richards′ equation for improving the mass conservative numerical scheme when the water table is within the soil column,

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K\left(\frac{\partial(\psi - \psi_e)}{\partial z} \right) \right] - Q \qquad (2.7)$$

where ψ_e is the equilibrium soil matric potential (cm). The hydraulic conductivity is derived from Eq. (2.3), and equilibrium soil matric potential and equilibrium volumetric water content are shown in Eqs. (2.8 and 2.9) based on *Clapp and Hornberger* [1978],

$$\psi_e = \psi_{sat}\left(\frac{\theta_e(z)}{\theta_{sat}}\right)^{-b} \tag{2.8}$$

$$\theta_e(z) = \theta_{sat}\left(\frac{\psi_{sat} + z_\nabla - z}{\psi_{sat}}\right)^{-\frac{1}{b}} \tag{2.9}$$

where $\theta_e(z)$ is the equilibrium (e) volumetric water content ($cm^3\ cm^{-3}$) at depth z (z_∇ is the water table depth).

In CLM, 10 soil layers discretized unevenly include a thin top soil layer (1.75 cm) needed to better simulate infiltration and evaporation fluxes [*Sahoo et al.*, 2008]. Furthermore, CLM considers the variability in ground water table as the lower boundary condition using the SIMple Groundwater Model (SIMGM) [*Niu et al.*, 2007]. A groundwater component is defined as an unconfined aquifer below the soil column (343 cm). To obtain the water table depth, the model parameterizes groundwater discharge and recharge with various constants derived from sensitive analysis [*Niu et al.*, 2007]. On the other hand, the model assumes that runoff generation is controlled by saturation area derived from topographic information and its parameterization is based on an exponential form, which is obtained from observations of the upper soil layers over

small watersheds. However, this runoff generation could be also driven by the relationship between rainfall intensity and soil infiltration capacity, especially in regions with thick soils or deep groundwater. The assumption of dominant topographic control might lead to erroneous simulations for the subsurface runoff [*Li et al.,* 2011].

2.3.2.4 Soil Parameters of the Hydrological Models

Parameter optimization using a modified-microGA was implemented to identify the soil hydraulic properties as unknown parameters whose variation has large effects on the model outputs [*Musters et al.,* 2000; *Hupet et al.,* 2002; *Ritter et al.,* 2003]. Several major input parameters related to soil moisture dynamics were selected as shown in Table 2.2 (Noah LSM - θ_{sat}, K_{sat}, *psisat* ($\partial\psi_{sat}/\partial z$), *b*, *q;* SWAP - θ_{sat}, K_{sat}, θ_{res}, α, *n;* CLM - θ_{sat}, K_{sat}, ψ_{sat}, *b, WATDRY*). Feasible ranges of the parameters (i.e., search spaces in a modified-microGA) for each model were defined based on literature related to the model parameter sensitivity and to accommodate a diversity of soils ranging from clay to sandy loam [*Leij et al.,* 1999; *Liu et al.,* 2004; *Ines and Mohanty,* 2008; *Rosero et al.,* 2010; *Shin, et al.,* 2012].

Table 2. 2 Summary of soil hydraulic parameters and its feasible ranges used in the modified-*micro*GA for the three hydrological models.

LSMs	Parameters ($\mathbf{p}=p_{i=1,\ldots,I}$)	Descriptions	Unit	Min.[a]	Max.[a]	No. of bits	Binaries (2^L)
SWAP ($i=1$)	θ_{sat}	Saturated water contents	$cm^3\ cm^{-3}$	0.37	0.55	5	32
	K_{sat}	Saturated hydraulic- conductivity	$cm\ d^{-1}$	1.84	55.70	10	1,024
	θ_{res}	Residual water contents	$cm^3\ cm^{-3}$	0.06	0.16	7	128
	α	Empirical shape parameter	cm^{-1}	0.01	0.03	5	32
	n	Empirical shape parameter	-	1.20	1.61	6	64
Noah LSM ($i=2$)	θ_{sat}	Saturated water contents	$cm^3\ cm^{-3}$	0.35	0.55	5	32
	K_{sat}	Saturated soil hydraulic-conductivity	$cm\ d^{-1}$	8.64	86.4	6	64
	psisat	Saturated soil matric potential ($\partial\psi_{sat}/\partial z$)	$cm\ cm^{-1}$	0.10	0.65	6	64
	b	Clapp-Hornberger b parameter	-	4.00	10.00	6	64
	q	Quartz content	-	0.10	0.82	5	32
CLM ($i=3$)	θ_{sat}	Saturated water contents	$cm^3\ cm^{-3}$	0.33	0.66	5	32
	K_{sat}	Saturated soil hydraulic-conductivity	$cm\ d^{-1}$	0.09	864	8	256
	ψ_{sat}	Saturated soil matric potential	cm	-75.00	-3.00	7	128
	b	Clapp-Hornberger b parameter	-	3.00	10.00	6	64
	WATDRY	Soil water content (wilting point)	-	0.02	0.30	5	32

[a]Feasible ranges of the parameters [*Liu et al.*, 2004; *Ines and Mohanty*, 2008; *Rosero et al.*, 2010].

2.3.3 Modified-microGA (Genetic Algorithms) for Estimating Optimal Parameters and Their Uncertainty

GAs are powerful algorithms based on the mechanics of nature (i.e., survival of the fittest mechanism) for searching optimal solutions from the unknown space [*Holland*, 1975]. GAs are basically composed of the GA operators such as selection, crossover, and mutation. New GA algorithms have been developed to improve the searching ability and save the computational time. *Krishnakumar* [1989] developed the so-called microGA to allow more micro population restarts in order to overcome the relatively

poor exploitation characteristic of the original GA. The micro population restarts searching solutions at the search space when most of the new parameter sets through the GA operator in a generation are similar up to 90%. It allowed that the GA can find solutions more efficiently saving the computational time. *Ines and Droogers* [2002] modified the microGA (i.e., modified-microGA) to consider interjecting new genetic materials to the micro population adopting a creep mutation. The creep mutation operator suggested by *Carroll* [1998] alters the parameter sets to minimize the effect of perturbation included in the converged parameter sets. In this study, we used a modified-microGA [*Ines and Droogers*, 2002] in searching the optimized soil parameters for the three selected hydrological models. The modified-microGA was integrated with the hydrological models to optimize each model input parameter sets, p = $\{p_{i=1,\ldots,J}\}$, as shown in Figure 2.1 based on the inversion model. The number of bits and binary used in the modified-microGA were decided by the degree of discrete divisions between the minimum and maximum values for each parameter range (Table 2.2). The objective (Z(p)) functions were formulated in equation (2.10),

$$Objective(Z(\mathbf{p})) = Minimize\left(\frac{1}{T}\sum_{t=1}^{T}\left|\theta_{i,t}^{sim} - \theta_{t}^{obs}\right|\right) \forall i \qquad (2.10)$$

where $\theta_{i,t}^{sim}$ and θ_{t}^{obs} are the simulated and observed surface soil moisture, respectively.

For the parameter uncertainty analysis, we used the multipopulation with different random number seeds (-1000, -950, and -750) in the modified-microGA

process. After the given generations, the individual and average fitness of all the parameter sets (i.e., chromosomes) from the multiple populations were calculated. The parameter sets which have the fitness values above the average were then selected as the probable solutions. Further, we carried out the perturbation analysis in order to account for the variations of the model parameters estimating the approximated solutions (**p′**) for each parameter set (**p**).

The perturbation analysis has been used to evaluate how variations of the model input parameters affect model outputs [*de Kroon et al.*, 1986; *Caswell*, 2000; *Benke et al.*, 2008]. The perturbed parameters were calculated as

$$\mathbf{p}' = \mathbf{p}_{Avg} \times (1 \pm x_i \xi) \tag{2.11}$$

where **p′** and $\mathbf{p}_{Avg}$ are the perturbed and averaged parameter set, $x_i \sim Norm(0, \sigma_i^2)$ is the normal random deviate with the mean and standard deviation calculated by the parameter sets selected (above the average fitness), ξ is the error term related to uncertain parameter (30% was applied in this study to account for the variations of the model parameters estimating the approximated solutions for each parameter sets).

The surface soil moisture dynamics were simulated using the perturbed parameters, and their uncertainties with the ±95% confidence interval (PCI) were evaluated for each model.

2.3.4. Bayesian Model Averaging Scheme Based on the Land Surface Wetness Conditions and Model Structural Uncertainty

The BMA scheme estimates weights for various model predictions based on their probabilistic likelihood measures [*Raftery et al.,* 2005]. Here, the variable '*y*' indicates the BMA prediction, namely predictive (weighted) soil moisture, and $f_{i=1,\ldots,J}$ is the individual model prediction (surface soil moisture dynamics) from the selected hydrological models (i=1,…,J) using the optimized parameters (section 2.3.3). The BMA posterior distribution of *y* given the model predictions can be formulated in Eq. (2.12) as following,

$$P(y \mid f_1, \cdots, f_i) = \sum_{i=1}^{J} P_i(f_i \mid D) P_i(y \mid f_i, D) \tag{2.12}$$

where the PDF, ($P_i(f_i \mid D)$) is the posterior probability for model prediction given the training data (i.e., observations, *D*) and can be dealt with as weights (a single set of weights, w_i) defining the individual model's relative contributions to the BMA prediction, and *J* is the number of hydrological models used (i.e., 3). The conditional PDF, ($P_i(y \mid f_i, D)$) denotes the posterior distributions of *y* given model prediction and observations, which is approximated by a normal distribution with mean ($\overline{f_i}$) and standard deviation (σ_i). The assumption of normal distribution could be inappropriate for soil moisture primarily driven by precipitation, while the gamma distribution is more

reasonable to represent the highly skewed predictive distribution of soil moisture (*Sloughter et al.*, 2006). However, when we tested the two assumptions (normal and gamma distribution), the assumption of normality improved more the BMA method for soil moisture prediction. In the study of *Vrugt and Robinson* (2007), they also found an improvement of BMA method with the assumption of normal distribution for streamflow forecasting instead of the gamma distribution. The posterior mean (*E*) and variance (*Var*) of the BMA prediction (*y*) can be computed in Eqs. (2.13 and 2.14).

$$E\left[y \mid f_{i=1,\ldots,J} \right] = E\left[P(y \mid f_{i=1,\ldots,J}) \right] = E\left[\sum_{i=1}^{J} w_i P_i \ (y \mid f_i \ ) \right] = \sum_{i=1}^{J} w_i \ f_i \tag{2.13}$$

$$Var\left[y \mid f_{i=1,\ldots,J} \right] = \sum_{i=1}^{J} w_i \left[\ f_i \ - \sum_{l=1}^{J} w_l \ f_l \ \right]^2 + \sum_{i=1}^{J} w_i \ \sigma_i^2 \tag{2.14}$$

The BMA approach then estimates the weights and variances of simulated surface soil moisture dynamics from the three models. The variance of BMA prediction consists of the between-model variance and the within-model error variance in the BMA procedure. The values of w_i and σ^2 were estimated by the maximum likelihood (*L*) as described in Eq. (2.15),

$$L(w_{i=1,\ldots,J}, \sigma^2 \mid f_{i=1,\ldots,J}, y) = \sum_{t=1}^{T} \log\left(\sum_{i=1}^{J} w_i P_i \ (y_t \mid f_{i,t}) \right) \tag{2.15}$$

where *T* is the time domain. To find the maximum likelihood for the weights and variances, we used the DiffeRential Evolution Adaptive Metropolis – Markov Chain Monte Carlo (DREAM-MCMC) algorithm [*Vrugt et al.,* 2009]. The BMA weights are highly correlated with the model performance indicating that higher weights are assigned to a model which performed better than others. This algorithm has been used for estimating the BMA parameters (weight and variance) and is unique in solving complex, multi-modal, and high-dimensional sampling problems [*Vrugt et al.,* 2008, 2009]. Thus, we estimated the weights (a single set of weights) for different hydrological models using the DREAM-MCMC algorithm and the effective surface soil moisture dynamics were calculated by aggregating the three model outputs based on the estimated weights. Hydrological models can predict the hydrologic response well during the dry or wet season based on their own model parameters and structures [*Hsu et al.*, 2009]. In order to reflect the strengths of individual models for certain land surface wetness conditions, we categorized soil moisture measurements based on the land surface wetness conditions (e.g., wet, moderately wet, and dry conditions, etc.) using the K-mean clustering algorithm [*MacQueen,* 1967]. Near surface soil moisture can involve several state variables of climate and physical properties (e.g., soil texture, vegetation cover, precipitation events, etc.) with respect to the wetness conditions so that the thresholds of wetness conditions can be identified using the measurements [*Narasimhan et al.*, 2005; *D'Odorico et al.,* 2007; *Brocca et al.,* 2008]. Thus, the thresholds based on the soil moisture measurements can be also applicable to other locations having similar soil type, land cover, and hydro-climatic characteristics. The clustering algorithm determines the

land surface wetness conditions based on the degree of variability between available soil moisture measurements (note that the number of wetness conditions, *G*, was manually determined). Different weights ($w_{i=1,\ldots,J}^{g=1,\ldots,G}$, multiple sets of weights) of model outputs corresponding to the land surface wetness conditions were calculated by the BMA scheme (Eqs. 2.12-2.15), respectively. The estimated weights were assigned to the individual model output, and then the weighted soil moisture simulations were aggregated to estimate the effective surface soil moisture dynamics reducing error due to the model structural uncertainties. In this study, we evaluated the performance of BMA scheme using a single (S-BMA) and multiple (M-BMA) sets of weights in modeling. Then, we quantified the model structural uncertainties with the ±95 PCI estimated from the posterior distribution of the BMA parameters (i.e., weight and variance).

2.3.5. Statistical Analysis

Three performance criteria were selected to evaluate the performance of individual model predictions and of the multiple model simulation compared to the observations. They are Pearson's correlation (*R*), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) as Eqs. (2.16, 2.17, and 2.18),

$$R = \frac{\sum_{t=1}^{T}(\theta_t^{sim} - \overline{\theta}_t^{sim})(\theta_t^{obs} - \overline{\theta}_t^{obs})}{\sqrt{\sum_{t=1}^{T}(\theta_t^{sim} - \overline{\theta}_t^{sim})^2 \sum_{t=1}^{T}(\theta_t^{obs} - \overline{\theta}_t^{obs})^2}} \tag{2.16}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{T}(\theta_t^{obs} - \theta_t^{sim})^2}{T}} \tag{2.17}$$

$$\text{MAE} = \frac{1}{T}\sum_{t=1}^{T}\left|\theta_t^{obs} - \theta_t^{sim}\right| \tag{2.18}$$

where $\overline{\theta}_t^{sim}$ and $\overline{\theta}_t^{obs}$ are the average of θ_t^{sim} and θ_t^{obs}, respectively.

2.3.6 Study Area and Description of Model Conditions

In this study, the Little Washita (LW 13) site in Oklahoma (sub-humid) and Walnut Gulch (WG 82) site in Arizona (semi-arid) were selected for evaluating the model parameters and structural uncertainties under two different hydro-climatic conditions (Figure 2.2). The LW 13 site has a sub-humid climate with an average annual rainfall of approximately 750 mm with most precipitation during Spring and Fall. Daily-mean maximum temperature is 30˚C in July with annual-mean temperature of 16˚C. The climate of WG 82 is semi-arid with an average annual rainfall of approximately 350 mm, which is mostly received from July to September. Daily-mean maximum temperature of 35˚C occurs in June with annual-mean temperature of 17.7˚C. Both study sites have a native grass cover, and their soil types are silty loam and sandy loam for LW 13 and WG 82, respectively. The three hydrological models require common weather data (i.e., precipitation, temperature, relative humidity, solar radiation, and wind speed) and Noah LSM and CLM additionally need the air pressure values. They were collected from the

USDA Agricultural Research Service (ARS 136) Micronet and the Oklahoma Mesonet weather stations (Ninnekah station) from January 1 to December 31, 1997 for the LW 13 site. The weather datasets for the WG 82 site were obtained from the Arizona Meteorological Network [*Keefer et al.,* 2009] and the Soil Climate Analysis Network (SCAN, Walnut Gulch #1) sites from January 1 to December 31, 2004.

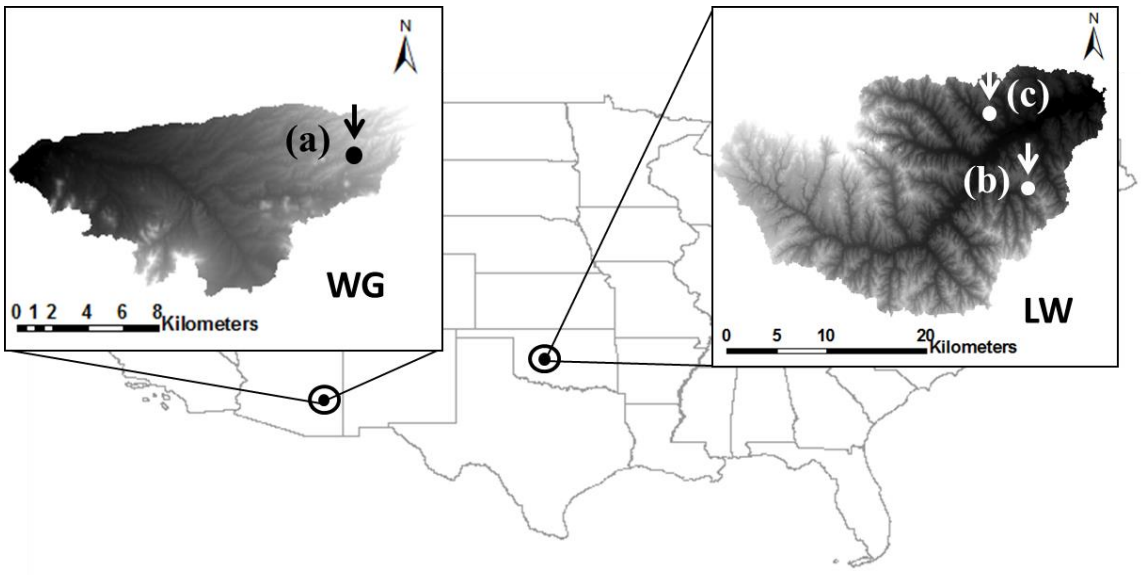

Figure 2. 2 Study sites; (a) Walnut Gulch (WG) 82 in Arizona, (b) Little Washita (LW) 13, and (c) SCAN 2023 in Oklahoma.

We validated our approach with the *in situ* soil moisture measurements (0-5 cm) during the Southern Great Plains experiment 1997 (SGP97, DOY: 170 ~ 197) [*Mohanty and Skaggs,* 2001; *Mohanty et al.,* 2002] for the LW 13 site and Soil Moisture Experiment 2004 (SMEX04, DOY: 216 ~ 238) [*Jackson et al.,* 2009] for the WG 82 site. Here, we calibrated the multi-model approach using the measurements during the

simulation periods (DOY: 170 ~ 183 for LW 13 and DOY: 216 ~ 227 for WG 82), and the validations were conducted in the given periods (DOY: 184 ~ 197 for LW 13 and DOY: 228 ~ 238 for WG 82), respectively. These experiment data sets have been validated significantly and used widely in various studies, but the experiment periods are limited. Thus, we also tested our approach using longer data sets (April 1 to December 31, 2011) measured at USDA-SCAN 2023 site (Figure 2.2) in Little Washita watershed, in the close proximity of LW13. The site is close to the LW 13 site having the same climate condition (sub-humid) and has a grass cover and silty clay soil. The weather data sets were collected from the SCAN 2023 site from January 1 to December 31, 2011.

In order to reflect the impacts of various land surface wetness conditions in modeling as mentioned above, *in situ* measurements were categorized using the K-mean clustering algorithm at the LW 13 and WG 82 sites. Thresholds of the clustering ranges could be different with site conditions such as hydro-climates due to the different climate forcing and hydrologic responses which can influence the model performance. In order to determine the appropriate range of weight sets we tested several different weight sets (e.g., 2, 3, and 4 sets) clustered using K-mean algorithm for each site. Comparing the BMA predictions of each set to the measurements including at least 5 data in each class, we found the suitable sets of weights representing the highest correlation and reflecting the models' characteristics properly discussed in section 2.3.2. The *in situ* data was then clustered into the three (G=3: wet, moderately wet, and dry conditions) and two (G=2: wet and dry conditions) classes for the LW (13 and SCAN 2023) and WG (82) sites, respectively.

Hydrological models have different initial and bottom boundary conditions due to their own structural characteristics. In the study sites, actual depth to the groundwater table was not available during the experiment periods, so we assumed that the bottom boundary condition is defined with free-drainage at 2 m depths from the soil surface for the Noah LSM and SWAP models, while the bottom boundary condition for CLM was decided with the water table dynamics calculated from aquifer water storage via the SIMGM [*Niu et al.*, 2007] after spinning up the model. For the initial condition, the Noah LSM and CLM models performed a spinning up to initialize the soil profile. Uniform initial soil water pressure head distributions ($h(z,t=0)$ = -100 and -500 cm for the LW (13 and SCAN 2023) and WG (82) sites indicating the shallow/deep groundwater levels, respectively) were used for the SWAP model.

2.4 Results and Discussion

2.4.1 Estimation of Optimized Model Parameters and Their Uncertainties

The optimized model parameters and the uncertainties of each model were estimated using the modified-*micro*GA under two different hydro-climatic regions. Figure 2.3 shows the probability distributions and their quantile box charts for the

estimated soil hydraulic parameters of each model using multiple random number seeds (i.e., -1000, -950, and -750) at the LW 13 site during the calibration period (DOY: 170 ~ 183, 1997). The estimated parameters for individual models showed the unimodal distributions indicating a probable optimized parameter value. Further, some of the parameters represented discontinuous distributions because the modified-*micro*GA searched for the possible parameter sets from the multi-population and different random number seeds exploring the complete search space. The optimized values for each model were used for evaluating the model parameter uncertainty, estimating the effective soil moisture dynamics for the study site. Based on these results, we found that the optimized soil hydraulic parameters (θ_{sat}, b, ψ_{sat}, and K_{sat}) and their ranges (i.e., search spaces) for the three models showed differences under the same modeling conditions (i.e., atmospheric forcings, soils, vegetations, etc.). The discrepancy between the models may be attributed to different parameterizations and structures that can also provide different model performances.

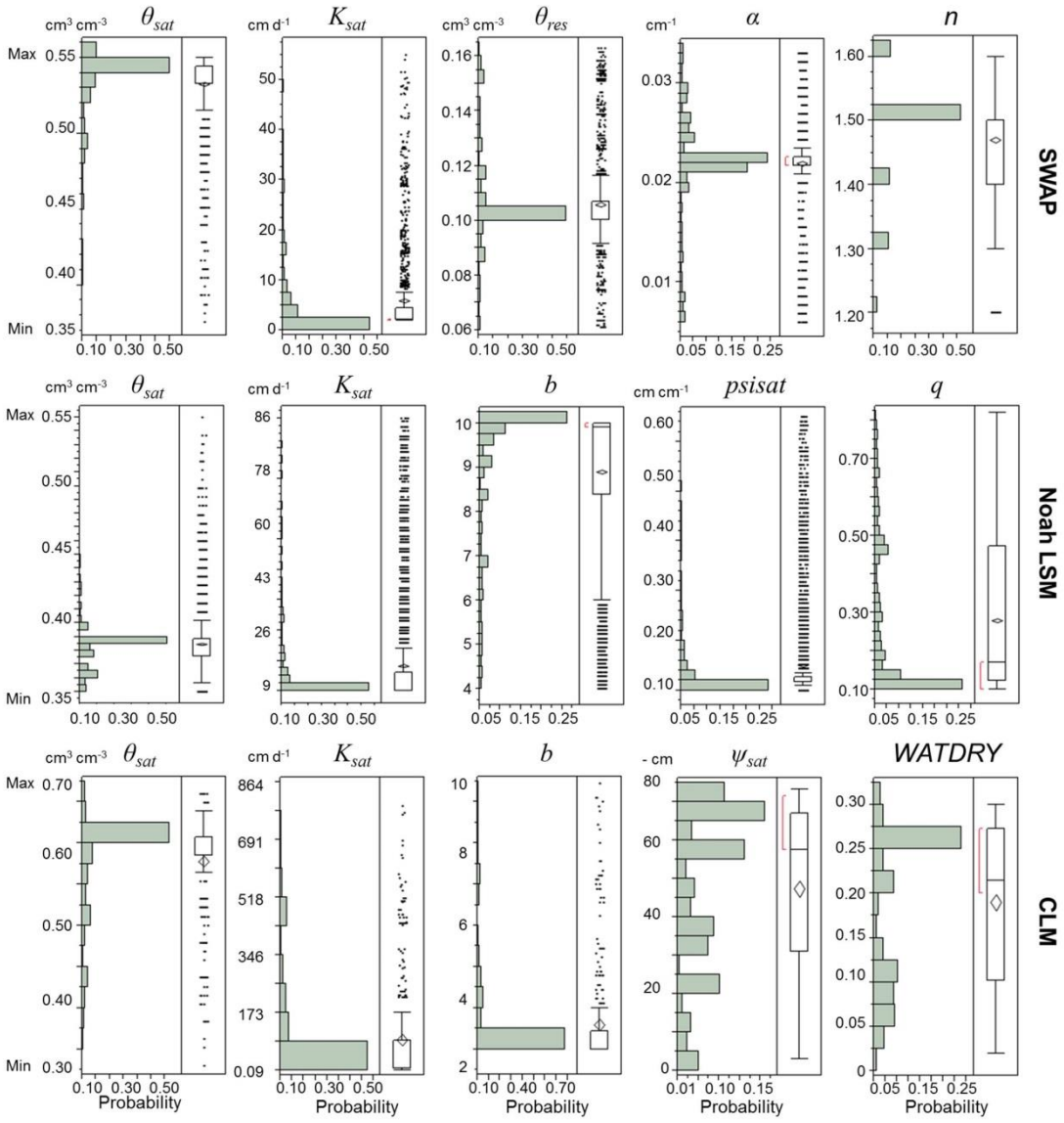

Figure 2. 3 Probability distributions and quantile box plots of the searched soil parameters of the three hydrological models using the multiple random number seeds (i.e., -1000, -950, and -750) for the LW 13 site.

In order to quantify parameter uncertainties of each model, we generated ten perturbed parameter ensembles using the statistics (mean and standard deviation) of estimated parameters based on the multiple populations and random number seeds. Then, the surface (0-5 cm) soil moisture dynamics were simulated using the perturbed parameter ensembles for each model in a forward mode. Figure 2.4a-c present the comparison of *in situ* and simulated surface soil moisture dynamics and their uncertainty band for SWAP, Noah LSM, and CLM during the calibration and validation periods. The results showed very narrow uncertainty boundaries because the possible parameter sets searched by the modified-*micro*GA using the different populations were very similar for the cases of SWAP and Noah LSM (Figures 2.4a and 2.4b). Some observations deviated from the narrow boundaries of the simulated soil moisture from SWAP and Noah LSM. It can be inferred that the single model could not predict properly for a certain period due to their model structural error.

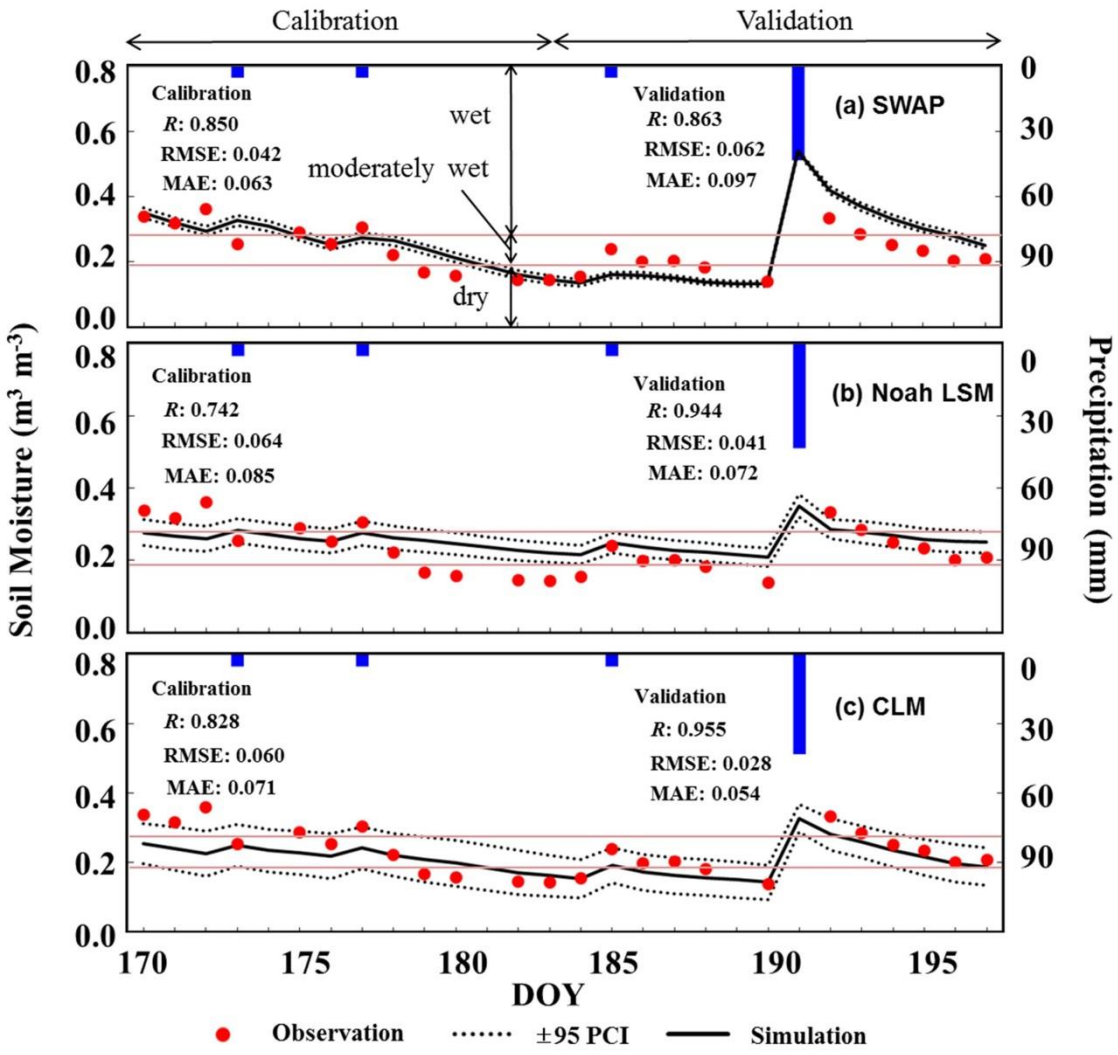

Figure 2. 4 *In situ* and simulated surface soil moisture (0~5 cm) dynamics using the optimized soil parameters derived by the modified-*micro*GA for (a) SWAP, (b) Noah LSM, and (c) CLM at the LW 13 site during calibration and validation periods.

Overall, however, three different models predicted the surface soil moisture dynamics well in comparison with the measurements (*R*: 0.742 ~ 0.850, RMSE: 0.042 ~ 0.064, and MAE: 0.063 ~ 0.085 during the calibration period; *R*: 0.863 ~ 0.955, RMSE: 0.028 ~ 0.062, MAE: 0.054 ~ 0.097 during the validation period). The SWAP model showed better performance than others at the LW 13 site during the calibration period, while CLM performed better during the validation period. On a closer view, the simulated surface soil moisture dynamics by SWAP matched well with the measurements during DOY 170 ~ 177 (volumetric water content above 0.280 $m^3 m^{-3}$), but the CLM results were identified better during DOY 177 ~ 190 (volumetric water content below 0.190 $m^3 m^{-3}$). Also, both Noah LSM and CLM performed well during DOY 192 ~ 197 (for volumetric water content 0.190 ~ 0.280 $m^3 m^{-3}$). The simulated surface soil moisture from SWAP was more sensitive to the precipitation which can be associated directly with the wet surface condition, compared to those of Noah-LSM and CLM, because of a thin top soil layer (1 cm) which can capture the dynamic change of surface soil moisture. Thus it showed rather good agreement with measurements than other models during wet condition (Figure 2.4a). In contrast, CLM showed poor performance during wet condition (Figure 2.4c). In CLM, the simulated surface soil moisture was underestimated due to the unreliable surface runoff generation and high

sensitivity of evaporation to precipitation. During the dry condition, CLM predicted the surface soil moisture well whereas Noah LSM somewhat overestimated the surface soil moisture. This may be attributed to the layer thickness of the models. The thicker top layer of Noah LSM holds more soil water after precipitation events than the thin soil layers of the other models (Figure 2.4b). These findings support those of *Hsu et al.* [2009] as they state that the performances of different models has their own strengths and weaknesses for certain processes, and we found that the performances of different hydrological models (Noah LSM, CLM, and SWAP) might vary based on the different land surface wetness conditions (e.g. wet, moderately wet, and dry conditions).

Figure 2.5 shows the probability distributions of estimated effective soil hydraulic parameters based on the multiple random number seeds for the WG 82 site during the calibration period (DOY: 216 ~ 227, 2004). Most of the probability distributions were unimodal for the Noah LSM and CLM parameters, except *q* (in Noah LSM) and *WATDRY* (in CLM) variables.

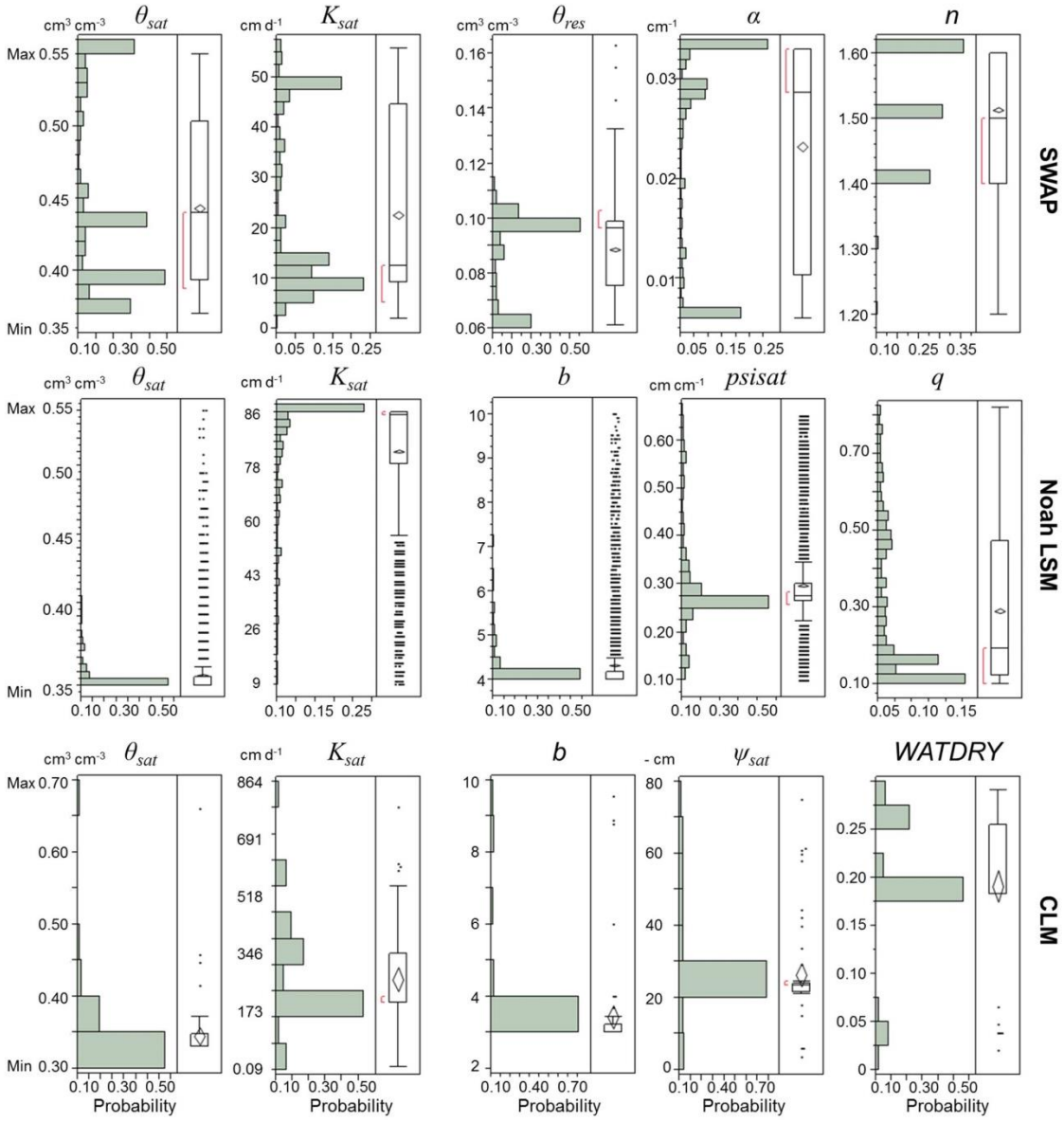

Figure 2. 5 Probability distributions and quantile box plots of the searched soil parameters of the three hydrological models using the multiple random number seeds (i.e., -1000, -950, and -750) for the WG 82 site.

However, the model parameter distributions of SWAP have multiple modes indicating local minima that can be derived by the modified-*micro*GA in the search space. Thus the local minima that significantly deviated from the referenced parameter ranges (UNSODA [*Leij et al.,* 1999], Soil Survey [*Wösten et al.,* 1994], Rosetta [*Schaap et al.,* 1999], and Clapp and Hornberger table [*Clapp and Hornberger,* 1978]) of the sandy loam soil type (predominant at the WG 82 site) were excluded. Also, we found a response time lag of 1 day between observed precipitation and simulated soil moisture that could be attributed to the difference of actual measurement time during the day and model time steps (starting at 12 midnight) at the WG 82 site. Figures 2.6a-c show the comparison of measured and simulated surface soil moisture dynamics with ±95 PCI after a 1-day lag was corrected. The simulated soil moisture dynamics from the three models agreed well with the measurements.

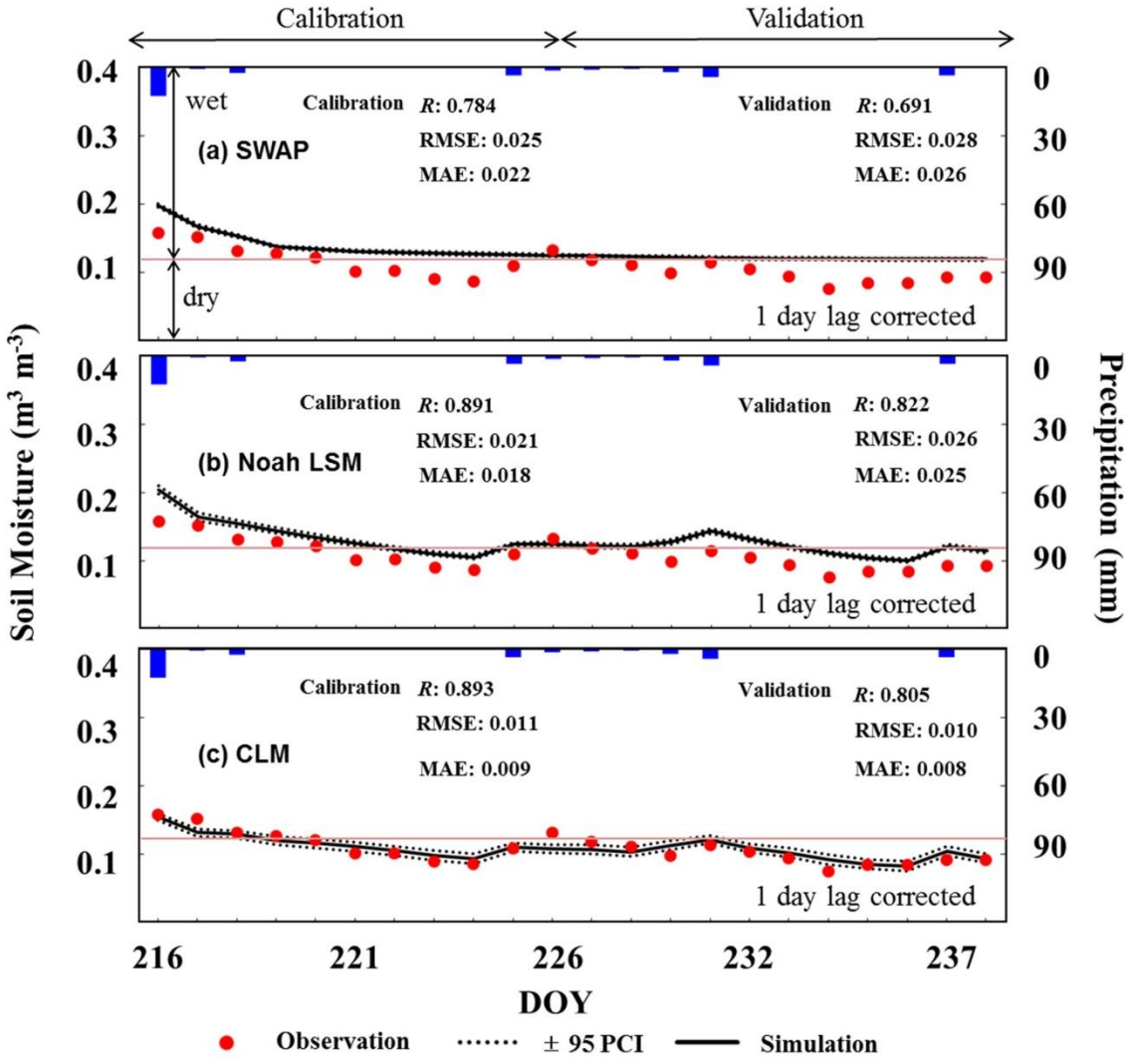

Figure 2. 6 *In situ* and simulated surface soil moisture (0~5 cm) dynamics using the optimized soil parameters derived by the modified-*micro*GA for (a) SWAP, (b) Noah LSM, and (c) CLM at the WG 82 site during calibration and validation periods.

Statistical analyses demonstrated that CLM performed better than others during the calibration and validation period as shown in the figures. The outputs of SWAP showed more uncertainties compared to the results of the other two models under the prevailing condition (e.g. relatively small precipitation and high solar radiation) as shown in Figure 2.6 (DOY: 222 ~ 238). SWAP tends to overestimate the surface soil moisture when the soil is relatively dry along with small precipitation and high evapotranspiration rate estimated using Penman-Monteith method [*Baroni and Tarantola,* 2012]. We also found that the SWAP results matched the measurements during DOY 216 ~ 221 (above 0.125 $m^3 m^{-3}$, wet condition) with higher correlation (R=0.945) than others (R=0.911 for Noah LSM and R=0.889 for CLM) at the WG 82, while the CLM model identified better during the period of DOY 222 ~ 238 (below 0.125 $m^3 m^{-3}$, dry condition). In general, CLM showed a good performance for this site considering the water table dynamics as a bottom boundary condition, but the model underestimated the surface soil moisture during wet conditions that can be associated to more moisture loss through evaporation. Noah LSM appeared to somewhat overestimate the surface soil moisture because of the thick top soil layer, but the model showed a similar tendency as CLM compared to the field observations (Figure 2.6b). Compared to the results at LW 13, the parameter uncertainty boundaries of each model were smaller, because of the low variability of surface soil moisture estimations. It may indicate that the relatively low rainfall amounts at the WG 82 site (semi-arid) cause the low surface soil moisture variability in modeling.

For the longer period simulation at SCAN 2023 site, the three models integrated with the modified-*micro*GA predicted the surface soil moisture well representing a good agreement with the measurements (*R*: 0.75, RMSE: 0.052, and MAE: 0.039 for SWAP; *R*: 0.89, RMSE: 0.033, and MAE: 0.023 for Noah LSM; *R*: 0.78, RMSE: 0.046, and MAE: 0.035 for CLM). Yet the predictions from the models indicated different trends under the different land surface wetness conditions defined with the same thresholds of the LW 13 site (Figure 2.7). SWAP shows good response to precipitation events in predicting the surface soil moisture peaks better than others during the wet condition, whereas the simulated surface soil moisture decreased rapidly during the dry-down phase (i.e., moderately wet and dry conditions) after the precipitation event. On the other hand, CLM and Noah LSM showed best performance in moderately wet and dry condition, respectively. Evaporation in CLM is very sensitive to the precipitation on short time scale (the case of LW 13) so that the evaporation removes soil water from the top soil layer. This is the reason why CLM predicted well the low surface soil moisture during dry condition at the LW 13 site. In contrast, on long time scale, more soil water can be retained from previous precipitation events that may cause the overestimation of surface soil moisture.

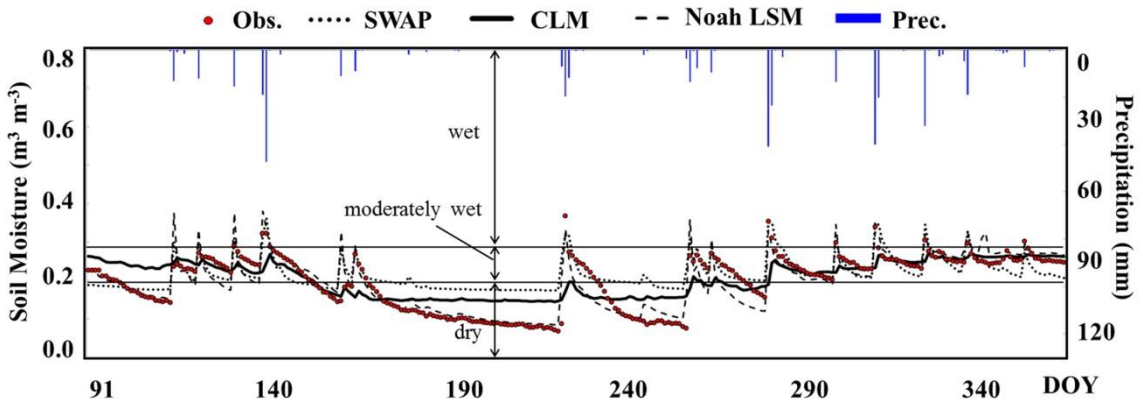

Figure 2. 7 *In situ* and simulated surface soil moisture (0~5 cm) dynamics using the optimized soil parameters derived by the modified-*micro*GA for SWAP, Noah LSM, and CLM at the SCAN 2023 site.

Overall, the predicted surface soil moisture dynamics using the three models based on the optimized parameters derived by the modified-*micro*GA matched well with the measurements in two different hydro-climatic regions. However, the measured soil moisture dynamics could not be captured adequately by the parameter uncertainty boundaries of SWAP and Noah LSM. Furthermore, the performances of different hydrological models in estimating the surface soil moisture showed different trends under various wetness conditions and different hydro-climatic conditions. It infers that uncertainties due to the different model structures are reflected significantly in model outputs.

2.4.2 Estimation of Effective Surface Soil Moisture and Its Uncertainty

In order to reduce bias due to model structural uncertainties (i.e., different model parameterizations, governing equations, etc.) mentioned above, we assigned a single (S-BMA) and multiple (M-BMA) sets of weights derived by the BMA scheme to the individual surface soil moisture predictions. A single set of weight ($w_{i=1,\ldots,J}$) was estimated based on the simulation results from the three models for the LW 13 site during the whole simulation period as shown in Table 2.3.

Table 2. 3 A single and multiple sets of the Bayesian Model Average (BMA) weights for the three hydrological models at the LW 13 site.

BMA set		Weights		
		SWAP (*i*=1)	Noah LSM (*i*=2)	CLM (*i*=3)
S-BMA[a] ($w_{i=1,\ldots,J}$)		0.291	0.005	0.704
M-BMA[b] ($w_{i=1,\ldots,J}^{g=1,\ldots,G}$)	*g*=1	0.533	0.466	0.001
	g=2	0.001	0.293	0.706
	g=3	0.008	0.002	0.990

S-BMA[a] means a single set of the weights for the three models (*i*=1,2,3).
M-BMA[b] means multiple sets of the weights corresponding to three land surface wetness conditions (*g*=1,2,3 represent the wet, moderately wet, and dry conditions, respectively).

Note that *G* and *J* are the number of land surface wetness conditions and hydrological models, respectively.

The highest weight (0.704) was assigned to CLM, which showed the best model performance (*R*: 0.837, RMSE: 0.047, MAE: 0.036) over the simulation period, while SWAP (*R*: 0.789, RMSE: 0.053, MAE: 0.044) and Noah LSM (*R*: 0.806, RMSE: 0.054, MAE: 0.046) had relatively lower weights of 0.291 and 0.005, respectively. The aggregated surface soil moisture dynamics using a single set of weights (R: 0.823, RMSE: 0.040, and MAE: 0.061) for the three models matched better with the measurements than SWAP and Noah LSM predictions in Figure 2.8.

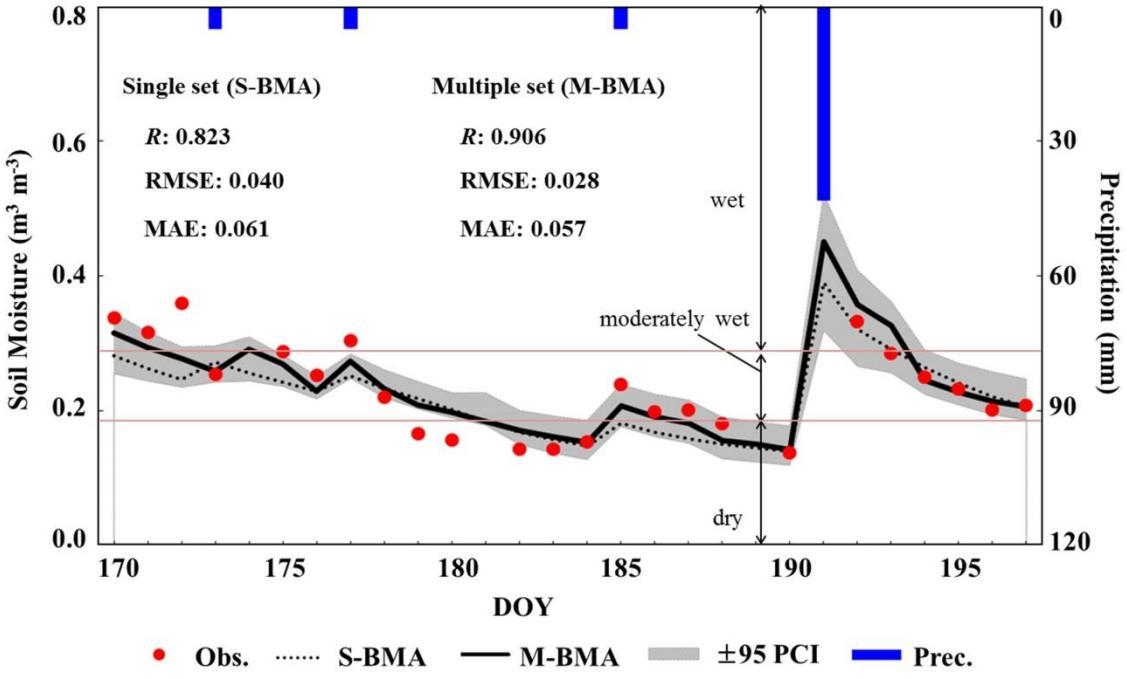

Figure 2. 8 *In situ* and simulated surface soil moisture using a single (S-BMA, dotted line) and multiple (M-BMA, black line) sets of the BMA weights and ±95 PCI at the LW 13 site.

However, there was no significant improvement of the single-weighted prediction compared to CLM prediction. This was because the single-weighted based surface soil moisture dynamics were considerably biased towards the CLM results assigned with the highest weight uniformly along the whole period and did not reflect a good performance of other models during a certain condition (e.g. wet and moderately wet). As shown in Figure 2.4, the SWAP model performed better during DOY 170 ~ 177 (defined as the wet condition), while Noah LSM and CLM predicted the surface soil moisture estimates better under the moderately wet and dry conditions, representing the advantages and disadvantages of each model structure. These findings demonstrated that we need to classify the simulation period for assigning different weights to the model predictions based on the land surface wetness conditions. For these reasons, we categorized the *in situ* measurements using the K-mean clustering algorithm as the wet (above 0.280 $m^3 m^{-3}$), moderately wet (0.190 ~ 0.280 $m^3 m^{-3}$), and dry (below 0.190 $m^3 m^{-3}$) conditions, respectively. Then, we estimated multiple sets of the weight (i.e. $w_{i=1,\ldots,J}^{g=1}$ -wet, $w_{i=1,\ldots,J}^{g=2}$ -moderately wet, and $w_{i=1,\ldots,J}^{g=3}$ -dry conditions) based on the categorized soil moisture measurements for the LW 13 site (see Table 2.3). The highest weight (0.533) was assigned to the SWAP results during the wet condition at the LW 13 site, while CLM had the highest weights (0.706 and 0.990) during the moderately wet and dry conditions, respectively. These multiple weight values can be seen as the performance of individual models based on the advantages of each model structure. The effective (multiple-weighted) surface soil moisture dynamics showed a better match with the measurements (*R*: 0.906, RMSE: 0.028, and MAE: 0.057) in Figure 2.8. Compared to

the single-weighted results (Figure 2.8), the BMA scheme based on the multiple sets of weight (based on wetness thresholds) also improved the surface soil moisture estimations and their uncertainties, especially on DOY 170-177. Thus, our findings demonstrated that the BMA-based multi-model simulation approach with multiple sets of weights is more suitable for addressing model structural uncertainties than those with a single set of weights.

Table 2. 4 A single and multiple sets of the Bayesian Model Average (BMA) weights for the three hydrological models at the WG 82 site.

BMA set		Weights		
		SWAP (i=1)	Noah LSM (i=2)	CLM (i=3)
S-BMA[a] ($w_{i=1,\ldots,J}$)		0.001	0.053	0.946
M-BMA[b]	g=1	0.936	0.001	0.063
($w_{i=1,\ldots,J}^{g=1,\ldots,G}$)	g=2	0.002	0.002	0.996

S-BMA[a] means a single set of the weights for the three models (i=1,2,3).
M-BMA[b] means multiple sets of the weights corresponding to two land surface wetness conditions (g=1,2 represent the wet and dry conditions, respectively).

Note that G and J are the number of land surface wetness conditions and hydrological models, respectively.

We estimated a single set of weights ($w_{i=1,\ldots,J}$) for the whole simulation period for the WG 82 site as shown in Table 2.4. The highest weight value (0.946) was assigned to the CLM results that show the best prediction (*R*: 0.856, RMSE: 0.014, and MAE: 0.011) and then in turn the low weights of 0.001 and 0.053 were assigned to SWAP and Noah LSM, respectively. The aggregated (single-weighted) surface soil moisture dynamics agreed with the measurements, but they were also biased to the CLM results representing that the predictions have uncertainties during the wet period (DOY 216 ~ 221, Figure 2.9) as shown in the results of LW 13 site.

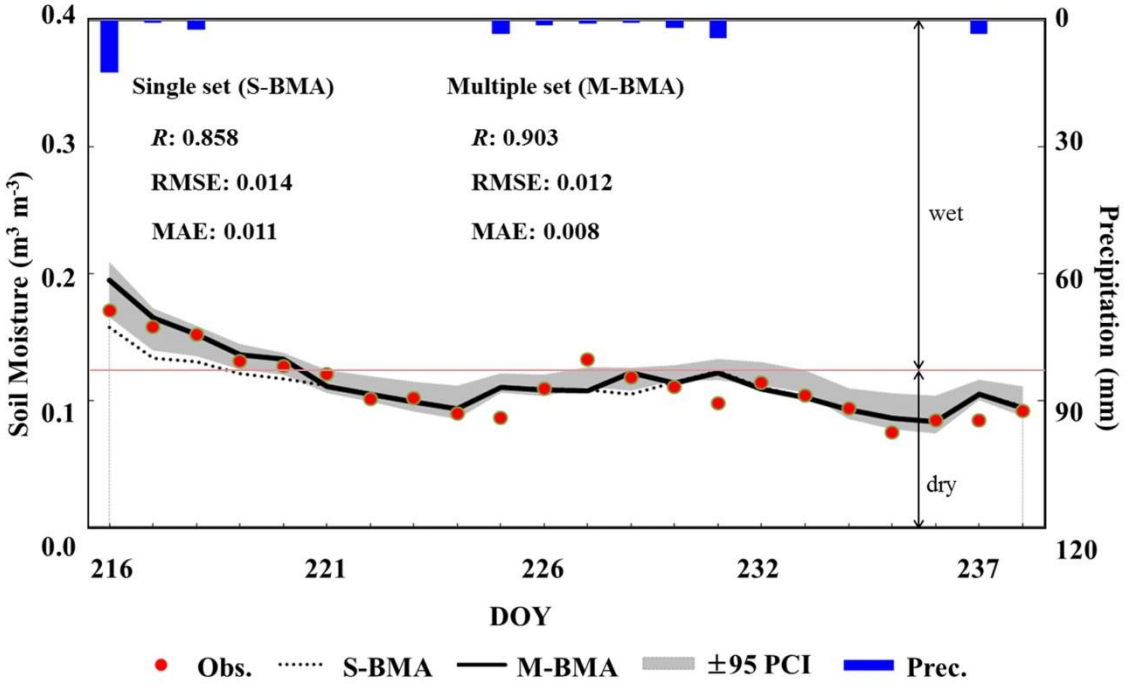

Figure 2. 9 *In situ* and simulated surface soil moisture using a single (S-BMA, dotted line) and multiple (M-BMA, black line) sets of the BMA weights and ±95 PCI at the WG 82 site.

Thus, we categorized the simulation period into the two classes (i.e. wet and dry conditions) and estimated the two sets of the weight ($w_{i=1,\ldots,J}^{g=1}$ -wet and $w_{i=1,\ldots,J}^{g=2}$ -dry conditions, see Table 2.4) for the WG 82 site. As shown in previous section, the simulated surface soil moisture dynamics from the SWAP model were closer to the measurements during the wet condition, while CLM performed better along the dry period. Thus the highest weight values (0.936 and 0.996 for the wet and dry conditions) were assigned to the results of SWAP and CLM models, respectively. The aggregated surface soil moisture dynamics using multiple sets of weights (*R*: 0.903, RMSE: 0.012, and MAE: 0.008) identified better with the measurements than the individual models and single-weighted results. Further, a poor performance due to the structural errors of single model could be compensated by good performances of other models indicating that the measured soil moisture data were mostly located within the ±95 PCI.

Table 2. 5 A single and multiple sets of the Bayesian Model Average (BMA) weights for the three hydrological models at the SCAN 2023 site.

BMA set		Weights		
		SWAP (i=1)	Noah LSM (i=2)	CLM (i=3)
S-BMA[a] ($w_{i=1,\ldots,J}$)		0.204	0.650	0.146
M-BMA[b] ($w_{i=1,\ldots,J}^{g=1,\ldots,G}$)	g=1	0.592	0.406	0.002
	g=2	0.432	0.125	0.443
	g=3	0.001	0.934	0.065

S-BMA[a] means a single set of the weights for the three models (i=1,2,3).
M-BMA[b] means multiple sets of the weights corresponding to three land surface wetness conditions (g=1,2,3 represent the wet, moderately wet, and dry conditions, respectively).

Note that G and J are the number of land surface wetness conditions and hydrological models, respectively.

We also tested our proposed approach using long period data (DOY 91-365) at SCAN 2023 site. The long period soil moisture measurements were categorized into the three classes (wet, moderately wet, and dry conditions) by the same range of the wetness conditions for LW 13 and multiple sets of weights were estimated using the BMA scheme (Table 2.5). The highest weights were assigned to SWAP (0.592 for wet condition), CLM (0.443 for moderately wet), and Noah LSM (0.934 for dry condition), respectively. The surface soil moisture prediction based on the multiple sets of weights showed better improvement (R: 0.940, RMSE: 0.025, and MAE: 0.018) compared to the individual model performances and single-weighted prediction (Figure 2.10).

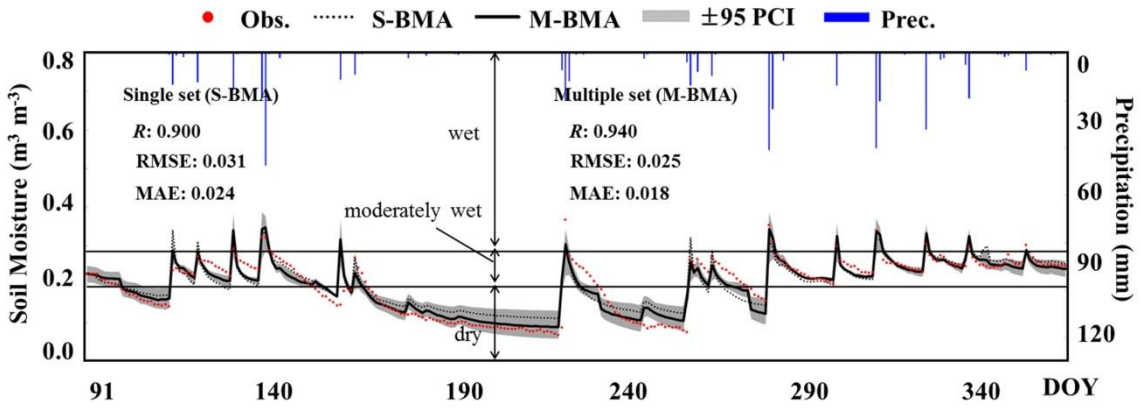

Figure 2. 10 *In situ* and simulated surface soil moisture using a single (S-BMA, dotted line) and multiple (M-BMA, black line) sets of the BMA weights and ±95 PCI at the SCAN 2023 site.

Based on these findings, we suggest that model structural uncertainties can be addressed by the BMA-based multi-model simulation approach using multiple sets of weight corresponding to soil wetness conditions for the two different study sites.

2.5 Conclusions

Soil moisture dynamics estimated by different hydrological models are affected by their own model parameters and structures. Without identifying these uncertainties, the robustness of model outputs from various hydrological models may be elusive. Our study was focused on improving parameter and structural uncertainties caused by different hydrological models in predicting surface soil moisture. In this study, we adapted three different hydrological models (i.e. Noah LSM, SWAP, and CLM) for estimating surface (0-5 cm) soil moisture integrated with a modified-*micro*GA (advanced version of original genetic algorithm (GA)) to search optimized model

parameters for each model. Here, we simulated the surface soil moisture dynamics using the optimized soil parameters of each model in a forward mode. In order to address the effects of model structural uncertainties, we applied a Bayesian Model Averaging (BMA) scheme to the multi model outputs based on the land surface wetness conditions. By aggregating the weighted model outputs for each model, the newly developed approach estimates the effective surface soil moisture dynamics and quantifies model parameter and structural uncertainties. To test our approach, we selected the Little Washita (LW 13 and SCAN 2023) in Oklahoma (sub-humid) and Walnut Gulch (WG 82) in Arizona (semi-arid) sites under the two different hydro-climatic conditions.

For the uncertainty analysis of soil parameters, we used the multi-population for the modified-*micro*GA process with different random number seeds (-1000, -950, and -750). Overall, the estimated parameter distributions for individual models at the LW 13 and WG 82 sites were unimodal which represent the optimized soil hydraulic parameters. However, the (common) optimized parameters of the three different models at the study sites had variations under the similar modeling conditions (i.e., atmospheric forcings, soils, vegetations, etc.) indicating that the individual model performances were affected by their own model parameterization and structural uncertainties.

We derived the surface soil moisture dynamics from the estimated soil parameters using the three models. Mostly, the simulated results of each model matched well with the measurements, but the SWAP and Noah LSM results still had uncertainties showing that a few soil moisture measurements were out of the uncertainty bounds at the LW 13 and WG 82 sites. Furthermore, the outputs from the three hydrological models

showed different model performance under the land surface wetness (i.e., wet, moderately dry, and dry) conditions depending on their inherent model structures. In general, the SWAP model performed better than other models during the wet condition, while CLM and Noah LSM predicted better during the dry period. Thus, we applied the BMA scheme to assign a single or multiple sets (corresponding to various land surface wetness conditions) of weights to each model output for the two study sites. The results showed that the effective surface soil moisture estimates based on multiple sets of weights were more identifiable with the measurements compared to both the original model and single-weighted outputs. It suggests that each model's limitations under certain wetness conditions or hydro-climatic conditions can be compensated by other model strengths. Based on these findings, our proposed methodology can be useful for predicting the effective surface soil moisture estimates and better addressing model parameter and structural uncertainties in soil moisture modeling. Further, this multi-model simulation approach will be applicable to other locations for forecasting soil moisture dynamics effectively using multiple sets of weights derived properly based on wetness conditions or several climate and physical properties.

CHAPTER III

INFLUENCE OF LATERAL SUBSURFACE FLOW AND CONNECTIVITY ON SOIL WATER STORAGE IN LAND SURFACE MODELING[2]

3.1 Synopsis

Lateral surface/subsurface flow and their connectivity play a significant role in redistributing soil water which has a direct effect on biological, chemical, and geomorphological processes in the root zone (~1m). However, most of the land surface models (LSMs) neglect the horizontal exchanges of water at the grid or subgrid scales, focusing only on the vertical exchanges of water as one-dimensional process. To develop better hydrologic understanding and modeling capability in complex landscapes, in this study we added connectivity-based lateral subsurface flow algorithms in the Community Land Model (CLM). To demonstrate the impact of lateral flow and connectivity on soil water storage we designed three cases including: (1) complex surface topography only, (2) complex surface topography in upper soil layers and soil hydraulic properties with uniform anisotropy and (3) complex surface topography and soil hydraulic properties with spatially-varying anisotropy. The connectivity was considered an indicator of the variation of anisotropy in case 3, which was created by wetness conditions or geophysical controls (e.g., soil type, NDVI, and topographic index). These cases were

---

[2] Reprinted with permission from "Influence of lateral subsurface flow and connectivity on soil water storage in land surface modeling" by Jonggun Kim and Binayak P. Mohanty (2016), *J. Geophys. Res. Atmos.*, *121*, doi:10.1002/2015JD024067, Copyright 2016 American Geophysical Union.

tested in two study sites (ER 5 field and ER-sub watershed in Oklahoma) comparing to the field (gravimetric and remote sensing) soil moisture observations. Through the analysis of spatial patterns and temporal dynamics of soil moisture predictions from the study cases, surface topography was found to be a crucial control in demonstrating the variation of near surface soil moisture, but not significantly affected the subsurface flow in deeper soil layers. In addition, we observed the best performance in case 3 representing that the lateral connectivity can contribute effectively to quantify the anisotropy and redistributing soil water in the root zone. Hence, the approach with connectivity-based lateral subsurface flow was able to better characterize the spatially distributed patterns of subsurface flow and improve the simulation of the hydrologic cycle.

3.2 Introduction

Lateral surface/subsurface flow is an important hydrologic process and a key component of the water budget. Through its direct impacts on soil moisture, it can affect water and energy fluxes at the land surface and influence the regional climate and water cycle [*Gochis and Chen,* 2003; *Kumar,* 2004]. Further, the lateral flow and its connectivity play significant role in redistributing soil water which have a direct effect on biological, chemical, and geomorphologic processes in the root zone [*Lu et al.*, 2011; *Western et al*., 2001]. In spite of the importance of lateral flow, most of the land surface models (LSMs: Community Land Model (CLM), Noah Land Surface Model (Noah LSM), Variable Infiltration Capacity (VIC), etc.) neglect the horizontal exchanges of

water at the grid or sub-grid scales, focusing only on the vertical exchanges of water as a one-dimensional process. Surface routing models (e.g., River Transport Model, RTM) are already included to reflect the lateral movement of surface water in land surface modeling, but the lateral subsurface flow is excluded because the models generally assume that lateral transfers of subsurface moisture are fairly marginal in soil water budgets of a regional scale. Recently, 3-D hydrological surface-subsurface models were developed by coupling LSMs with distributed hydrological models to account for interactions between atmospheric, hydrological, and ecological processes (CATHY/NoahMP [*Niu et al.*, 2014] and PARFLOW/CLM [*Maxwell and Miller*, 2005]). Although these hydrological models include a process for the lateral subsurface flow, they still have limitations for deriving lateral hydraulic parameters (e.g., lateral hydraulic conductivity) that might be related to connected patterns of subsurface properties. Furthermore, spatial variability of soil moisture in the unsaturated zone cannot be described successfully without relevant understanding of how the subsurface flow is distributed or connected vertically or laterally in complex landscapes [*Hatton*, 1998; *Zhang et al.*, 1999; *Jana and Mohanty*, 2012a,b,c; *Shen et al.*, 2013]. More realistic understanding of surface and subsurface water movement at large scales can be resolved through a hyper-resolution land surface modeling that allows for better representation of spatially heterogeneous land surfaces [*Wood et al.*, 2011]. Thus, the lateral subsurface flow should be accounted for in hydrological modeling, characterizing vertical and lateral flow components effectively in the unsaturated zone.

Various studies have been conducted to account for the lateral flow in the unsaturated soil. *Zaslavsky and Sinai* [1981] explained a theory of unsaturated lateral flow with the major causes such as soil surface slope, anisotropy, and layering. *Famiglietti and Wood* [1994] developed a land surface modeling approach based on the TOPMODEL framework to address the impact of topographic configuration on soil moisture heterogeneity at a watershed scale. They showed a significant role of the topographic control in development of soil moisture heterogeneity and improved the simulation of hydrologic cycle using the modeling approach. *Chen and Kumar* [2001] explored the role of the topographic control in the seasonal and inter-annual variations of energy and water balances using statistical moments of topographic wetness indices and observed an improvement of stream-flow predictions. Gravity and gradients in matric potential are also critical mechanisms in the unsaturated zone, causing soil water movements from high to low potential [*McCord and Stephens,* 1987; *Jana and Mohanty,* 2012a,b,c]. Water moving vertically through a heterogeneous soil profile can be influenced by the heterogeneity of soil hydraulic properties between soil layers, which can cause lateral flow at the interface [*Zhu and Lin,* 2009]. In process-based SVAT (Soil-Vegetation-Atmosphere Transfer) models, soil hydraulic properties (e.g., saturated soil water content, soil matric potential, and saturated hydraulic conductivity) are critical inputs to account for water movement in soil. The soil hydraulic properties are normally derived using several empirical equations (e.g., van Genuchten, Cosby, and Clapp and Hornberger) according to soil texture. Among the soil properties, an estimation of lateral hydraulic conductivity is more challenging because of the lack of available information.

Thus, anisotropy has been used to derive the lateral hydraulic conductivities from the relationship between vertical and lateral permeability because soil behaves as an anisotropic medium which can cause lateral subsurface flow [*Zaslavsky and Sinai*, 1981; *Wang et al.*, 2011]. In the previous studies related to soil anisotropy, statistical or empirical anisotropy ratios were used at various scales [*Chen and Kumar*, 2001; *Kumar*, 2004; *Assouline and Or*, 2006; *Maxwell and Kollet*, 2008]. However, available experimental data and information for the anisotropy ratio in unsaturated soils might be limited to be applied in heterogeneous landscapes of large land areas. In order to overcome the limitations, the anisotropy ratio can be derived by spatially distributed patterns of wetness condition or its dominant physical controls such as soil texture, vegetation (*NDVI*), and topographic index (*TI*) to characterize the spatial pattern of subsurface flow in the unsaturated zone [*Chen and Kumar*, 2001].

A hydrologic connectivity has been proposed to address not only hydrologic flow paths but also spatial patterns of soil moisture variability at a catchment scale [*Western et al*., 2001; *Hwang et al*., 2009; *Gaur and Mohanty*, 2015]. The lateral connectivity is critically important for representing connected pathways of runoff in the landscapes and understanding movements of surface/subsurface flow [*Mueller et al.*, 2007; *Smith et al.*, 2010]. *Jencso et al*. [2009] derived hydrologic connectivity between catchment landscapes and channel networks to identify runoff source areas based on the topographic characteristics. *Hwang et al*. [2012] found significant relationships between annual hydrologic metrics (e.g., runoff and ET) and HVG (Hydrologic Vegetation Gradient) used as an indicator for lateral hydrologic connectivity at a watershed scale.

Lateral subsurface flow connectivity can be derived from spatially distributed patterns of wetness condition or dominant physical factors and used to quantify the spatially varied anisotropy ratios in heterogeneous landscapes. In this study, we explored the influences of lateral subsurface flow and its connectivity on soil water storage in the unsaturated zone using a land surface model (Community Land Model: CLM). None of previous studies have considered spatially varying anisotropy ratios derived from lateral connectivity to consider the lateral subsurface flow in hydrological modeling.

Thus, the objectives of this study are: 1) to develop better hydrologic understanding and modeling capability in complex landscapes using a connectivity-based lateral subsurface flow algorithm and 2) to demonstrate the subsurface flow variability effectively using spatially distributed patterns of root zone wetness conditions and its physical controls at field and sub-watershed scales. Although this study was focused on smaller scale hydrological processes compared to large scale climate models, it still can provide insights for large scale land surface modeling to enhance their capability. In this study, the concept of lateral flow was used for the unsaturated zone that can be governed by topography and gradients in matric potential.

3.3 Methodology

3.3.1 Study Area

El-Reno site 5 (ER 5: field scale) and El-Reno sub-watershed (ER-sub: sub-watershed scale) located in the North Canadian River basin in Oklahoma were selected to evaluate the proposed approach in this study (Figure 3.1). The ER 5 site (area: 0.8 km

× 0.8 km) is located within the ER-sub boundary (area: 27 km^2). These sites have a sub-humid climate with an average annual rainfall of approximately 805 mm. Daily-mean maximum temperature is 34˚C in July with annual-mean temperature of 15˚C. The topography of the ER 5 is generally flat with average slopes less than 4.0%, while the ER-sub site has a variety of slopes from 11.0% to 0.001%. The ER 5 site has a native grass with 1m root depth and mostly silty loam across the study domain. Vegetation in the ER-sub ranges from short and tall grasses (predominant) and forest in the north and central area to cropland in the south. Various soil types (e.g., silty loam (dominant), loam, and clay loam) are represented across the region.

Our proposed approach was validated with daily *in situ* soil moisture (49 sampling points) measured in the top 5 cm soil (June 18th – July 17th) and in depths of 0-15, 15-30, 30-45, 45-60, and 60-90 cm (July 6th and July 15th) during the Southern Great Plains experiment 1997 (SGP97) [*Mohanty et al*., 2002] for the ER 5 site. Using a truck mounted Giddings probe, soil samples between the land surface and 90 cm depth were collected on a 7 × 7 square sampling grid (100 m spacing between sampling points) across the ER 5 field (Figure 3.1(a)). For the ER-sub site, we validated model predictions with Electronically Scanning Thin Array Radiometer (ESTAR) pixel-based (800 × 800 m) near surface soil moisture products [*Jackson et al.*, 1999] obtained during Southern Great Plains Experiment 1997, SGP97 (June 18th – July 17th) (Figure 3.1(b)).

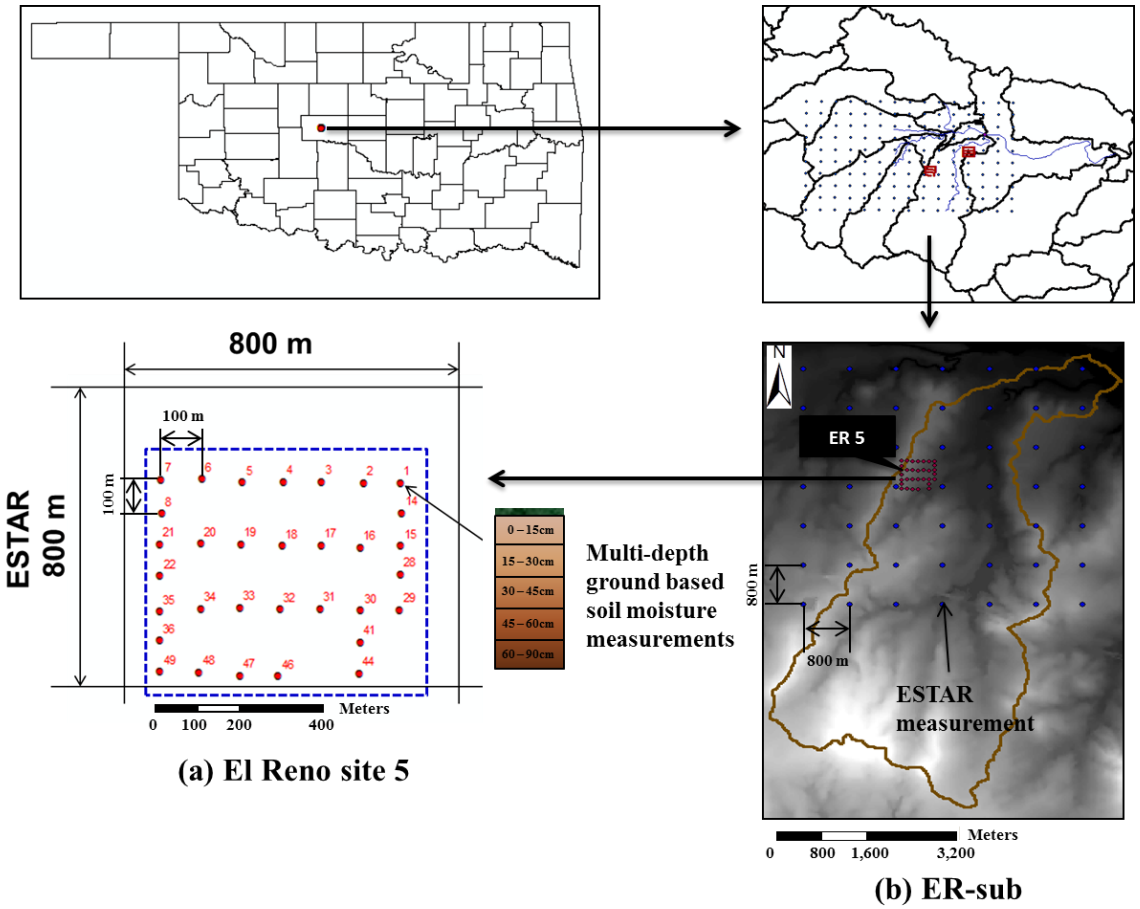

Figure 3. 1 Study sites for (a) El Reno 5 (ER 5) matching the ESTAR remote sensing footprint with multi-depth ground based soil water measurements using truck-mounted Giddings probe (100 m spacing) and (b) El Reno sub-watershed (ER-sub) in Oklahoma.

3.3.2 Description of Model Condition and Forcing Data

Community Land Model (CLM, *Oleson et al.*, [2010]) serves as the dynamic land surface model component of Community Earth System Model (CESM, *Oleson et al.*, [2010]), which consists of various processes such as biogeophysics, hydrologic cycle, biogeochemistry, and dynamic vegetation. The model can be run in offline mode with

prescribed forcing data or in a mode fully coupled to CESM with output from Community Atmosphere Model (CAM, *Collins et al.*, [2006]) which is the atmospheric component of CESM. CLM simulates surface and subsurface runoff based on the simple TOPMODEL-based runoff model (SIMTOP) [*Niu et al*., 2005]. The model considers water table dynamics as the lower boundary using the SIMple Groundwater Model (SIMGM, *Niu et al.*, [2007]). Bare soil evaporation is simulated based on the *Philip and De Vries* [1957] diffusion model, and transpiration process uses an aerodynamic approach based on the Biosphere Atmosphere Transfer Scheme (BATS) model [*Dickinson et al*., 1993] and a stomatal resistance from the LSM model [*Bonan*, 1996]. River Transport Model (RTM) is coupled to CLM for the runoff routing process over a domain [*Oleson et al.*, 2010]. In this study, we used CLM4.0 and ran the model with RTM in offline mode. The soil column in CLM consists of ten soil layers with the thickness of 1.75, 2.76, 4.55, 7.5, 12.36, 20.38, 33.60, 55.39, 91.33 and 113.7 cm (total depth of 343 cm). Soil water flow in CLM is simulated by the modified one-dimensional (1-D) Richards' equation [*Zeng and Decker*, 2009]. CLM has been enhanced to improve hydrological cycle (water balance), vegetation dynamics, and computational performance in the last decade. Nevertheless, the model still simplifies complex processes for the root zone soil hydrology considering only vertical flow using a 1-D Richards's equation. In this study, we modified soil water flow process including a lateral flow component in the unsaturated zone to improve the model performance (as described in section 2.3).

We ran the model in offline mode with atmospheric forcing data (precipitation, temperature, specific humidity, wind speed, surface air pressure, and solar radiation) collected from North American Land Data Assimilation System (NLDAS) which were applied uniformly for the study sites. In this study, we generated model input at spatial resolutions of 50 m and 100 m for the ER 5 site and the ER-sub, respectively. As required input datasets, land cover, soil types with depth, and topographic information were obtained from NLCD (National Land Cover Database), SSURGO (Soil Survey Geographic database), and NED (National Elevation Dataset), respectively. The bottom boundary condition of the model is decided with the water table dynamics calculated from aquifer water storage via the SIMGM [*Niu and Yang*, 2007], and then the model performed a spinning up to initialize the soil profile for the initial condition. In CLM, soil hydraulic properties are determined based on percentages of clay and sand using an empirical equation developed by *Clapp and Hornberger* [1978]. However, CLM tend to simulate the soil moisture lower than the observations in this study because the parameters estimated from the model input (percentages of clay and sand) for the ER 5 site deviated from the referenced parameter ranges (Clapp and Hornberger table) of silty loam soil (predominant in the ER 5 site). Thus, we adjusted the parameters (trial and error) to satisfy the possible ranges of parameters and applied in CLM and modified CLM.

3.3.3 Lateral Subsurface Flow Process

CLM (based on one-dimensional simulation) assumes that soil water drains only vertically to the water-table and there are no interactions between parallel soil columns.

To improve the simplified subsurface flow process in the unsaturated zone by CLM, we modified the one-dimensional vertical soil water flow with three-dimensional flow based on Richards's equation to consider the lateral subsurface flow in the model. The three-dimensional water flow can be expressed as follows,

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial X_c} - Q = \frac{\partial}{\partial X_c}\left[k_{X_c}\left(\frac{\partial \psi - \psi_E}{\partial X_c} \right) \right] - Q \tag{3.1}$$

where θ is the soil moisture content, t is time, q is the water flux in soil, X_c is $\{x,y,z\}$, x and y represent the horizontal directions, z represents the vertical direction, Q is a sink term (evapotranspiration (ET) loss), k_{Xc} is the unsaturated hydraulic conductivity in the direction X_c, ψ is the soil matric potential, ψ_E is the equilibrium (E) soil matric potential, which means there exists a constant hydraulic potential above the water table, when the water table is within the specified soil column/depth.

To estimate soil moisture content at each layer, the model solves a numerical solution based on Eq. 3.1. A new lateral flow term (q_h) is added into the numerical solution of the model, and then the fluxes are calculated at time $n+1$ as follows,

$$\frac{\Delta X_{h,i} \Delta \theta_{h,i}}{\Delta t} = -q_{h-1,i-1}^{n+1} + q_{h,i}^{n+1} - Q_{h,i} \tag{3.2}$$

where $Q_{h,i}$ is a sink (e.g., ET loss), h and i represent the number of soil columns (i.e., x and y direction) and layers (i.e., z direction), respectively.

The vertical and lateral fluxes in Eq. 3.2 are calculated as follows (Eqs. 3.3 and 3.4),

$$q_{h,i}^{n+1} = q_i^n + q_h^n + \frac{\partial q_i + \partial q_h}{\partial \theta_i}\Delta\theta_i + \frac{\partial q_i + \partial q_h}{\partial \theta_{i+1}}\Delta\theta_{i+1} \tag{3.3}$$

$$q_{h-1,i-1}^{n+1} = q_{i-1}^n + q_{h-1}^n + \frac{\partial q_{i-1} + \partial q_{h-1}}{\partial \theta_{i-1}}\Delta\theta_{i-1} + \frac{\partial q_{i-1} + \partial q_{h-1}}{\partial \theta_i}\Delta\theta_i \tag{3.4}$$

Let

$$q_{i-1}^n = -k_{V,i-1}\left[\frac{(\psi_{i-1} - \psi_i) + (\psi_{E,i} - \psi_{E,i-1})}{z_i - z_{i-1}}\right], \qquad q_i^n = -k_{V,i}\left[\frac{(\psi_i - \psi_{i+1}) + (\psi_{E,i+1} - \psi_{E,i})}{z_{i+1} - z_i}\right],$$

$$q_{h-1}^n = -k_{H,h-1}\left[\frac{(\psi_{h-1} - \psi_h)}{x_h - x_{h-1}}\right], \qquad q_h^n = -k_{H,h}\left[\frac{(\psi_h - \psi_{h+1})}{x_{h+1} - x_h}\right].$$

where k_V and k_H represent vertical and lateral hydraulic conductivity (LT^{-1}), respectively.

To investigate the influences of lateral subsurface flow and its connectivity on soil water storage we designed three cases (Figure 3.2). Case 1 is to determine the lateral subsurface flow by slope of surface topography for all soil layers. In case 2 and 3, the lateral subsurface flow is estimated by topography in the upper soil layers and heterogeneous hydraulic properties in the lower soil layers. One of the most challenging parameters in case 2 and 3 is lateral hydraulic conductivity (k_H) which should be

identified appropriately to consider the lateral movement of soil water in the unsaturated zone. The term anisotropy was employed to derive the parameter (k_H) using uniform and spatially-varying ratios (connectivity-based) for case 2 and 3, respectively. Detailed explanations for each case are discussed in the following sections.

To evaluate the performance of modified model predictions for the study cases, we selected three performance criteria such as Pearson's correlation (R), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE).

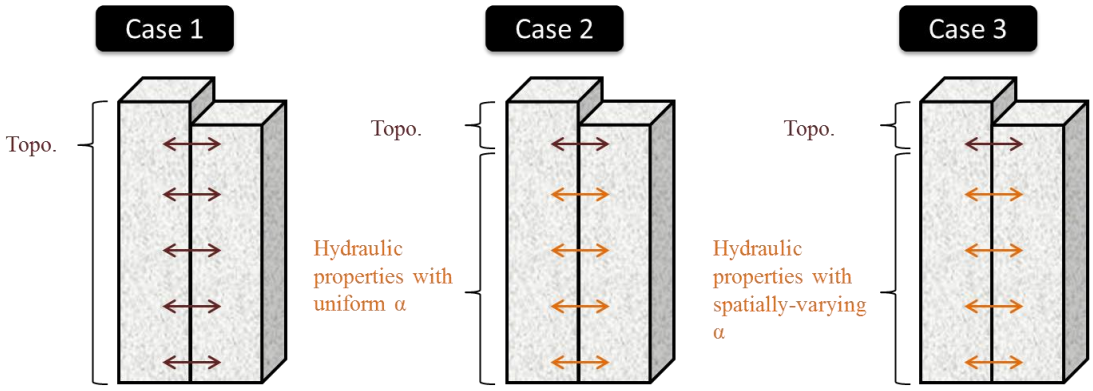

Figure 3. 2 Three study cases designed for the lateral subsurface flow process. Anisotropy (α) is used to derive the saturated hydraulic conductivity in vertical and lateral directions as uniform (case 2) or connectivity-based spatially-varying (case 3) ratio.

3.3.3.1 Case 1: Topography

Surface topographic configuration plays a significant role in determining the soil water flow vertically and laterally near the surface indicating that the changes of flow

direction based on the topography coincide with the changes in the rate of moisture content [*Chen and Kumar*, 2001; *Fan et al.*, 2007]. *Zaslavsky and Sinai* [1981] developed a simple relationship between the vertical and lateral component of soil water movement using the slope of surface topography and found that the lateral component was proportional to the slope and the vertical component of flow. In case 1, we assumed that the lateral subsurface flow moves parallel to the slope of surface topography. The lateral flux (q_h) can be estimated based on the relationship using surface slope as follows (Eq. 3.5),

$$q_h = -k\left[\frac{\partial(\psi - \psi_E)}{\partial z}\right]\tan\beta \tag{3.5}$$

where β is the slope angle.

In addition, flow directions derived from digital elevation method (DEM) using a single-direction algorithm (D8) in GIS hydrologic modeling were included to determine the direction of flow out of each soil column. Thus, the soil water flow process in CLM was modified using Eq. 3.5 with surface slope and flow direction for all soil layers to evaluate the influence of surface topography on the lateral subsurface flow in the unsaturated zone (Figure 3.2 (case 1)).

3.3.3.2 Case 2: Topography and Heterogeneous Hydraulic Properties with Uniform Anisotropy

Surface topography can be a dominant factor to determine the lateral component of subsurface flow near the slope surface, while the lateral subsurface flow in deep soil layers can be more influenced by heterogeneity of hydraulic properties [*Lu et al.*, 2011]. Thus, the two hydrologic processes (surface topography for 1st to 3rd layers and heterogeneous hydraulic properties for 4th to 10th layers) were considered together in case 2 (Figure 3.2). To take into account the lateral subsurface flow based on heterogeneous hydraulic properties, vertical and lateral hydraulic conductivity must be determined across a domain. However, the lateral hydraulic conductivity for spatially heterogeneous landscapes is unavailable and difficult to measure, especially for large areas. Due to the limitations, an anisotropy ratio has been proposed to derive the saturated hydraulic conductivity in vertical and lateral directions that is defined as a directionally dependent property of soil [*Chen and Kumar*, 2001; *Choi et al.*, 2007]. The lateral saturated hydraulic conductivity ($K_{s,H}$) for each soil layer can be derived by multiplying the vertical saturated hydraulic conductivity ($K_{s,V}$) with the anisotropy ratio (α) as,

$$K_{s,H}(z) = \alpha K_{s,V}(z) \tag{3.6}$$

The anisotropy ratio (α) can be obtained from published results or via model calibration through sensitivity analysis. In this study, we run the model adjusting the

anisotropy ratio by trial and error within the possible ranges from the literature [*Chen and Kumar,* 2001] that compared to the soil moisture measurements for each depth. In turn, the appropriate ratio selected was applied to estimate the lateral hydraulic conductivity uniformly across the study sites in the modified CLM model.

3.3.3.3 Case 3: Topography and Heterogeneous Hydraulic Properties with Connectivity-based Spatially-varying Anisotropy

In the previous section, the anisotropy ratio (α) was applied with a constant value across the whole domain. However, anisotropy can be varied for different directions in accordance with various landscape conditions such as soil, vegetation, and topography configuration [*Chen and Kumar,* 2001]. In case 3, we added a connectivity-based lateral subsurface flow algorithm in the subsurface process of CLM to quantify the spatially varying anisotropy ratio for the two study sites. Hydrologic connectivity is critically important for understanding spatial patterns of subsurface flow and can play a significant role in redistributing soil water in the unsaturated zone. It represents how a certain cell in a domain is connected to another cell through an indicator map. The indicator map (I) is used to identify the spatial patterns (connectivity) of the variable of interest (u, e.g., wetness condition or physical controls) above a threshold value (s) in the hydrologic connectivity process (Eq. 3.7). The connectivity is calculated based on the indicator map using the connectivity function ($\tau(d)$) expressed as the probability that a certain cell (x) in a domain (X) is connected to another cell with a distance ($x+d$) in X (Eq. 3.8).

$$I(u) = \begin{cases} 0 & if \quad u < s \\ 1 & if \quad u \geq s \end{cases} \tag{3.7}$$

$$\tau(d) = P(x \leftrightarrow x + d \mid x, x + d \in X) \tag{3.8}$$

Spatially-varying anisotropy can be quantified using the lateral connectivity pattern derived by describing spatially distributed patterns of wetness conditions (e.g., soil moisture measurements) for the ER 5 site and various physical controls (e.g., soil type, vegetation, topography) for the ER-sub site.

The connected patterns of wetness conditions above a certain threshold can be considered as preferred flow paths resulting from connected pixels or concentrated subsurface flow paths, assuming that higher wetness regions produce greater and faster flow in the unsaturated zone [*Western et al*., 2001]. For the ER 5 site, the near surface soil moisture (~ 5 cm) observed on 5 days (June 19th, 25th, July 2nd, 6th, and 12th in 1997) with no rainfall were used to investigate the spatial patterns of soil moisture (wetness) (Figure 3.3(a)).

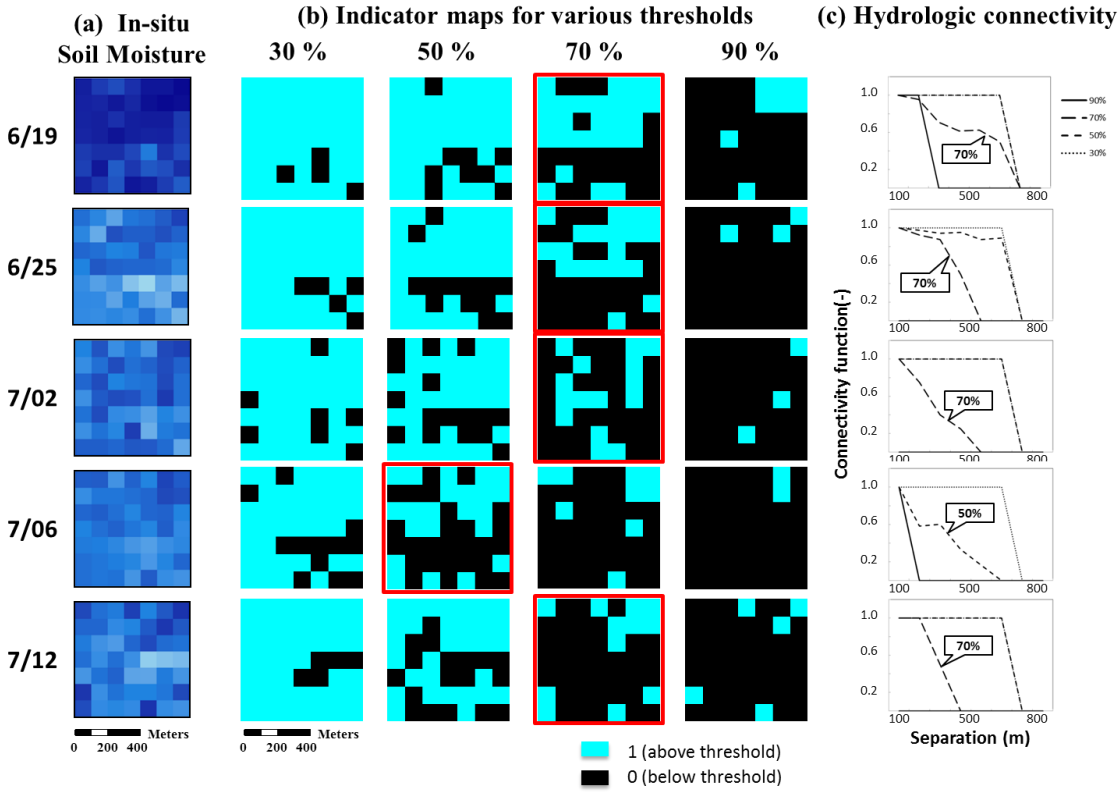

Figure 3. 3 (a) *In situ* measurements at top 5 cm (pixel size: 100 × 100 m), (b) indicator maps for various thresholds of degree of soil wetness (θ/θ_s) on sampling dates, and (c) hydrologic connectivity for 5 sampling dates. Optimum threshold values for daily soil wetness were identified based on visual examination of the connectivity function vs. separation distance plots. Note that selected red boxes around indicator maps correspond to the optimum thresholds selected from the connectivity functions, representing distinct connected patterns on various sampling dates.

Table 3. 1 Thresholds of wetness (*s*) of the near surface soil moisture measurements for the ER 5 site.

	Thresholds of wetness			
Date	30%	50%	70%	90%
6 / 19	0.40	0.44	0.48	0.53
6 / 25	0.23	0.29	0.35	0.41
7 / 02	0.30	0.34	0.39	0.44
7 / 06	0.31	0.35	0.38	0.42
7 / 12	0.26	0.32	0.38	0.44

Indicator maps (binary maps coded 0 or 1) for 4 different thresholds of wetness (30, 50, 70, and 90%) were then created using the soil moisture measurements indicating that pixels of soil moisture above the thresholds are assigned '1' and others are assigned '0' as shown in Figure 3.3(b) and Table 1. Using the indicator maps representing various connected patterns of soil moisture, we calculated the hydrologic connectivity for each map to find an optimum threshold value (or indicator map) that reflects the soil moisture connectivity well for the ER 5 site (Figure 3(c)) following the analysis in *Western et al*. [2001] study. The selected indicator maps for the 5 days (red boxes in Figure 3.3(b)) were combined to consider the possible patterns from the different measurement days and determine how the lateral flow can be distributed across the domain. In turn, we derived spatially-varying anisotropy ratio maps in 8 directions through assigning the ratios ranging from 30 to 0.01 according to the combined indicator map ranging from 0 to 5 (Figure 3.4).

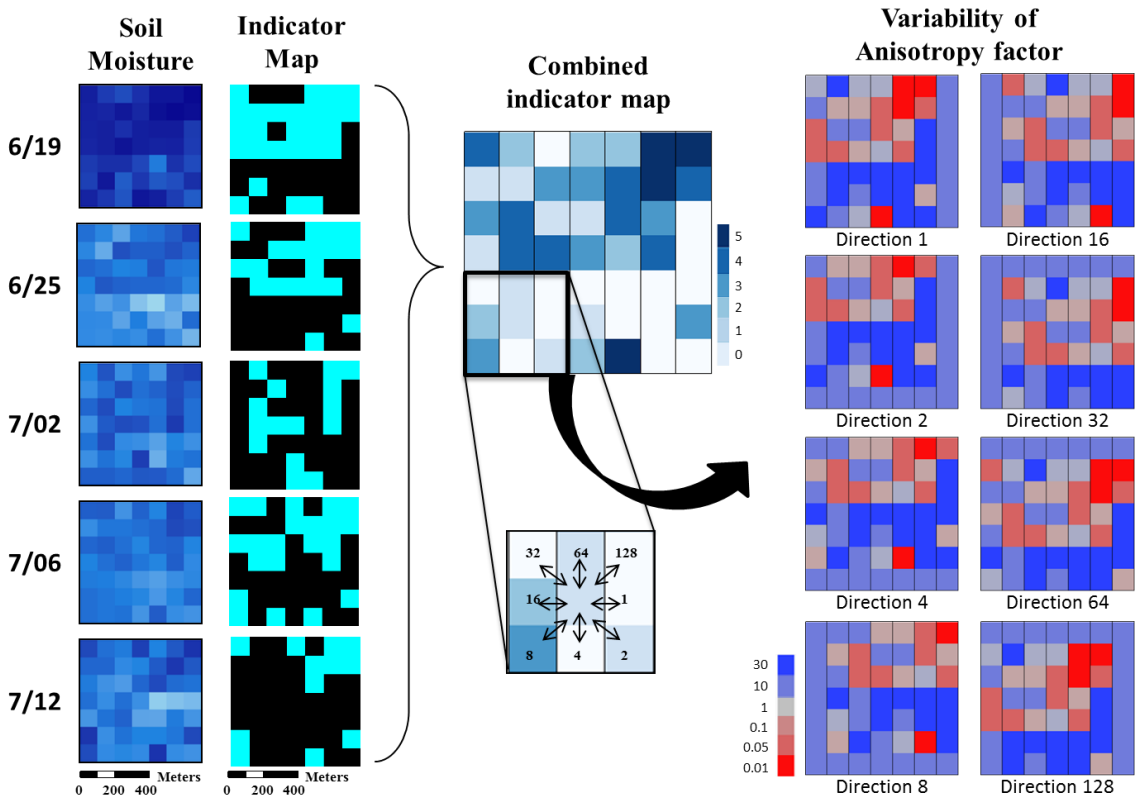

Figure 3. 4 Spatially-varying anisotropy ratio maps (pixel size: 100 × 100 m) (in 8 directions) derived from the connectivity patterns for the near surface soil layers (1st to 3rd) by combining optimum indicator maps for all sampling dates. Similar maps of the other layers (not shown) were derived from the soil moisture measured at deeper soils (up to 90cm).

The possible ranges of the anisotropy ratio were obtained from the literature and the numerical experiments conducted in previous section for the study site. In general, hydraulic conductivity in lateral directions is higher than that in vertical directions, but this is not always true because the unsaturated zone is highly complex with various flow processes such as preferential flow (macropore flow) which might cause soil water movement quickly in vertical direction ($\alpha < 1$) [*Dabney and Selim*, 1987]. The spatially-

varying anisotropy ratio maps for the upper soil layers (1st to 3rd layers) were generated using the near surface soil moisture measurement. For the deep soil layers (4th ~ 10th), the anisotropy ratios were derived from the soil profile measurements (0-15, 15-30, 30-45, 45-60, and 60-90 cm) in a similar way. However, the measurements for deep soil are available only for two days (July 6th and 15th) during the SGP 97 campaign period. Thus, indicator maps for 5 thresholds of wetness (40, 50, 60, 70, and 80%) were estimated (Table 3.2) and combined by adding their binary values to represent the spatially distributed soil moisture patterns and quantify the anisotropy ratios. The spatially-varying anisotropy ratios were then estimated based on the combined map for each soil layer. Thus, the lateral component of subsurface flow was calculated using the anisotropy ratios in the modified CLM for the ER 5 site.

Table 3. 2 Thresholds of wetness of the root zone soil moisture measurements with depth for the ER 5 site.

	Thresholds of wetness				
Depth	40%	50%	60%	70%	80%
0 – 15	0.35	0.38	0.41	0.45	0.48
15 – 30	0.30	0.31	0.33	0.34	0.35
30 – 45	0.21	0.24	0.27	0.29	0.32
45 – 60	0.32	0.34	0.35	0.36	0.38
60 – 90	0.25	0.27	0.30	0.32	0.34

In addition to the wetness condition (soil moisture), various physical controls such as soil, vegetation, and topographic configuration have been identified as dominant controls on the variability of soil moisture at watershed scales [*Mohanty and Skaggs*, 2001; *Joshi and Mohanty*, 2010; *Gaur and Mohanty*, 2013]. These factors can be also used to describe how soil water flows and redistributes in heterogeneous landscapes with regard to the anisotropy. For example, the clay content in soil has a significant effect on anisotropy due to its low permeability retaining more water in soil. Root density in vegetation area could be also related to anisotropy in soil, leading to non-uniform lateral hydraulic conductivity [*Yang and Musiake*, 2003]. The spatial pattern of vegetation density within a watershed is a good estimator for spatial patterns of root zone moisture dynamics and lateral connectivity within watersheds [*Hwang et al.*, 2009]. In this study, soil moisture measurements with depth are not available for the ER-sub site, hence we derived the subsurface connectivity patterns using the dominant physical controls (percentage of clay and sand, NDVI, and Topographic Index) for quantifying the anisotropy ratios (Figure 3.5(a)-(d)).

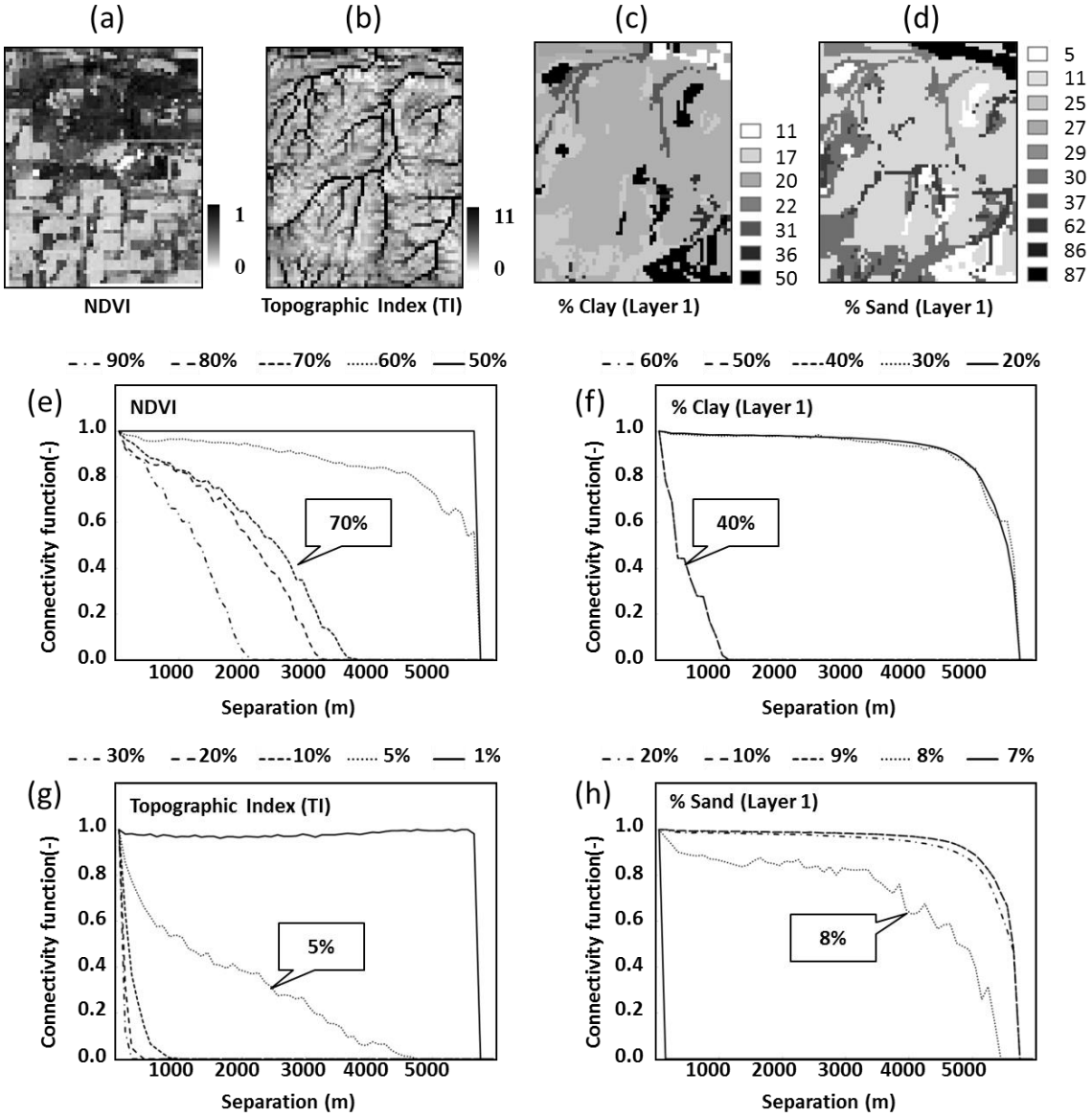

Figure 3. 5 Dominant physical controls ((a) NDVI, (b) %Clay, (c) %Sand, and (d) Topographic Index) and (e-f) their connectivity functions for the ER-sub site. Thresholds values for different physical controls were defined based on its range and numerical analyses. Optimum threshold values for individual physical controls were identified based on visual examination of the connectivity function vs. separation distance plots.

Recent studies explored the combined effects of topography and vegetation on connectivity of runoff source areas and shallow groundwater and showed the potential for improving the estimation of hydrologic connectivity [*Mayor et al.*, 2008; *Hwang et al.*, 2009; *Emanuel et al.*, 2014]. Thus, we considered connected patterns of the combination of physical controls as landscape descriptors or potential predictors for redistribution of soil moisture in the unsaturated zone. Using the connectivity function, we found an optimum threshold for each variable reflecting connected patterns across the ER-sub site (Figure 3.5e-h) and generated their indicator maps using Eq. 3.8. In turn, the indicator maps for the physical controls were combined to reflect the effects of physical controls jointly on hydrological processes that represents unique configurations of the physical components like the concept of hydrological response units (*HRUs*) [*Flügel*, 1995] as expressed in Eq. 3.9.

$$CombinedMap = I(\%clay) + I(\%sand) + I(NDVI) + I(TI) \tag{3.9}$$

where, *I(%clay), I(%sand), I(NDVI), and I(TI)* represent the indicator maps (binary maps) for the percentage of clay and sand, *NDVI* ($(R_{NIR}-R_{red})/(R_{NIR}+R_{red})$), and Topographic Index (*TI*, $\mathrm{Ln}(a / \tan\beta)$), respectively, R_{NIR} and R_{red} are the reflectance of Near InfraRed (NIR) radiation and visible red radiation, respectively, a represents the upslope area, and $\tan\beta$ is the local down slope. The physical controls may not contribute equally to describing the soil moisture variability in the unsaturated zone, but in this study we assumed that the variables have equal effects on hydrological processes because it is

difficult to identify which physical control contributes more to the redistribution of subsurface soil moisture that can vary with complex landscape characteristics. The spatially-varying anisotropy ratio maps for each soil layer were then estimated for the ER-sub (Figure 3.6) and applied in the modified CLM model to estimate the lateral component of subsurface flow.

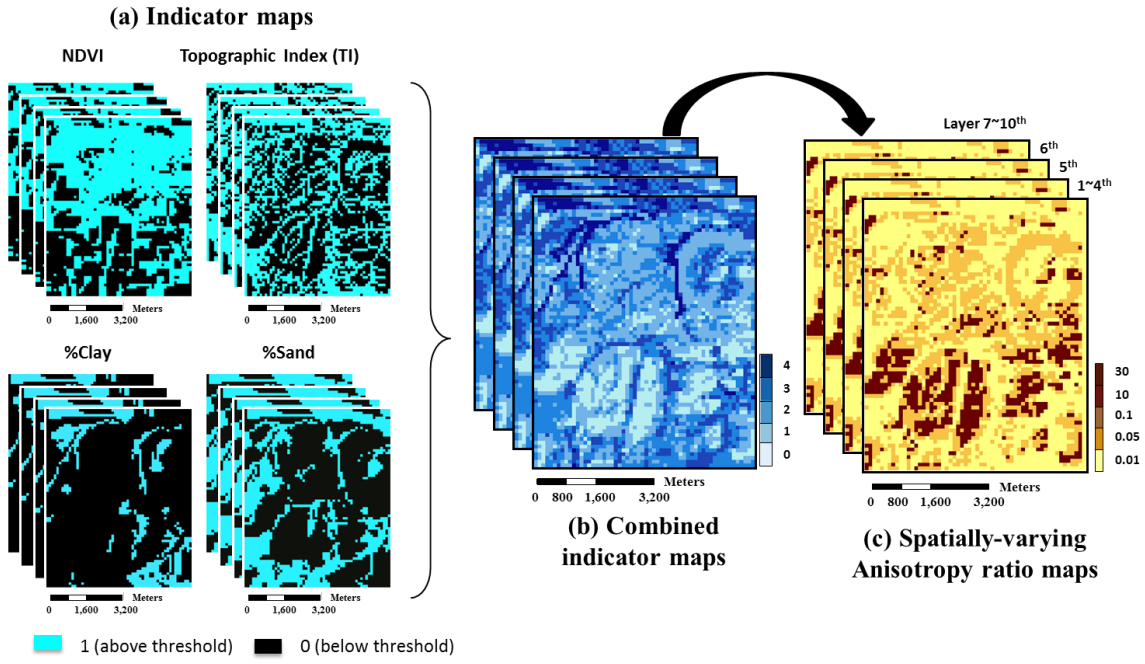

Figure 3. 6 (a) Optimum indicator maps of various physical controls (NDVI, TI, %Clay, %Sand) for various soil layers (1st – 10th), (b) combined indicator maps for each soil layer, and (c) corresponding spatially-varying anisotropy ratio maps (pixel size: 100 × 100 m) at the ER-sub site.

3.4. Results and Discussion

CLM was modified through three different cases designed in this study taking into account the effects of lateral subsurface flow and its connectivity on soil water storage in the unsaturated zone. In order to validate the proposed approach, the simulated near surface and root zone (up to 90 cm) soil moisture using the modified CLM model in the three cases were compared to that of original CLM model and observations at the two study sites (field and sub-watershed scale).

3.4.1 Field Scale (El-Reno Site 5)

Near surface and root zone soil moisture was simulated using the modified model including the lateral subsurface flow based on three different cases at the ER 5 site. Figure 3.7 shows the comparison of observed and simulated root zone soil moisture using the original model and modified model for the three cases with depth (0 ~ 15, 15 ~ 30, 30 ~ 45, and 45 ~ 90 cm) on July 6th 1997. Although the study site has almost uniform land cover and soil type, the observations for all depths showed the variability in the soil moisture distribution that can be attributed to the influence of lateral subsurface flow (Figure 3.7(a)).

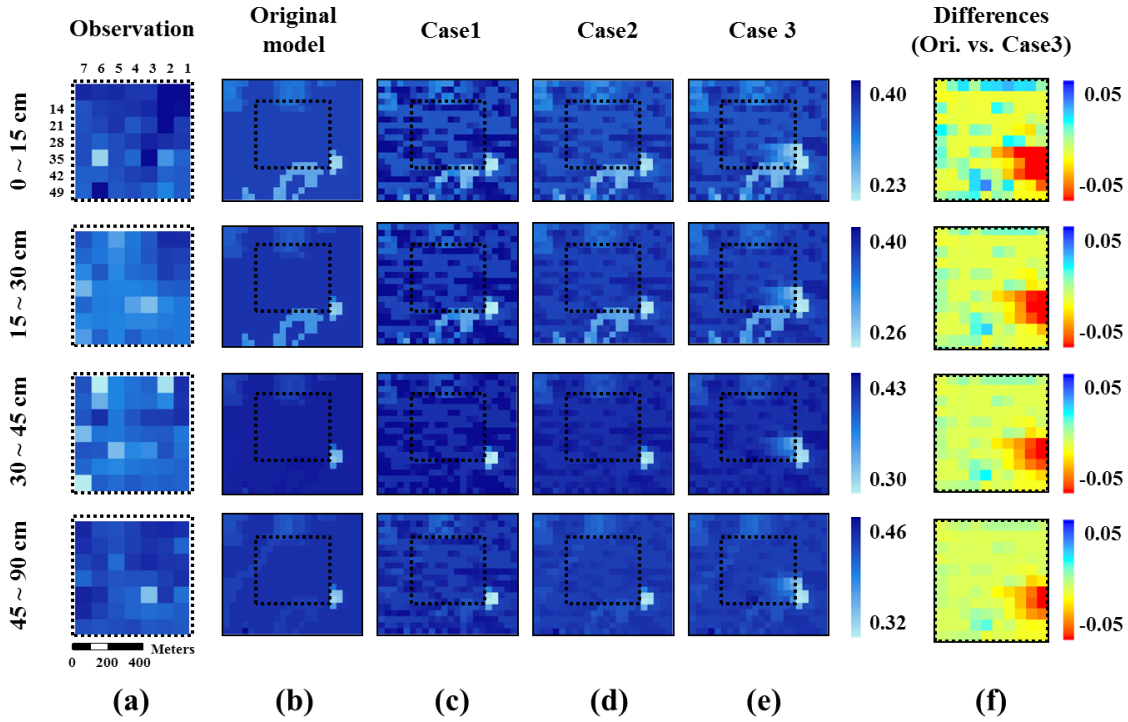

Figure 3. 7 Comparison of the root zone soil moisture of (a) ground observation (pixel size: 100 × 100 m), (b) original CLM model, and modified CLM model (pixel size: 50 × 50 m) through (c) case 1, (d) case 2, and (e) case 3 at the ER 5 site, and (f) differences between the original and modified CLM model of case 3.

However, the original model output represented almost uniform patterns across the site because one-dimensional model estimates the root zone soil moisture identically under the same input data (e.g., vegetation and soil), ignoring the interactions between soil columns (Figure 3.7(b)). When we included the lateral flow component based on the slope, the modified model (case 1) showed spatially distributed soil moisture patterns indicating higher moisture content on the area of low elevation (Figure 3.7(c)). This was because the modified model simulated the root zone soil moisture considering that soil water flows from the upstream to the downstream according to the flow direction as the lateral subsurface flow.

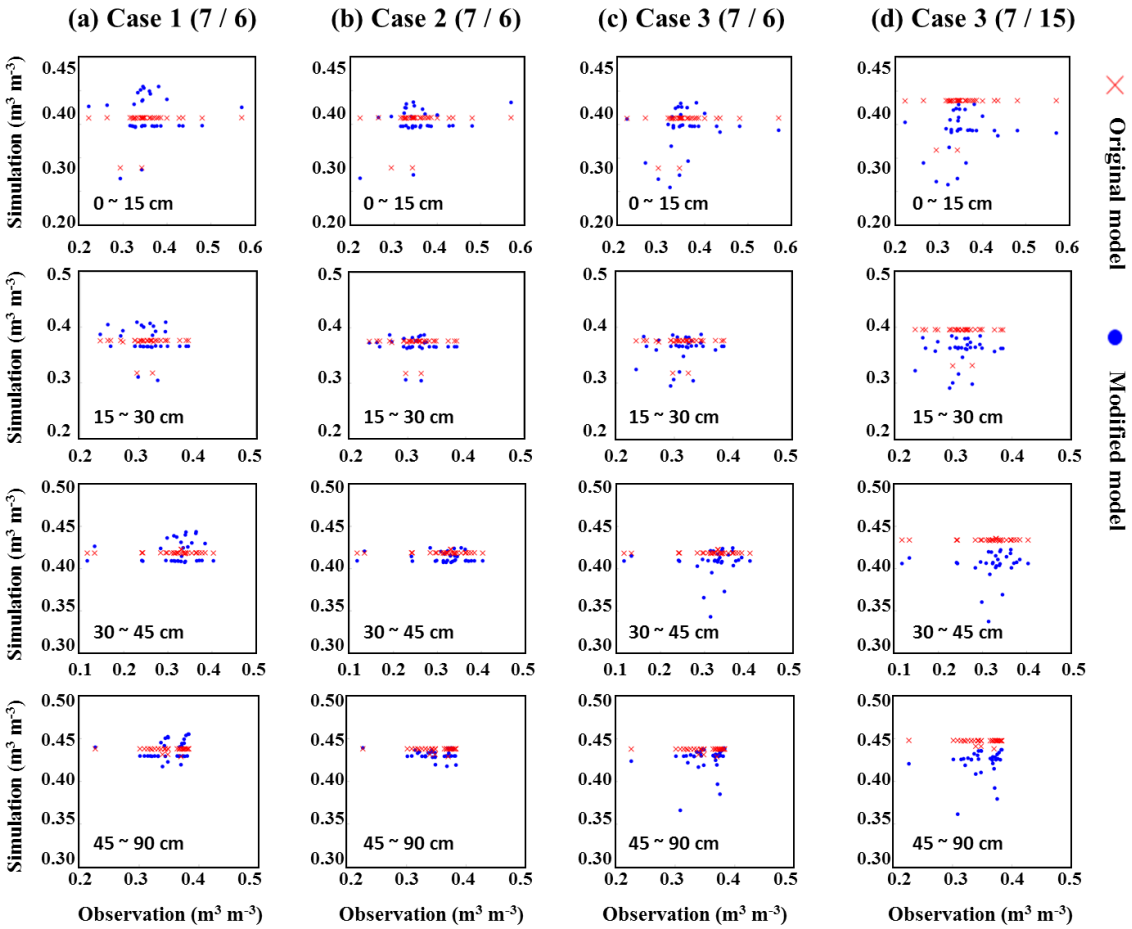

Figure 3. 8 Comparison of the simulated root zone soil moisture using the original and modified model against the observations; (a) case 1, (b) case 2, and (c) case 3 in July 6th and (d) case 3 in July 15th for the ER 5 site.

We also confirmed an improvement of describing the soil moisture variability with the lateral subsurface flow in Figure 3.8(a) which shows the comparisons of simulated root zone soil moisture using the original and modified model against the observations with depth. The original model showed the uniform patterns of root zone soil moisture across the domain, while the modified model (case 1) showed the variation in root zone soil moisture indicating small improvement compared to the original model,

especially at the depth of 0 ~ 30 cm. Based on the results of case 1, we found that the subsurface flow prediction can be improved by considering the lateral subsurface flow based on the topography, but there was still uncertainty predicting the root zone soil moisture in deep soil layers (30 ~ 45 and 45 ~ 90 cm) causing overestimations for the study site. It can be inferred that considering the surface topography only is not enough to account for the root zone soil moisture variability in deep soil because surface and subsurface topography may differ and the lateral subsurface flow in deep soil layers may be governed more by heterogeneous hydraulic properties.

In case 2, the lateral flow component was estimated by topography for the upper layer (1^{st} to 3^{rd}) and heterogeneous hydraulic properties with uniform anisotropy for the lower layers (4^{th} to 10^{th}) together. In this study, we performed the numerical experiments to find a proper (optimum) anisotropy ratio (α) within the possible range (0.01 ~ 30) for the study site. When the anisotropy ratio of 0.05 was applied, the model output (soil moisture with depth) was most similar to the observations through the numerical experiments for the ER 5 site. The ratio was applied uniformly across the domain to estimate the lateral hydraulic conductivity. The modified model (case 2) also predicted the root zone soil moisture better than the original model (Figure 3.7(d)). Figure 3.8(b) shows the predicted root zone soil moisture using the modified model against the multi-depth ground-based observation for all the grid cells. The results of the case 2 were slightly improved than that of case 1 indicating that the average model predictions were closer to the observations. The root zone soil moisture predicted in case 1 (considering surface topography only) tend to be overestimated in all depths, while the modified

model including heterogeneous hydraulic properties with uniform anisotropy ratio showed better performance. This is because high moisture content in certain grid cells can be redistributed effectively into the neighboring cells depending on the heterogeneous hydraulic properties of soil as a lateral subsurface flow. Although case 2 showed more improvement for predicting near surface soil moisture variability (0 ~ 30 cm), it could not capture the soil moisture patterns in deep soil layers using the uniform anisotropy ratio (Figure 3.8(b)). In previous studies, an anisotropy ratio has been applied uniformly in a domain to calculate the lateral component of subsurface flow [*Chen and Kumar*, 2001; *Kumar*, 2004; *Choi et al.*, 2007]. However, we found that the lateral flow component using the constant anisotropy ratio could not identify the subsurface flow successfully in deep soil at the ER 5 site.

In order to overcome the limitations observed in case 1 and 2, the spatially-varying anisotropy ratios were derived from the observed soil moisture patterns (wetness) through the hydrologic connectivity and the optimal thresholds. Figure 3.7(e) and 3.8(c) show the comparison of observed and simulated root zone soil moisture across the study site with depth for July 6th 1997. Compared to the case 1 and 2 with no connectivity, the results of the case 3 with connectivity presented better performances to predict the root zone soil moisture patterns within the domain, even showing improvement in deeper soil layers (30~45 and 45~90 cm). The improvement was also confirmed with a validation in July 15th 1997 (Figure 3.8(d)) representing better agreement with the variability of observations than the original model. It can be inferred that the lateral connectivity derived from the wetness conditions can describe the spatial patterns of subsurface flow

effectively with quantifying the spatially-varying anisotropy ratios. The lateral subsurface flow resulted in the differences between original and modified model prediction that might lead to affect the simulation of the hydrological cycle and various components significantly (Figure 3.7(f)).

Furthermore, we compared the simulated near surface soil moisture dynamics using the case 3 (DOY 170 ~ 197) with *in situ* measurements. To compare the observation and simulation, soil moisture data across the domain was averaged to match the grid based predictions with point-scale observations. The modified model of case 3 (*R*: 0.686, *RMSE*: 0.036, and *MAE*: 0.026) improved the near surface soil moisture predictions more than the original model (*R*: 0.383, *RMSE*: 0.056, and *MAE*: 0.044) (Figure 3.9). Based on these results for the ER 5 site, it was found that the lateral component of subsurface flow in the unsaturated zone is very important for predicting soil water storage successfully in land surface modeling and can be derived with the connectivity-based lateral subsurface flow algorithm. In addition, we can quantify the spatially-varying anisotropy ratios effectively and characterize the lateral subsurface flow variability using the connectivity patterns derived from wetness conditions and geophysical controls in the unsaturated zone.

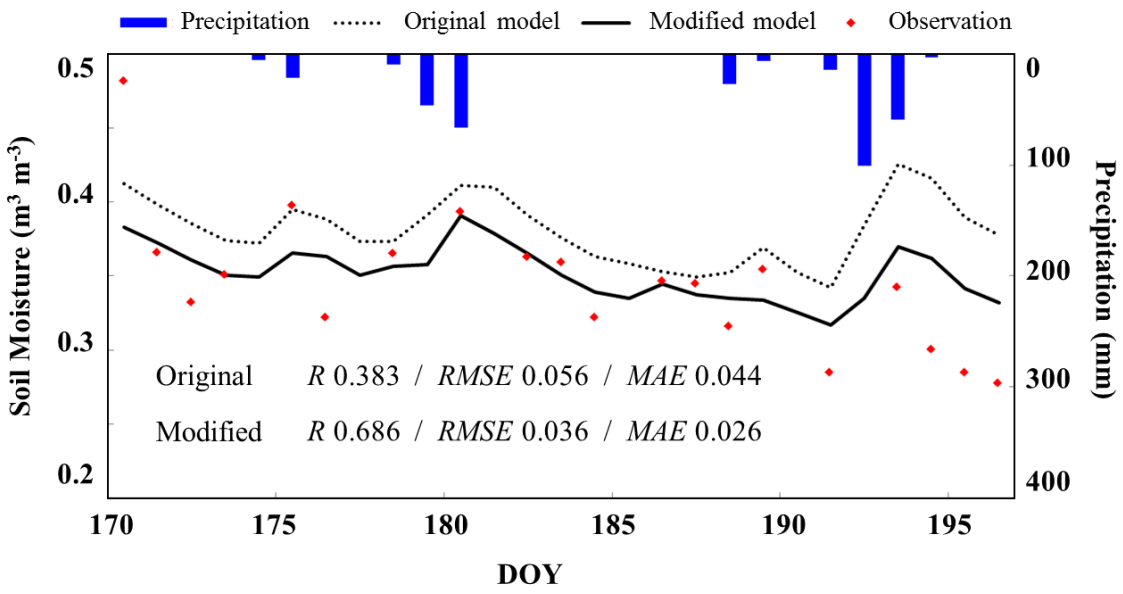

Figure 3. 9 Comparison of observed and simulated (case 3) near surface soil moisture dynamics (top 5 cm) using original (dotted line) and modified model (black line) at the ER 5 site.

3.4.2 Sub-watershed Scale (El-Reno Sub-watershed)

As shown in the previous section (field scale), we confirmed that the modified model with subsurface connectivity (quantifying the spatially-varying anisotropy ratios) performed better at ER-sub site than the original model and the case with spatially uniform anisotropy ratio. Further, we validated the modified model (case 3) in the ER-sub site located in North Canadian River basin to investigate the impacts of the lateral subsurface flow and its connectivity on water storage in soil at a much larger scale. The observed and predicted output was compared with their spatial patterns and temporal dynamics.

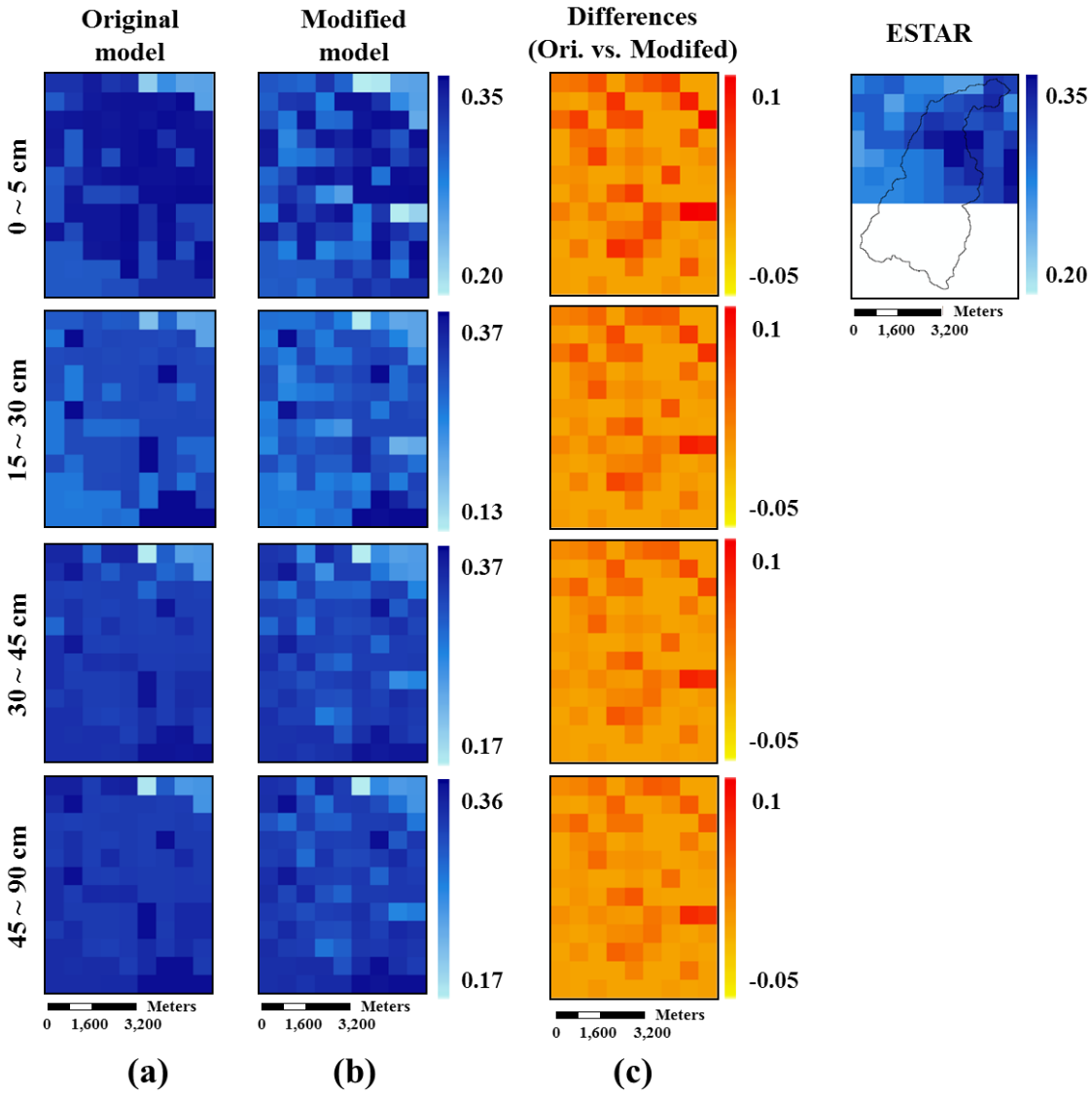

Figure 3. 10 Comparison of the root zone soil moisture (pixel size: 800 × 800 m) at various depths of (a) original model and (b) modified model (case 3), and (c) their differences for the ER-sub site.

Figure 3.10 presents the comparison of the simulated near surface and root zone soil moisture measured at discrete depths using the original and modified CLM model. Compared to the remotely sensed ESTAR observations (0 ~ 5cm), the original CLM model has a limitation in describing the soil moisture variability without lateral subsurface flow (Figure 3.10(a)). The model also tends to overestimate the soil moisture representing relatively uniform distribution in all layers and predicted identical soil water content in grid cells having the same soil type and vegetation due to the limitation of 1-D model. On the other hand, with the connectivity-based lateral subsurface flow the soil moisture prediction was improved representing spatially distributed patterns in all depths (Figure 3.10(b)). The connectivity with depth was derived from the combination of indicator maps (corresponding to their optimum thresholds selected using the connectivity function) of the dominant physical controls (% clay, % sand, NDVI, and TI). It was found that the connected pattern based on the various physical controls can provide significant hydrologic behaviors of subsurface flow to demonstrate the variability of subsurface flow and allow the model to redistribute soil water effectively at the ER-sub site (Figure 3.10(c)). To assess their similarity of spatial patterns quantitatively, the model output was compared to the observations through spatial moving window analysis which is useful to assess spatial patterns. Several different window sizes (1×1, 2×2, 3×3, and 4×4) were selected, and the average of model output within the moving window was used to measure the spatially distributed patterns. For 1×1 window size, the results of the three model evaluation criteria (*R*, *RMSE*, and *MAE*) were too low, although the modified model showed better performances than the original

model (Original CLM model - *R*: 0.30, *RMSE*: 0.090, and *MAE*: 0.083; Modified CLM model - *R*: 0.33, *RMSE*: 0.076, and *MAE*: 0.066). This was because the low values of model evaluation criteria were affected by mismatch between the same grid cells, even though they could be in close agreement at coarser scale (with neighboring grid cells). As the window size increases, the similarity of modified model output increased showing improvements in the model prediction (Table 3.3).

Table 3. 3 Comparison of spatial patterns of simulated near surface soil moisture using the original and modified model with various spatial moving window sizes.

	1 × 1			2 × 2			3 × 3			4 × 4		
	R	RMSE	MAE	R	RMSE	MAE	R	RMSE	MAE	R	RMSE	MAE
Original CLM	0.30	0.090	0.083	0.34	0.076	0.084	0.43	0.086	0.083	0.35	0.084	0.082
Modified CLM (case 3)	0.33	0.076	0.066	0.47	0.069	0.065	0.63	0.067	0.065	0.60	0.065	0.064

Overall, the spatial patterns could not be matched exactly in fine scale, however, the model was able to describe the variability of soil moisture through the connectivity-

based lateral subsurface flow. As shown in Figure 3.11(a), the modified model showed better agreement with the ESTAR observations (*R*: 0.90, *RMSE*: 0.076, and *MAE*: 0.065) than the original model (*R*: 0.88, *RMSE*: 0.089, and *MAE*: 0.077). Though the comparison is based on average soil moisture, it can be inferred that the lateral subsurface flow based on connectivity between subgrid cells within a large grid cell could enhance the modeling skill at large scales. Furthermore, in order to demonstrate the spatial and temporal comparisons within the sub-watershed, we calculated the differences of R values of soil moisture dynamics between the original and modified model (Figure 3.11(b)) in all grid cells. The positive difference (+) mean that the modified model performed better than original model in the grid cell. The results in most of the grid cells showed that the modified model with the connectivity-based lateral subsurface flow can predict the soil water content better spatially and temporally in the ER-sub watershed.

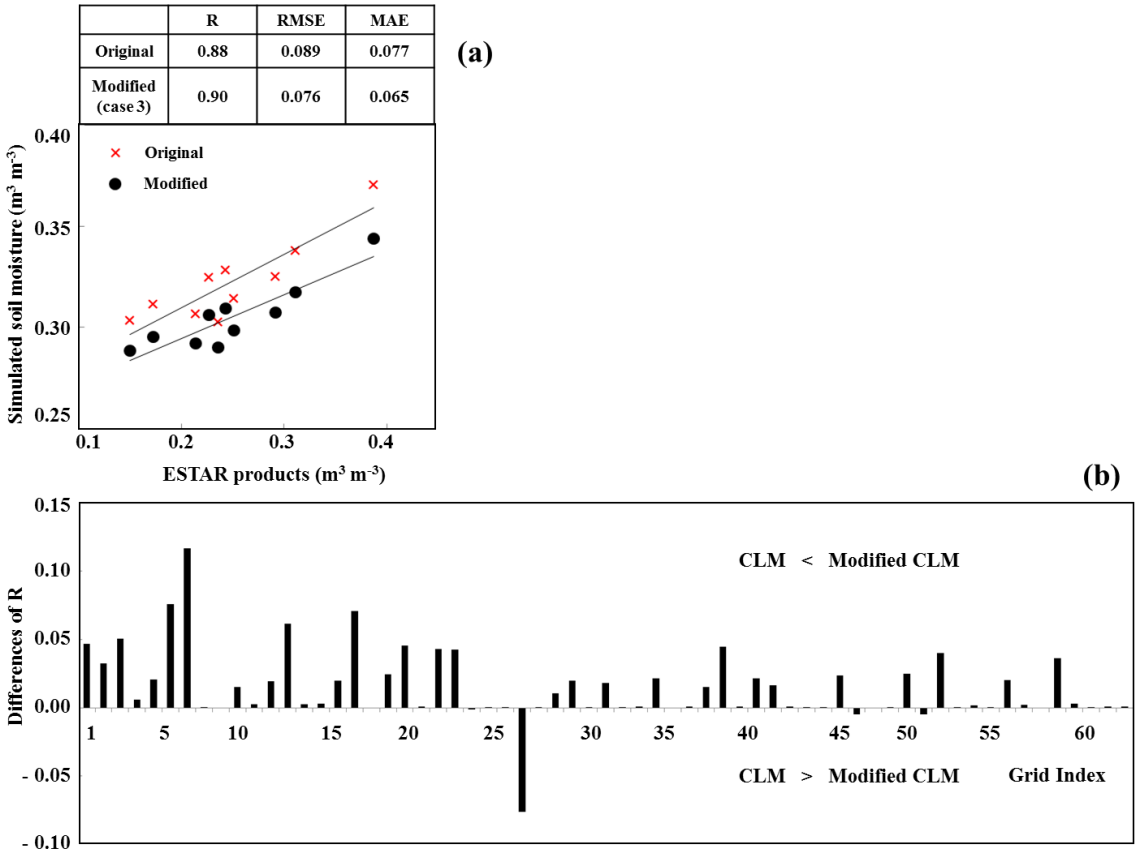

Figure 3. 11 Comparison of the simulated near surface soil moisture for (a) the average within the ER-sub site against ESTAR observations and (b) the differences of R values for each grid cell.

As shown in the comparison of spatial patterns and temporal dynamics of soil moisture prediction, there are differences between the model predictions with and without the lateral subsurface flow in land surface modeling, giving rise to the different soil water storage in the unsaturated zone (Figure 3.10(c)). In land surface modeling, soil moisture is an important component that affects considerably other components of the land surface water cycle (e.g., evapotranspiration, surface runoff, subsurface drainage, etc.) due to the interactions between them. The differences of soil moisture prediction

between the original and modified model led to significantly different surface runoff, subsurface drainage, and water storage (soil water + groundwater) (Figure 3.12).

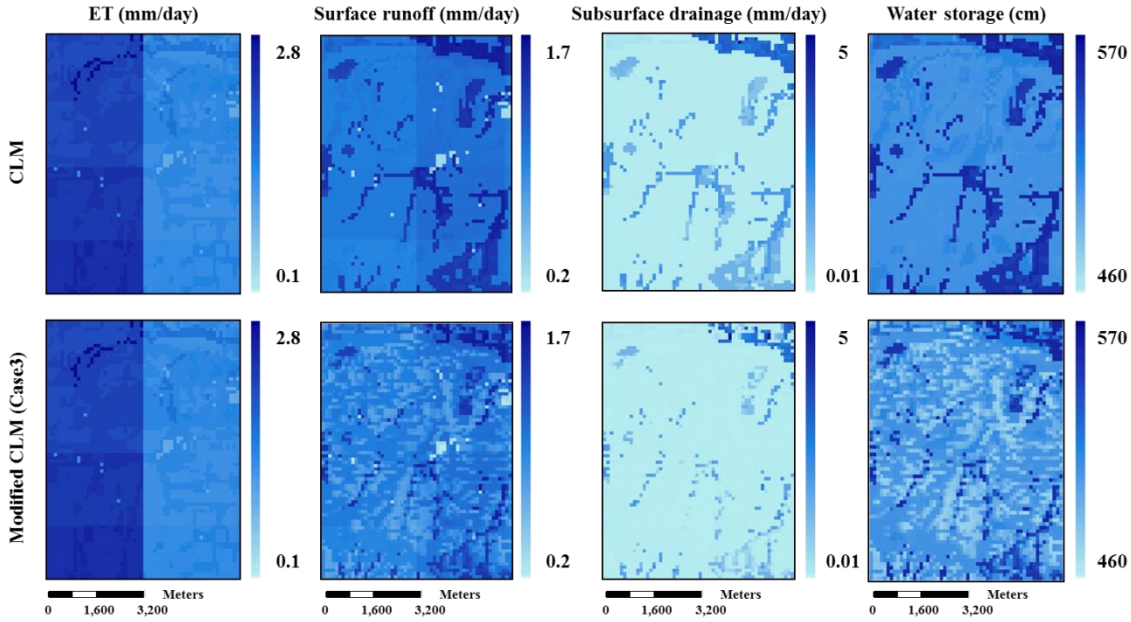

Figure 3. 12 Simulated evapotranspiration (ET), surface runoff, subsurface drainage, and water storage using the original and modified model (pixel size: 100 × 100 m) at the ER-sub site.

3.5. Conclusions

Most of the land surface models are one-dimensional which is not enough to explain the soil moisture variability in the root zone due to absence of interaction (lateral flow) between neighboring soil columns. There is a need to consider the lateral subsurface flow properly in hydrological modeling to account for spatially distributed soil moisture effectively and improve the prediction of subsurface redistribution of flow. Slope of surface topography and heterogeneity of hydraulic properties are considered to include the lateral subsurface flow in the unsaturated zone. One of the important factors is anisotropy ratio used for estimating the lateral hydraulic conductivity that varies spatially according to various landscape conditions such as wetness, soil, vegetation, and topographic configuration. The spatially-varying anisotropy ratios can be derived using a lateral connectivity pattern from wetness conditions and physical controls because the connectivity is a useful concept for understanding spatially distributed lateral subsurface flow and redistributing soil water in the unsaturated zone. In order to investigate the impacts of lateral subsurface flow and its connectivity on soil water storage, in this study we designed three cases (case 1 – surface topography; case 2 – topography and heterogeneous hydraulic properties with uniform anisotropy; case 3 – topography and heterogeneous hydraulic properties with spatially-varying anisotropy derived from connectivity patterns).

In ER 5 field site, the model predictions in case 1 showed the similar patterns to the observed near surface soil moisture distribution, but could not successfully describe the root zone soil moisture patterns in deep soils. It suggests that the surface topography

may not contribute to the lateral subsurface flow in deep soil at the site. The modified model in the case 2 also performed better than that in the case 1 and original model representing a good agreement with the observations. Nevertheless, the case 2 with uniform anisotropy ratio could not still capture the soil moisture variability in deep soil. On the other hands, the model prediction in the case 3 using the spatially-varying anisotropy ratio derived from the connectivity showed more improvements in all soil layers. We found that the connectivity derived from the wetness conditions could characterize the spatial patterns of lateral subsurface flow effectively and quantify the spatially-varying anisotropy ratio properly.

The modified CLM model with connectivity-based lateral subsurface flow (case 3) was validated at a sub-watershed site (ER-sub). The connectivity patterns were developed using the spatial patterns of physical controls (e.g., %sand, %clay, NDVI, and TI) to quantify the spatially-varying anisotropy ratio in this ER-sub site. The modified CLM model improved further the soil moisture prediction than the original CLM model leading to significant differences in performance between the models.

Based on these findings, we infer that the modified model with connectivity can characterize effectively the subsurface flow variability using spatially distributed patterns of wetness condition and physical controls. However, we also found limitations of the approach of deriving anisotropy ratio (α) and wetness connectivity due to their site-specific issue. The parameter and wetness connectivity obtained from combining indicator maps of various physical controls (assuming that the variables have equal effects on hydrological processes) may not be applicable in other sites such as forested

or low-lying areas. The limitations can be addressed to improve the applicability in future works by reflecting effectively site characteristics (dominant physical controls) in various landscapes and climate regions. Although this study has such limitations and was focused on relatively small-scale hydrological processes compared to large-scale climate models (e.g., 1 degree by 1 degree), these processes can be helpful to develop better understanding and modeling capability with the connectivity-based lateral subsurface flow in complex landscapes and allows for an improved simulation of the hydrologic cycle.

CHAPTER IV

A PHYSICALLY-BASED HYDROLOGICAL CONNECTIVITY ALGORITHM FOR DESCRIBING SPATIAL PATTERNS OF SOIL MOISTURE IN THE UNSATURATED ZONE

4.1 Synopsis

Hydrologic connectivity has been proposed as an important concept for understanding local processes in the context of catchment hydrology. It can be useful for characterizing the soil moisture variability in complex heterogeneous landscapes. The current process-based hydrological models could not completely account for flow path continuity and soil moisture spatial distribution in the unsaturated zone. In this study, we developed a physically-based hydrologic connectivity algorithm based on dominant physical controls (e.g., topography, soil texture, vegetation) to better understand spatially distributed subsurface flow and improve the parameterization of soil hydraulic properties in hydrological modeling. We investigated the effects of mixed physical controls on soil moisture spatial variability and developed hydrologic connectivity using various thresholds. The connectivity was used for identifying the soil moisture variability and applied in a distributed land surface model (Community Land Model, CLM) for calibrating soil hydraulic properties and improving model performance for estimating spatially distributed soil moisture. The proposed concept was tested in two watersheds (Little Washita (LW) in Oklahoma and Upper South Skunk (USS) in Iowa) comparing estimated soil moisture with the airborne remote sensing data (Electronically Scanning

Thinned Array Radiometer (ESTAR) and Polarimetric Scanning Radiometer (PSR)). Our finding demonstrated that the spatial variations of soil moisture could be described well using physically-based hydrologic connectivity, and the land surface model performance was improved by using the calibrated (distributed) soil hydraulic parameters. In addition, we found that the calibrated soil hydraulic parameters significantly affect model outputs not only on the water cycle, but also on surface energy budgets.

4.2 Introduction

Recently advanced approaches to catchment dynamics have been proposed through the examination of catchments emergent properties (i.e., spatially connected patterns of flow paths or variable source areas) [*Amoros and Bornette*, 2002; *Sivapalan*, 2005; *McDonnell et al.*, 2007; *Ali and Roy*, 2009]. Hydrologic connectivity has been developed as an important concept for understanding local processes in the context of catchment hydrology and can be defined as connected pathways of surface and subsurface flow and spatial patterns of soil moisture [*Western et al., 2001*; *Ali and Roy*, 2010; *Jencso and McGlynn*, 2011]. It can also provide a missing linkage for preferential flow inferred from unexpected water and chemical migration, which cannot be successfully accounted for through the current parameterization and process-based land surface modeling. Various connectivity metrics have been used in hydrology and ecology such as FRAGSTATS (e.g., cohesion, aggregation index, contagion, etc.) [*McGarigal et al.*, 2002], semivariogram range [*Western et al.*, 1998], gamma index

[*Ricotta et al.*, 2000], directional connectivity index [*Larsen et al.*, 2012], and integral connectivity scale [*Western et al.*, 2001]. The connectivity metrics are useful to better understand catchment hydrologic characteristics and identify runoff source areas at the hill-slope scale. Several studies explored the combined effects of topography and vegetation on connectivity of runoff source areas and shallow groundwater and showed the potential for improving the estimation of hydrologic connectivity [*Mayor et al.*, 2008; *Hwang et al.*, 2009; *Emanuel et al.*, 2014]. *Jencso et al.* [2009, 2010] derived the hydrologic connectivity between catchment landscapes and channel network to identify runoff source areas based on topographic characteristics. They explored the linkage between catchment structure and runoff characteristics and defined the connectivity from flow path continuity across hillslope, riparian, and stream (HRS) interfaces. Using this concept, *Smith et al.* [2013] developed the catchment connectivity model (CCM) to predict streamflow production using simulated hydrologic connectivity across HRS along a stream network. *Western et al.* [2004] demonstrated that saturation excess processes can be indicated by patterns of near surface soil moisture used for developing hydrologic connectivity using the integral connectivity scale technique. Based on these studies, hydrologic connectivity demonstrated significant hydrological behaviors using landscape information such as wetness condition, streamflow, and surface characteristics (e.g., topography and vegetation).

In the past, soil moisture variability has been extensively studied at different spatial scales using *in situ* and remote sensing data in various hydroclimate regions, which is crucial for understanding hydrological processes and catchment characteristics

across scales [*Gaur and Mohanty*, 2016]. The spatial variability of soil moisture can be a critical factor to develop the hydrologic connectivity characterizing spatial patterns of surface and subsurface flow. However, soil moisture information is very limited in deep soils as well as near surface soils for large regions. Soil moisture varies across space and time according to geophysical parameters (i.e., physical controls) such as topography, soil properties, and vegetation characteristics. The physical controls play a significant role in characterizing the heterogeneous landscape in surface and subsurface hydrology [*Famiglietti et al.*, 1999; *Mohanty and Skaggs*, 2001; *Joshi and Mohanty*, 2010]. *Gaur and Mohanty* [2013] explored the effects of physical controls on spatial patterns of soil moisture in humid and sub-humid climatic regions. They identified the dominant physical controls that strongly affect the soil moisture variability at various scales. Spatial patterns of soil moisture are dependent on a set of various (dynamic and static) physical controls which have been defined as precipitation, topography, soil, and vegetation. Thus, the spatial distribution of mixed physical controls can be considered to develop hydrologic connectivity as landscape descriptors or potential predictors for redistribution of surface and subsurface flow. Since precipitation and vegetation vary temporally, dynamic hydrologic connectivity can be also developed using the temporal aspect of physical controls. Recently, *Kim and Mohanty* [2016] developed the hydrologic connectivity algorithm for lateral subsurface flow processes based on the dominant physical controls to improve hydrological modeling at a sub-watershed scale. Their hydrologic connectivity based on the mixed physical controls (assuming that the variables have equal effects on hydrological processes) was successfully reflected to

account for subsurface lateral flow processes in land surface modeling. However, the equal contributions of different physical controls for describing the soil moisture variability may not be applicable in other regions or spatio-temporal scales, which may have different effects of dominant physical controls. Thus, it may be needed to investigate the effects of mixed (weighted) physical controls as well as the interactions between the controls on soil moisture distribution and subsurface flow.

In addition to improving the process modeling, hydrologic connectivity can be employed for improvement of existing parameterizations (especially for soil hydraulic properties) in land surface modeling. Land surface models estimate soil water content in soil profiles based on soil hydraulic properties which directly influence water holding capacity in the unsaturated zone [*Price et al*, 2010]. In land surface modeling, soil hydraulic properties are typically derived from empirical equations as their default parameters such as the pedo-transfer function by *Cosby et al.* [1984]. Although model parameter calibration is critical for achieving accurate model output, most land surface models use a set of default or spatially uniform model parameters [*Li et al.*, 2011]. The default soil hydraulic parameters derived empirically might not be enough to describe the soil moisture variability in spatially heterogeneous landscapes. Thus, in this study, we investigated the effects of mixed physical controls on soil moisture variability to develop physically based hydrologic connectivity and effectively calibrate the distributed soil hydraulic properties across large regions in land surface hydrological modeling.

The main objectives of this study are (1) to study spatially distributed patterns of physical controls which govern soil water redistribution in the unsaturated zone, (2) to develop a physically-based hydrological connectivity algorithm for better describing the spatial connection of subsurface flow in the unsaturated zone, and (3) to improve soil hydraulic parameterization schemes based on hydrologic connectivity in distributed hydrological modeling.

4.3 Methodology

4.3.1 Study Areas

The Little Washita (LW) watershed in Oklahoma and Upper South Skunk (USS) watershed in Iowa were selected as the test sites for this study (Figure 4.1). The study sites have different hydro-climatic conditions and watershed characteristics (e.g., soil properties, land cover, topography). The LW watershed is classified as sub-humid climate with a mean annual rainfall of approximately 926 mm and temperature of 16℃. The LW region (area of about 600 km^2) has rangeland and pastures dominated by patches of winter wheat and other crops, and soil textures ranging from fine sand to silty loam across the watershed. Several field campaigns were conducted in this watershed such as Washita ’94, Southern Great Plains 1997 (SGP97), Soil Moisture Experiments 2003 (SMEX03), and Cloud Land Surface Interaction Campaign 2007 (CLASIC07).

The climate of USS is humid with a mean annual rainfall of approximately 956 mm and temperature of 10.7℃. The region (area of about 2000 km^2) has mostly agricultural crops such as corn and soybean and mainly silty clay loam. The Soil

Moisture Experiments 2002 (SMEX02) and Soil Moisture Active Passive Vegetation Experiment 2012 (SMAPVEX12) field campaigns were conducted in this watershed. Our proposed approach was validated with Electronically Scanning Thin Array Radiometer (ESTAR) pixel-based (800 × 800 m) near surface soil moisture products [*Jackson et al.*, 1999] obtained during SGP97 (June 18th – July 17th, 1997) for the LW watershed and Aircraft Polarimetric Scanning Radiometer (PSR, *Bindlish and Jackson* [2002]) observed during SMEX02 (June 25th – July 12th, 2002) for the USS watershed. To evaluate the performance of land surface model with and without the subsurface hydrologic connectivity for the study watersheds, we selected several pixels on connected and unconnected regions with different characteristics and complexities (e.g., soil type, land use, and topography) as shown in Table 4.1 (Figure 4.1). In addition, the model performances were compared at various extent scales to evaluate the spatial variability of soil moisture prediction for large regions within the watersheds.

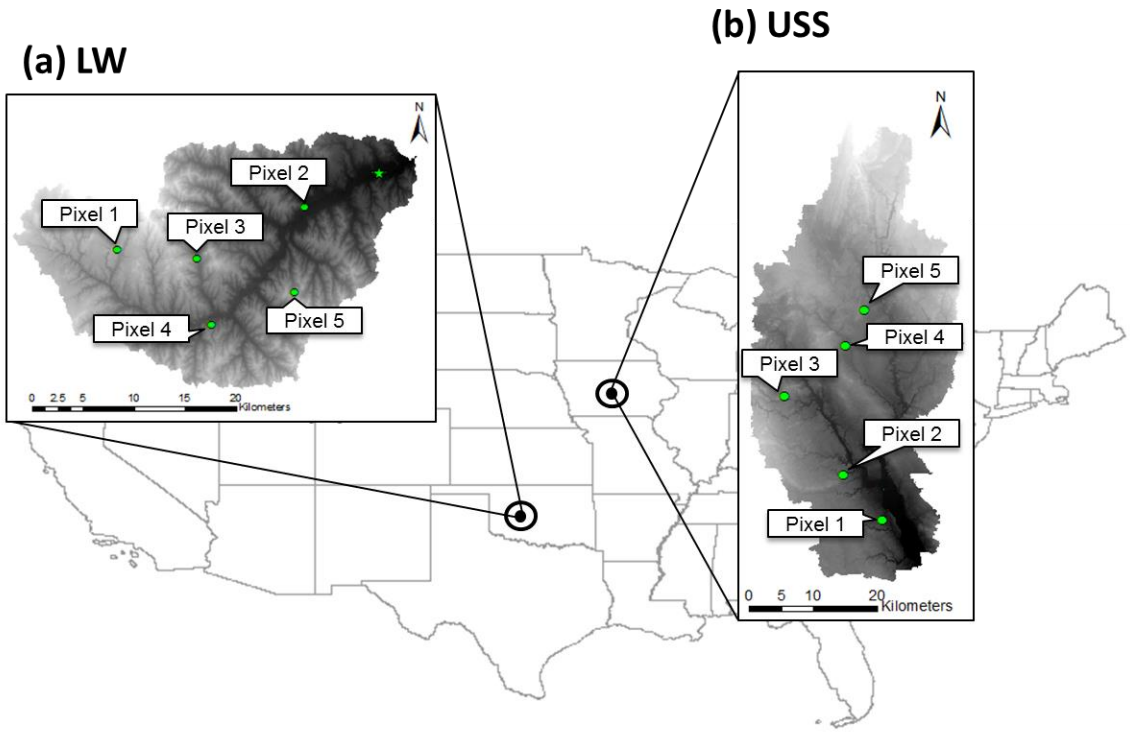

Figure 4. 1 Study sites of (a) Little Washita (LW) in Oklahoma and (b) Upper South Skunk (USS) in IOWA. The pixels represent connected and unconnected regions selected for analysis.

Table 4. 1 Characteristics of selected pixels in the study sites.

	LW			USS		
	Elevation	Soil texture	Landuse	Elevation	Soil texture	Landuse
Pixel 1	418 m	Silty clay loam	Crop	314 m	Sandy clay loam	Crop
Pixel 2	338 m	Loam	Crop	294 m	Loam	Forest
Pixel 3	391 m	Sandy clay loam	Forage	321 m	Loam	Crop
Pixel 4	379 m	Sandy clay loam	Alfalfa	310 m	Clay loam	Grass
Pixel 5	398 m	Clay loam	Pasture	311 m	Loam	Crop

4.3.2 Land Surface Model (Community Land Model)

Community Land Model (CLM) serves as the dynamic land surface model component of the Community Earth System Model (CESM, *Oleson et al.,* [2010]). CLM consists of various processes such as biogeophysics, hydrologic cycle, biogeochemistry, and dynamic vegetation. The model estimates surface and subsurface runoff based on the simple TOPMODEL-based runoff (SIMTOP) [*Niu et al*., 2005]. The SIMple Groundwater Model (SIMGM, *Niu et al.*, [2007]) is used for considering water table

dynamics as the lower boundary. Bare soil evaporation and plant transpiration are calculated using the *Philip and De Vries* [1957] diffusion model and an aerodynamic approach which is based on the Biosphere Atmosphere Transfer Scheme (BATS) model [*Dickinson et al*., 1993] and a stomatal resistance from the LSM model [*Bonan*, 1996]. CLM is coupled with the River Transport Model (RTM) for the runoff routing process [*Oleson et al.*, 2010]. The soil profile is divided into ten soil layers with the fixed thickness of 1.75, 2.76, 4.55, 7.5, 12.36, 20.38, 33.60, 55.39, 91.33 and 113.7 cm (total depth of 343 cm). The soil water flow is solved by the modified Richards' equation (4.1) [*Zeng and Decker*, 2009] which is derived by subtracting the hydrostatic equilibrium soil moisture distribution from the original Richards' equation for improving the mass-conservative numerical scheme when the water table is within the soil column,

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K\left(\frac{\partial(\psi - \psi_e)}{\partial z}\right)\right] - Q \tag{4.1}$$

where ψ and ψ_e are the soil matric potential and equilibrium soil matric potential (cm), z is soil depth (cm) taken positive upward, K is hydraulic conductivity(cm d^{-1}), Q is a soil moisture sink term, which is the root water extraction rate by plants. The hydraulic conductivity, equilibrium soil matric potential, and equilibrium volumetric water content are shown in Eqs. (4.2)-(4.4) based on *Clapp and Hornberger* [1978],

$$K(\theta) = K_{sat}\left(\frac{\theta}{\theta_{sat}}\right)^{2b+3} \tag{4.2}$$

$$\psi_e = \psi_{sat}\left(\frac{\theta_e(z)}{\theta_{sat}}\right)^{-b} \tag{4.3}$$

$$\theta_e(z) = \theta_{sat}\left(\frac{\psi_{sat}+z_{\nabla}-z}{\psi_{sat}}\right)^{-\frac{1}{b}} \tag{4.4}$$

where $K(\theta)$ and K_{sat} are the unsaturated and saturated hydraulic conductivity (cm d^{-1}), θ and θ_{sat} are the volumetric soil water content and saturated soil water content (cm^3 cm^{-3}), ψ_{sat} is the saturated soil matric potential (cm), $\theta_e(z)$ is the equilibrium (e) volumetric water content (cm^3 cm^{-3}) at depth z (z_{∇} is the water table depth), and b is the curve fitting parameter related to the pore size distribution (-), respectively. Primarily, the four soil hydraulic properties (θ_{sat}, K_{sat}, ψ_{sat}, b) are major input parameters for estimating soil moisture distribution in CLM [*Huang et al.*, 2013]. These soil properties are calculated based on the work by *Clapp and Hornberger* [1978] and *Cosby et al.* [1984], which are determined according to percent sand and percent clay contents (Eqs. (4.5)-(4.8)) (called default parameters in this paper). The means and standard deviations of the parameters are available from *Cosby et al.* [1984] as shown in Table 4.2.

$$\theta_{sat} = 0.489 - 0.00126 \times \%sand \tag{4.5}$$

$$b = 2.91 + 0.159 \times \%clay \tag{4.6}$$

$$\psi_{sat} = 10 \times 10^{(1.88-0.0131\times\%sand)} \tag{4.7}$$

$$K_{sat} = 0.0070556 \times 10^{(-0.884+0.0153\times\%sand)} \tag{4.8}$$

Table 4. 2 Means and standard deviations for the four hydraulic parameters for various textural classes (from Table 3 in *Cosby et al.* [1984]).

	b		Log ψ_{sat}		Log K_{sat}		θ_{sat}	
Class	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Sandy loam	4.74	1.40	1.15	0.73	-0.13	0.67	43.4	8.8
Sand	2.79	1.38	0.84	0.56	0.82	0.39	33.9	7.3
Loamy sand	4.26	1.95	0.56	0.73	0.30	0.51	42.1	7.2
Loam	5.25	1.66	1.55	0.66	-0.32	0.63	43.9	7.4
Silty loam	5.33	1.72	1.88	0.38	-0.40	0.55	47.6	5.4
Sandy clay loam	6.77	3.39	1.13	1.04	-0.20	0.54	40.4	4.8
Clay loam	8.17	3.74	1.42	0.72	-0.46	0.59	46.5	5.4
Silty clay loam	8.72	4.33	1.79	0.58	-0.54	0.61	46.4	4.6
Sandy clay	10.73	1.54	0.99	0.56	0.01	0.33	40.6	3.2
Silty clay	10.39	4.27	1.51	0.84	-0.72	0.69	46.8	6.2
Light clay	11.55	3.93	1.67	0.59	-0.86	0.62	46.8	3.5

After investigating the effects of mixed physical controls on soil moisture variability, the soil hydraulic parameters were calibrated using the physically-based hydrologic connectivity algorithm developed in section 4.3.4 (Figure 4.2). To evaluate the land surface model performance, we compared the model outputs (e.g., soil water content, evapotranspiration, surface runoff, and water storage) estimated by using the default soil hydraulic parameters versus using the calibrated soil hydraulic parameters.

We used CLM4.0 loosely coupled with RTM in an offline mode with atmospheric forcing data (precipitation, temperature, specific humidity, wind speed, surface air pressure, and solar radiation) collected from the USDA Agricultural Research Service (ARS) Micronet network for the LW watershed and SMEX02 Rain Gauge network for the USS watershed. To compare model predictions to the ESTAR and PSR soil moisture observations, model inputs for the two watersheds were generated at a spatial resolution of 800 m using land cover, soil types with depth, and topographic information obtained from NLCD (National Land Cover Database), SSURGO (Soil Survey Geographic database), and NED (National Elevation Dataset), respectively.

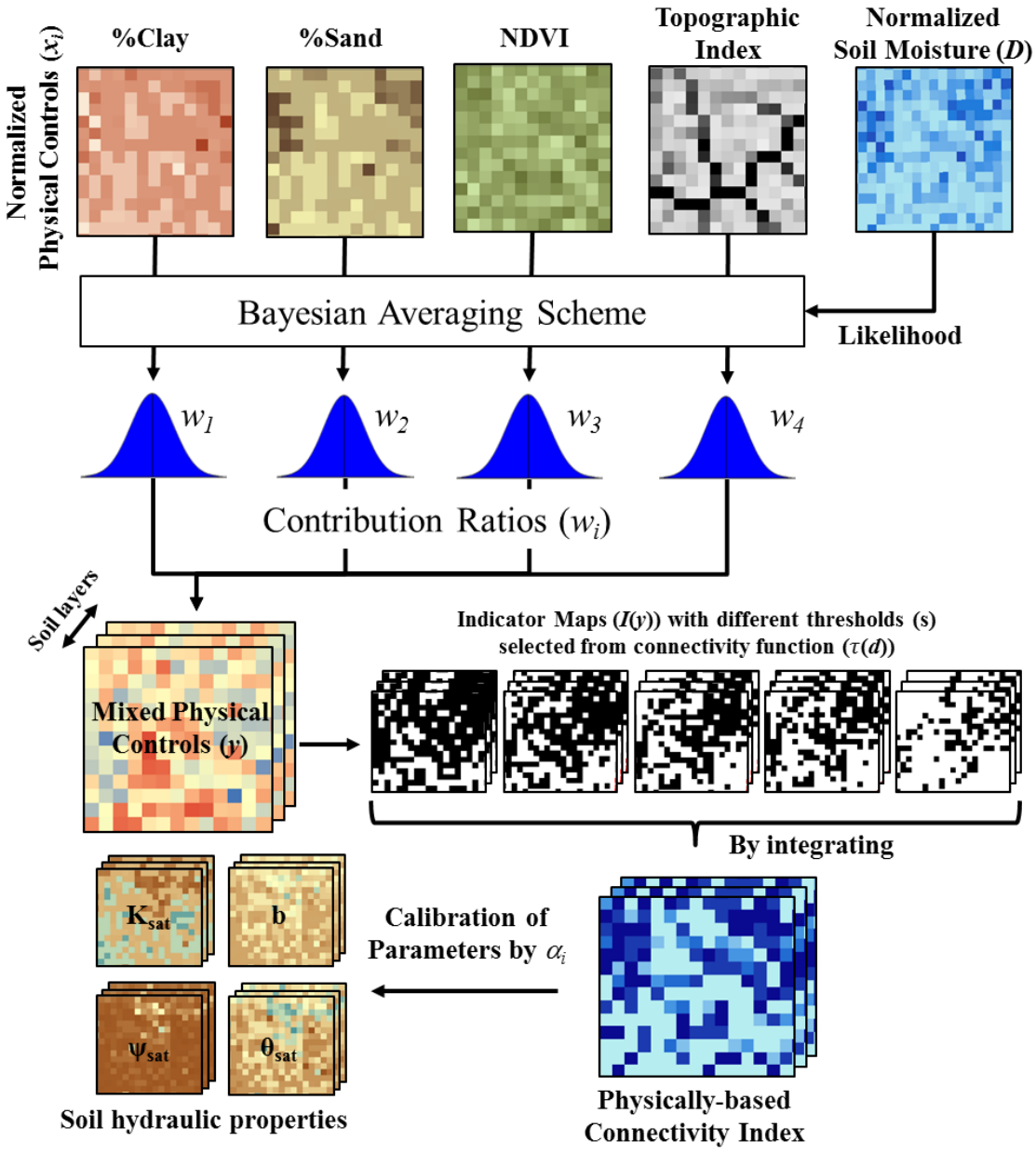

Figure 4. 2 Schematic diagram of information flow for developing connectivity index using Bayesian averaging of dominant physical controls and calibrating distributed soil hydraulic parameters. α_i is the calibrating factor for each parameter based on their standard deviation that is determined by the physically-based hydrologic connectivity index.

4.3.3 Mixed Physical Controls in Complex Landscapes

Kim and Mohanty [2016] developed hydrologic connectivity assuming that all physical controls are contributing equally to representing the soil moisture distribution in the unsaturated zone. However, that assumption has a limitation to be applied into other complex landscapes due to site-specific characteristics. In complex landscapes, spatial distribution of soil moisture varies and shifts with landscape characteristics such as spatial patterns of soils, vegetation, topography, and hydroclimates [*Gaur and Mohanty*, 2013, 2016]. To better characterize the spatial variability of soil moisture, the total contribution of various physical controls and their interactions need to be accounted. In this study, dominant physical controls (i.e., soil texture (%clay and %sand), topography (Topographic Index (TI), Ln(a/tanβ)), and vegetation (Normalized Difference Vegetation Index (NDVI), (R_{NIR}-R_{red})/(R_{NIR}+R_{red})) were considered. R_{NIR} and R_{red} are the reflectance of near infrared (*NIR*) radiation and visible red radiation, respectively; a represents the upslope area; and tanβ is the local downslope. Spatial data were collected from the Soil Survey (SSURGO), Landsat5 imagery, and USDA-NRCS Geospatial Data Gateway for the two watersheds (Figure 4.3). To effectively estimate the contributing ratios (weights) for the physical controls and their interactions, we used the Bayesian averaging scheme [*Hoetting et al.*, 1999] that can provide proper weights that show how the controls contribute to describing the spatial variability of soil moisture (Eq. (4.9)).

$$P(y|x_1, \cdots x_i) = \sum_{i=1}^{j} P_i\,(x_i|D) P_i(y|x_i, D) \tag{4.9}$$

where y is the combined (weighted) physical controls, x_i is the normalized physical controls (i=1,…,j), j is the number of physical controls and interaction terms used, PDF ($P_i(x_i|D)$) is the posterior probability for physical controls given the normalized soil moisture measurements (D) and defined as contributing ratios (w_i) of normalized physical controls (x_1, x_2, x_3, x_4 as %clay, %sand, NDVI, TI), the conditional PDF ($P_i(y|x_i,D)$) represents the posterior distributions of y given physical controls and measurements. In this study, interaction terms were also considered to examine the joint effects of physical controls (e.g., $x_{1\bullet2}$, $x_{1\bullet3}$, $x_{1\bullet4}$, $x_{2\bullet3}$, $x_{2\bullet4}$, $x_{3\bullet4}$). The estimated contributing ratios were used to combine the dominant controls and to develop the hydrologic connectivity for the study watershed.

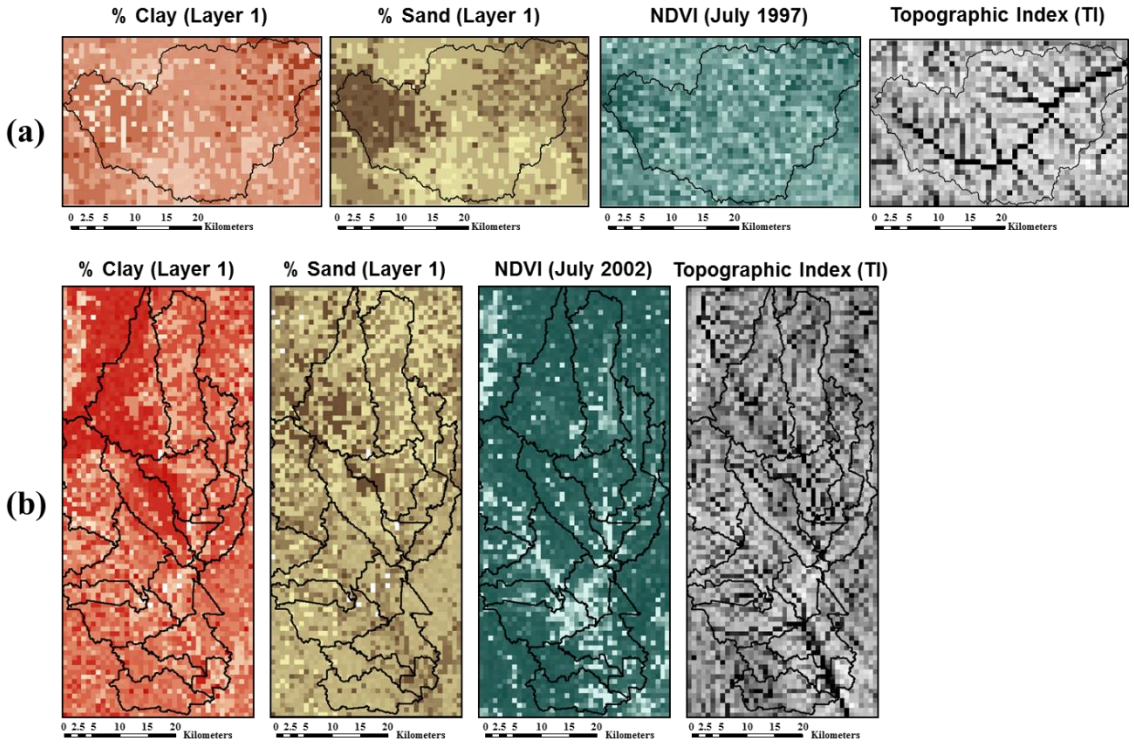

Figure 4. 3 Dominant physical controls (soil texture, vegetation, and topography) for the (a) LW and (b) USS watersheds (Spatial resolution of 800m).

4.3.4 Development of Physically-based Hydrologic Connectivity

By and large, hydrologic connectivity has been developed by patterns of wetness condition (e.g., soil moisture) or surface topography (e.g., contributing area) at a catchment scale [*Western et al., 2001*; *Jencso and McGlynn*, 2011]. However, information for surface wetness or root zone soil moisture is very sparse, and surface topography cannot sufficiently reflect the patterns of subsurface flow [*Kim and Mohanty*, 2016]. Thus, we developed physically-based hydrologic connectivity using the mixed physical controls (i.e., %clay, %sand, NDVI, TI) to identify the spatial variation of soil moisture. Hydrologic connectivity shows how cells are connected to each other across a

domain under a certain threshold of interest variable. Among various connectivity metrics, we used the integral connectivity scale technique which was successfully tested in a previous study [*Kim and Mohanty*, 2016] to describe the soil moisture spatial variability. The indicator map (I) is used to describe the spatial patterns of interest variable (y, mixed physical controls) above a certain threshold (s) in the hydrologic connectivity process (Eq. (4.10)). Connectivity is calculated using the indicator map $I(y)$ and the connectivity function ($\tau(d)$) expressed as Eq. (4.11).

$$I(y) = \begin{cases} 0 & if \quad y < s \\ 1 & if \quad y \geq s \end{cases} \tag{4.10}$$

$$\tau(d) = P(h \leftrightarrow h + d | h, h + d \in H) \tag{4.11}$$

where h is a certain cell in a domain (H), d is the distance between two cells.

Indicator maps ($I(y)$) for various thresholds (0 ~ 100%) were created using a mixed physical controls map generated with the contributing ratios of different physical controls showing that pixels above the thresholds on the mixed controls map were assigned to "1" and others assigned to "0". To consider various connected patterns of mixed physical controls, we selected 5 representative thresholds from the connectivity functions ($\tau(d)$) that reflect the connectivity patterns well across the watershed (Figure 4.4(a)) [*Western et al.*, 2001; *Kim and Mohanty*, 2016]. In turn, the indicator maps for the 5 thresholds were chosen (Figure 4.4(b)). The physically-based hydrologic connectivity index was developed by integrating the indicator maps ranging from 1 to 5

($\sum_{i=1}^{5} I(y)_{s_i}$, s_i is the selected thresholds) (Figure 4.4(c)). Pixels of higher index represent fairly connected and higher wetness regions, while lower index pixels indicate unconnected and drier regions.

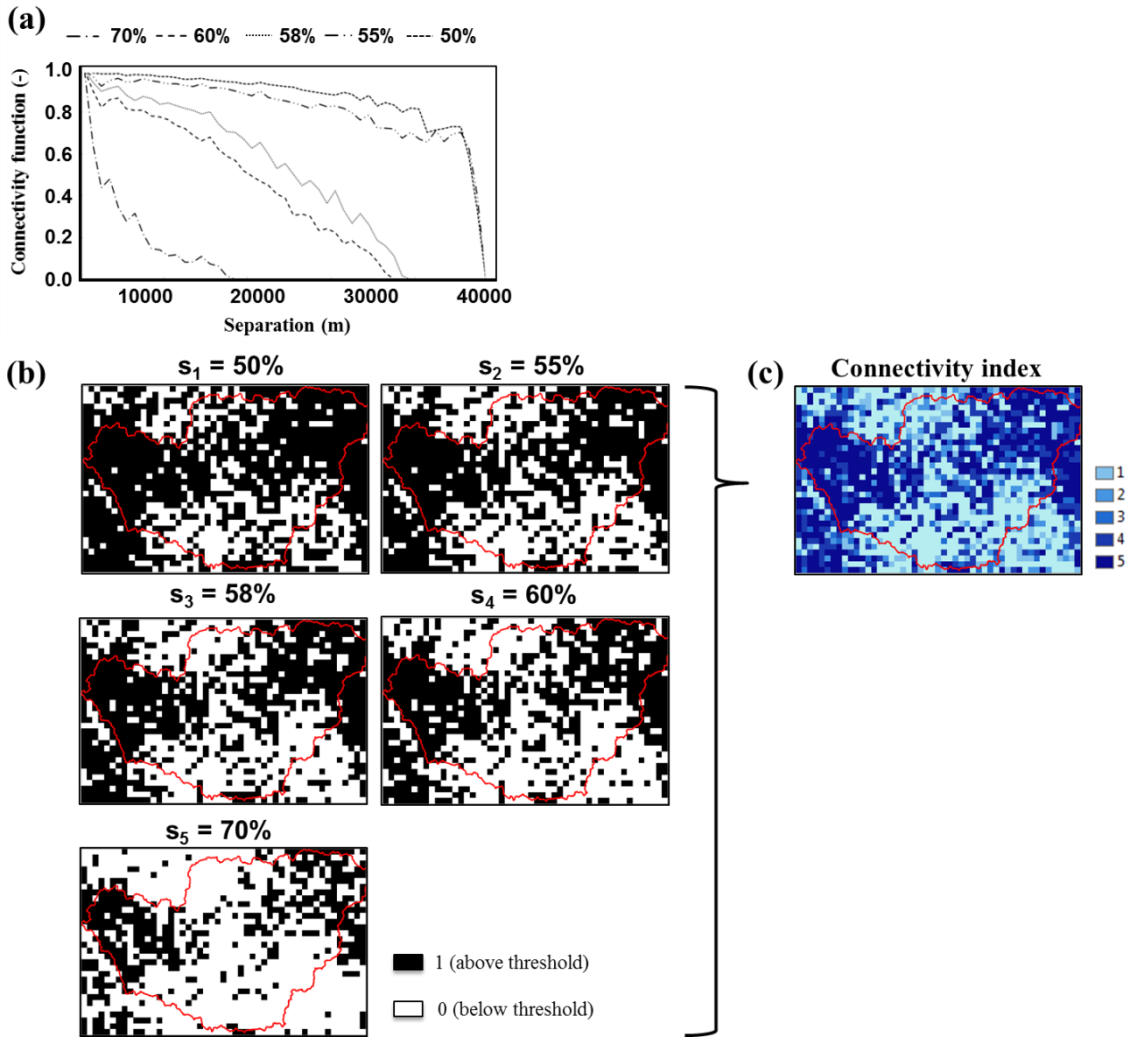

Figure 4. 4 Physically-based Hydrologic connectivity: (a) connectivity functions ($\tau(d)$) calculated using indicator maps ($I(y)$) of mixed physical controls, (b) indicator maps for 5 selected thresholds (s_i), and (c) physically-based hydrologic connectivity index map developed by integrating 5 indicator maps.

The connectivity index was developed for each soil layer of CLM using the collected soil information with depth, NDVI, and TI, and then used to calibrate soil hydraulic properties in land surface modeling. *Cosby et al.* [1984] developed a pedo-transfer function for estimating soil hydraulic properties (θ_{sat}, K_{sat}, ψ_{sat}, b) through a regression analysis using mean values of soil samples for various soil texture classes. The pedo-transfer function has been applied in CLM to model the soil parameters as a set of default parameters. However, the default parameters might not be enough to successfully describe the soil moisture distribution in all areas/regions because they were derived from the texture-based mean values of soil samples collected across the conterminous U.S. Thus, in this study, we calibrated the parameters within their possible ranges (implying various characteristics of sample sites such as texture, topography, vegetation, among others) by accounting for their standard deviation obtained in *Cosby et al.* [1984] study (Table 4.2) ($P_i' = P_{i,def} \pm \alpha_i$). P_i' is the calibrated parameter set (θ_{sat}, ψ_{sat}, b, K_{sat}); $P_{i,def}$ is the default parameter set; α_i is the calibrating factor for each parameter based on the standard deviation that is determined by the physically-based hydrologic connectivity index. The value of α is added to the default soil parameters of θ_{sat}, ψ_{sat}, and b and subtracted from the parameter of K_{sat} when the connectivity index is greater than 3 representing connected regions of physical controls and higher soil water content; on the contrary, for other regions that have a connectivity index less than 3 reverse operation is performed. The calibrated parameters based on physically-based hydrologic connectivity were applied in CLM to effectively predict spatially distributed

soil moisture. The model outputs using the default and calibrated parameters were compared to field observations.

4.4. Results and Discussion

In this study we investigated the effects of mixed dominant physical controls on soil moisture variability and developed the hydrologic connectivity algorithm to identify the spatial variations of soil moisture and improve the parameterization of soil hydraulic properties. The proposed approach was tested in two watersheds (LW and USS) and compared to airborne remote sensing near-surface soil moisture data (800 × 800 m). The physically-based hydrologic connectivity algorithm was applied to deeper soil layers as well as near surface soil layer. However, we compared to near-surface observations only because of the lack of soil moisture information for deeper soils at watershed scales.

4.4.1 Effects of Mixed Physical Controls on Soil Moisture Variability

The contributing ratios of the most dominant controls (up to 4) were derived using the Bayesian averaging scheme. Figure 4.5(a) and (b) shows the histograms of contributing ratios (w_1, w_2, w_3, and w_4) of the physical controls (e.g., %clay, %sand, NDVI, and TI) for the two study sites. For the LW watershed, NDVI (w_3 of 0.438), %clay (w_1 of 0.326), and %sand (w_2 of 0.235) represented higher contributions to soil moisture spatial distribution, while topography seldom contributes at the support scale of 800 × 800 m (Figure 4.5(a)).

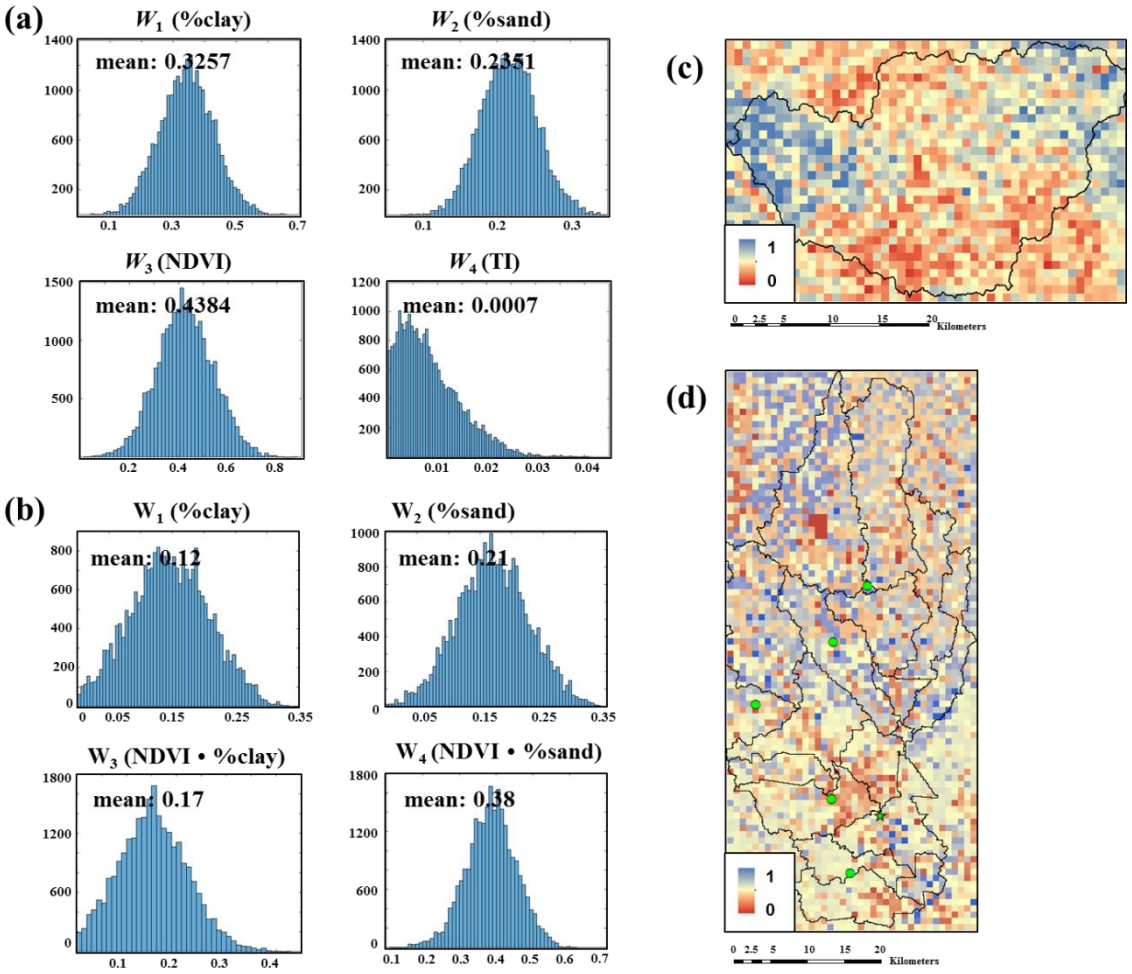

Figure 4. 5 Contributing ratios (weights, w_i) of physical controls ((a) for LW and (b) for USS) and mixed (weighted) physical controls maps ((c) for LW and (d) for USS).

The spatial distributions of soil texture and NDVI showed distinctive patterns across this watershed indicating that left and right hand sides of the watershed have higher values of %clay and NDVI and lower values of %sand corresponding to higher soil water content from ESTAR measurements. We also explored the effects of interactions between the physical controls (%clay · %sand, NDVI · %clay, NDVI

· %sand, TI · %clay, TI · %sand, NDVI · TI) on the contributions to soil moisture spatial variability. It was found that no significant contributions of the interactions existed in this watershed. It can be inferred that mixed response of individual physical controls based on their contributing ratios can predict the spatial variation of soil moisture well describing the distinctive patterns of landscape at the LW watershed. In addition, the spatial variation of mixed physical controls (normalized) was compared to the variability of measured saturated hydraulic conductivity (K_{sat}) from soil samples collected during the SGP97 hydrology experiment across the LW watershed. When K_{sat} measurements were rearranged according to soil types (CL - Clay Loam; L - Loam; LS - Loamy Sand; S – Sand; SiL - Silty Loam; SL - Sandy Loam), it showed high variations even for the same soil types representing a similar tendency as the variation of mixed physical controls with higher contribution of NDVI (Figure 4.6). This could be caused by other co-existing physical controls such as vegetation cover which may affect soil water flow, because of root distribution and organic matter content leading to different pore size distribution and water holding capacity in the unsaturated zone.

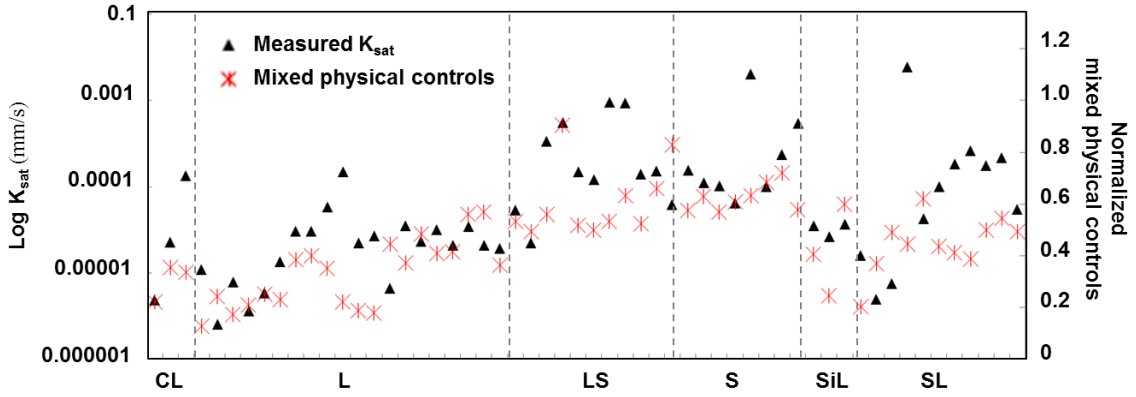

Figure 4. 6 Comparison of spatial variations of measured saturated hydraulic conductivity (K_{sat}) and normalized mixed physical controls (%clay, %sand, NDVI, TI) according to soil types across the LW watershed. (CL: Clay Loam, L: Loam, LS: Loamy Sand, S: Sand, SiL: Silty Loam, SL: Sandy Loam)

For the USS watershed, we found that contributing ratios of the dominant controls estimated using the Bayesian averaging scheme tend to be biased toward soil texture (0.50 and 0.33 for %clay and %sand, respectively) with no significant contributions of NDVI and TI. As with the results of LW, the surface topography showed no valid contribution at this support scale (800 × 800 m). On the other hand, when the interaction terms between the dominant controls (%clay · %sand, NDVI · %clay, NDVI · %sand, TI · %clay, TI · %sand, NDVI · TI) were included to account for the dependency of the physical controls, the interactions of %clay · NDVI and %sand · NDVI contributed significantly to the spatial distribution of soil moisture with resultant weights of 0.17 and 0.38, respectively (Figure 4.5(b)). It showed that NDVI influenced the description of the spatial variability of soil moisture as an interaction term with soil texture. In other words the mixed effects of interactions between physical controls as

well as the individual controls on soil moisture distribution can be a characteristic feature of larger and complex landscapes such as USS and LW watersheds.

Thus, we mixed the spatial patterns of physical controls based on their contribution ratios (w_1(%clay), w_2(%sand), w_3(NDVI), and w_4(TI) for LW watershed; w_1(%clay), w_2(%sand), w_3(NDVI · %clay), and w_4(NDVI · %sand) for USS watershed) (Figure 4.5(c) and (d)), and developed hydrologic connectivity maps.

4.4.2 Calibration of Soil Hydraulic Properties Based on Hydrologic Connectivity

Since soil moisture measurements with depth are not available at watershed scales, the contributing ratios of physical controls derived from near surface soil moisture were applied to combine the physical controls maps for deeper soil layers. The mixed physical controls maps for soil layers were used to calculate connectivity functions under various thresholds. The 5 representative thresholds (50%, 55%, 58%, 60%, and 70% for LW watershed and varying thresholds with depth for USS watershed) were found from connectivity functions for each soil layer that reflect connected patterns of the mixed physical controls well across the watersheds (Figure (4.7) and (4.8)). Using the 5 thresholds, the indicator maps were generated suggesting that the connectivity of mixed physical controls showed different patterns according to the thresholds which can reflect various spatial patterns of soil moisture in the unsaturated zone. In turn, the physically-based hydrologic connectivity index was developed by adding the indicator maps and quantifying the soil moisture variability. The hydrologic connectivity index

with soil depth was applied in calibrating the soil hydraulic properties in CLM, as depicted in Figure 4.2.

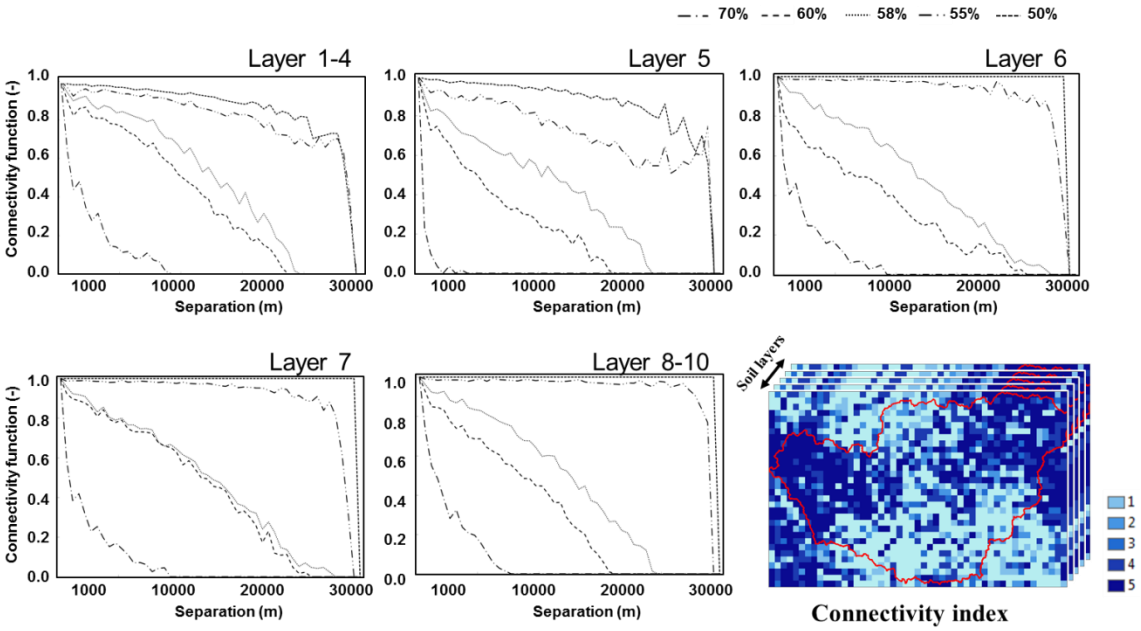

Figure 4. 7 Connectivity functions for 5 representative thresholds for soil layers and connectivity index for the LW watershed. Pixels of higher index represent highly connected and higher wetness regions; lower index pixels indicate unconnected and drier regions.

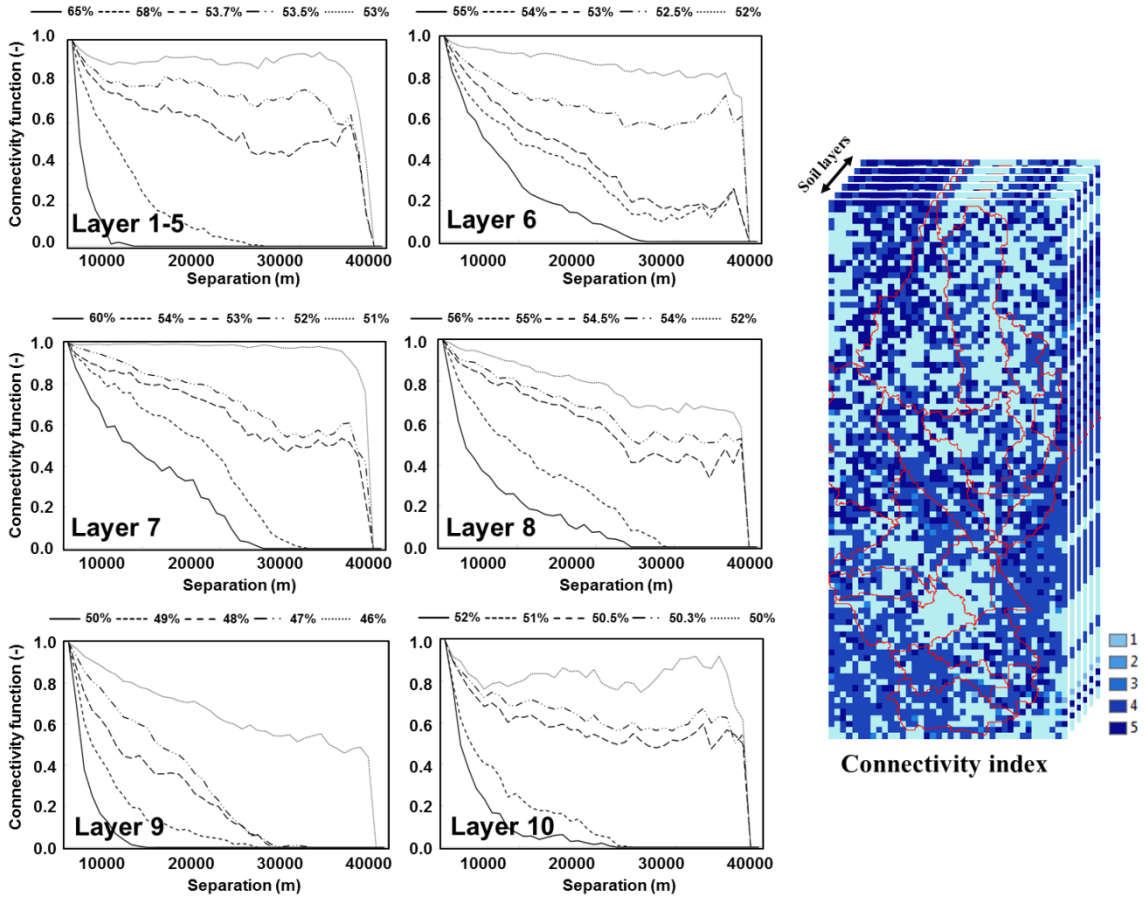

Figure 4. 8 Connectivity functions for 5 representative thresholds for soil layers and connectivity index for the USS watershed. Pixels of higher index represent highly connected and higher wetness regions; lower index pixels indicate unconnected and drier regions.

To analyze the spatial distributions of default and calibrated soil parameters, we selected a region which has complex landscape with relatively uniform soil types and heterogeneous vegetation cover and topography in the LW watershed. Figure 4.9 shows the comparison of spatial distributions of default and calibrated soil parameters. The default parameters have relatively uniform distributions depending on soil texture only (%sand and %clay) leading to low variation in soil moisture prediction. This is because CLM predicts soil hydraulic parameters from soil textural class alone. On the contrary, the parameters calibrated based on the physically-based hydrologic connectivity index showed the spatially distributed patterns across the region. Furthermore, the variations of default and calibrated soil hydraulic parameters were compared according to soil texture (Figure 4.10). The soil hydraulic parameters were uniformly predicted for the identical soil texture in the current model, while the calibrated parameters showed variations in space as shown in comparison of measured K_{sat} and mixed physical controls in Figure 4.6. Thus, the soil hydraulic parameters can be effectively calibrated using the hydrologic connectivity index to predict the variability of soil moisture in complex landscapes.

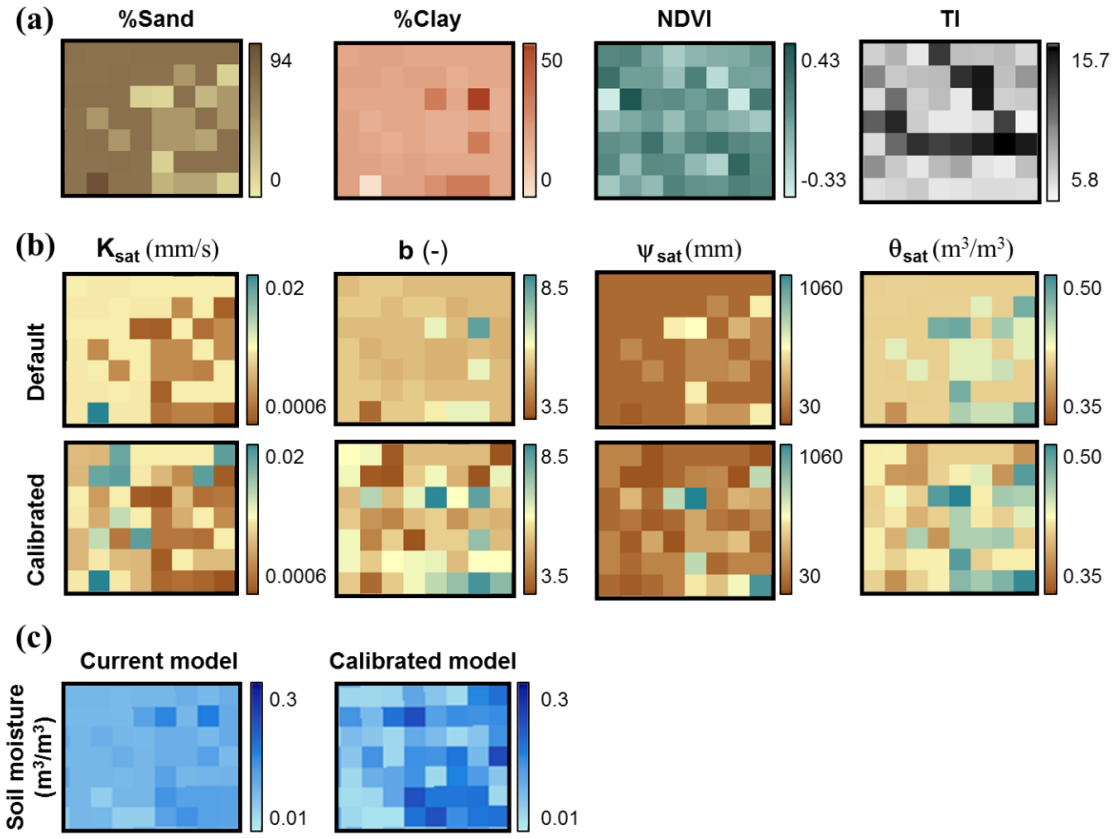

Figure 4. 9 (a) Dominant physical controls, (b) default and calibrated soil hydraulic parameters, and (c) soil moisture prediction in the current (CLM) and calibrated model for a selected region (800 × 800m resolution) in the LW watershed.

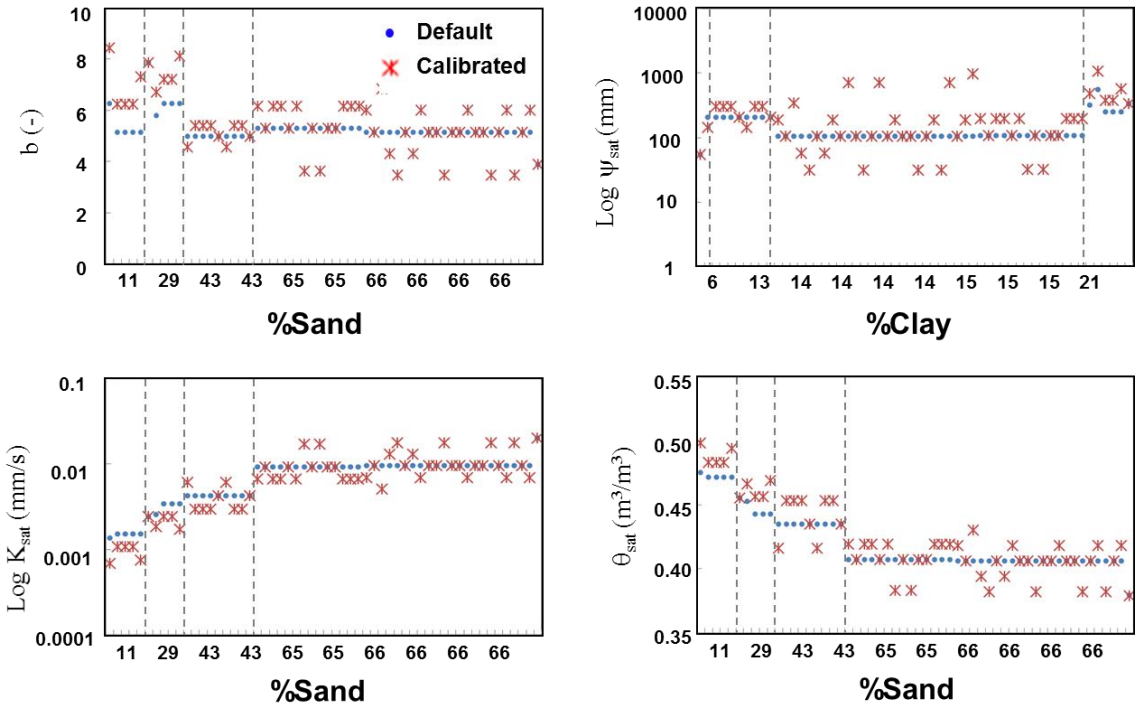

Figure 4. 10 Comparisons of default and calibrated soil parameters according to soil texture (%sand and %clay). θ_{sat}, K_{sat}, and b are estimated based on %Sand only; ψ_{sat}, is dependent on %Clay only in pedo-transfer function of CLM.

4.4.3 Comparison of CLM Output Using Default and Calibrated Soil Parameters

Model outputs (e.g., soil moisture, surface runoff, ET, and water storage) using the default and calibrated soil hydraulic properties were compared. Figure 4.11 shows the range of parameters based on their standard deviations in *Cosby et al.* [1984] study. The default and calibrated parameters were compared for the selected 5 pixels (Figure 4.1) which have various soil texture classes (loam, sandy clay loam, clay loam, and silty clay loam) and different vegetation in the LW watershed. After calibrating the parameters based on the physically-based hydrologic connectivity index, the soil

parameters of b, θ_{sat}, and ψ_{sat} were found to be higher than the default parameters, and K_{sat} lower than its default values in pixels 1, 2, and 3 located on connected pixels in the connectivity index map. In contrast, it showed lower b, θ_{sat}, ψ_{sat} and higher K_{sat} for the calibrated parameters than those of the defaults in pixels 4 and 5 which are on unconnected pixels. Several default and calibrated parameters were out of the ranges of parameters because the parameters were estimated with the pedo-transfer function derived through a regression analysis using mean values of soil samples. Using the calibrated parameters, the model can estimate higher soil water content in connected regions and lower soil water content in unconnected regions describing the spatially distributed soil moisture well across the LW watershed. Corroborating these findings, an improvement can be found by comparison of soil moisture dynamics on the selected pixels (Figure 4.12 (a) and (b)). On pixels 1, 2, and 3 (connected pixels), the current model with the default parameters tended to underestimate the soil moisture dynamics, while the model simulation using the calibrated parameters showed good agreement with ESTAR measurements (correlation coefficient of 0.81, 0.72, and 0.91; RMSE of 0.028, 0.054, and 0.029 in Figure 4.12(b)). On the other hand, the current model using the default parameters tended to overestimate the near surface soil moisture on pixels 4 and 5 (unconnected pixels) compared to the measurements. Although there are some uncertainties, the model prediction could be improved using the calibrated soil hydraulic parameters which match better with the measurements (correlation coefficient of 0.60 and 0.56; RMSE of 0.058 and 0.042 in Figure 4.12(b)) for the LW watershed.

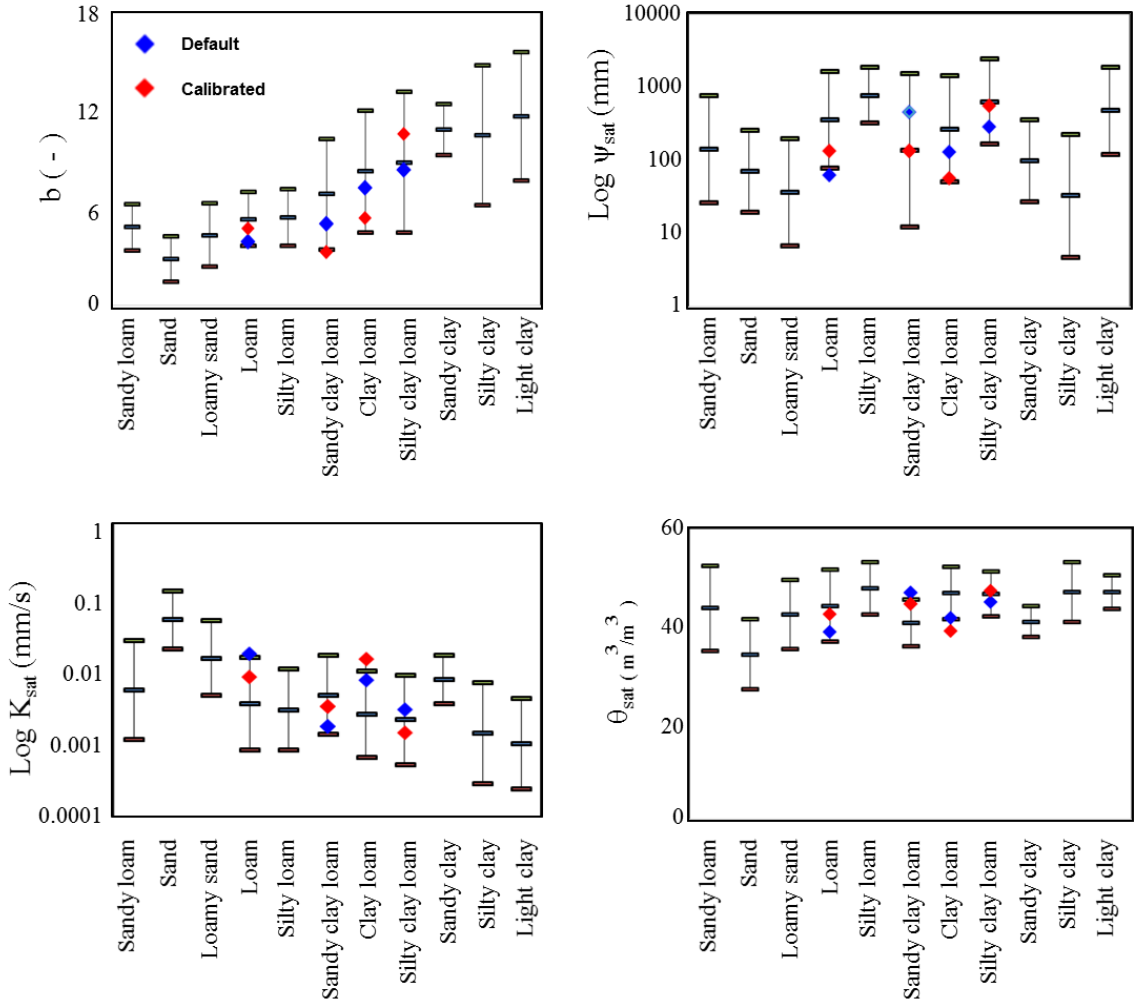

Figure 4. 11 Comparison of default parameters and calibrated parameters on 5 selected pixels for the LW watershed. The bar shows the ranges of parameters for 11 soil texture classes obtained from *Cosby et al.* [1984].

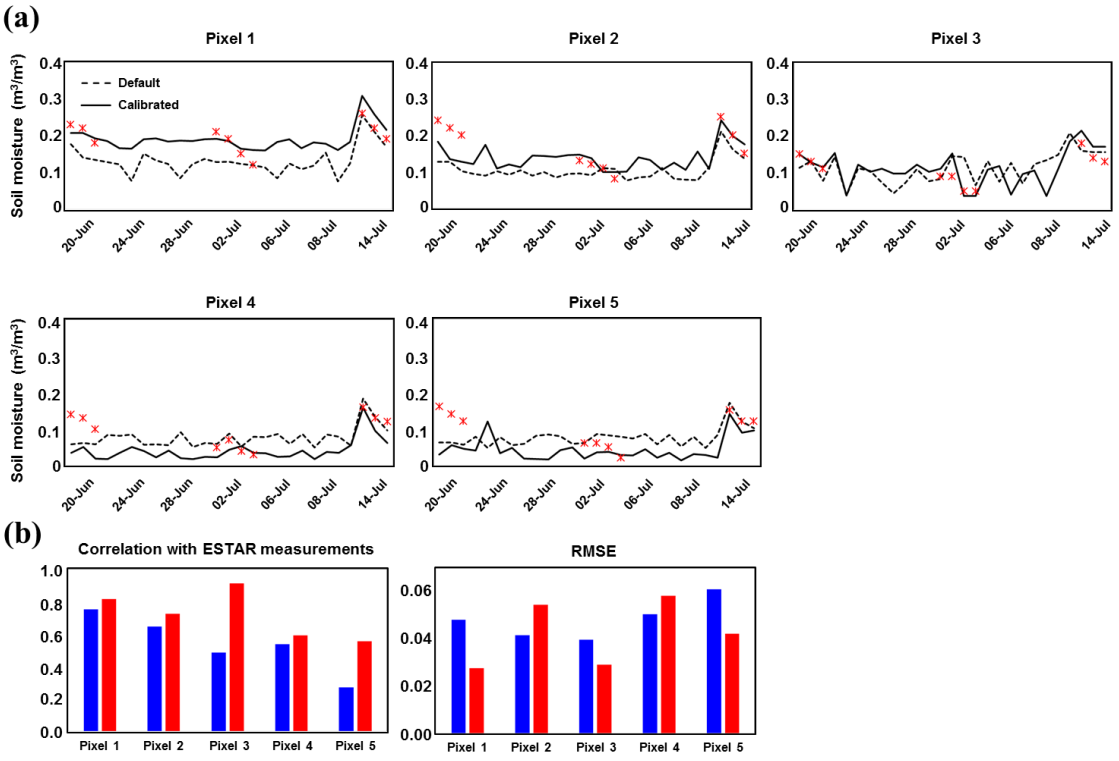

Figure 4. 12 (a) Comparisons of simulated and measured soil moisture dynamics on 5 pixels selected on connected and unconnected regions and (b) Correlation and RMSE with ESTAR measurements for 5 pixels in the LW watershed.

For the USS watershed, we compared the default and calibrated parameters on the selected 5 pixels (Figure 4.13), which were plotted on the ranges of parameters for various soil texture classes. As discussed above, after calibrating the parameters based

on the connectivity index it showed higher values of b, θ_{sat}, and ψ_{sat}, and lower K_{sat} on pixel 1 (Sand Clay Loam) and 5 (Loam) which are on connected pixels compared to the default parameters. On the other hand, lower values of b, θ_{sat}, and ψ_{sat}, and higher K_{sat} were assigned to the unconnected pixels (2 (Loam), 3 (Loam), and 4 (Clay Loam)). When we compared the simulated soil moisture dynamics using the default and calibrated parameters on the selected pixels, the improvement of model performance was found (Figure 4.14 (a)). Most of the pixels, except pixel 5, showed higher correlation with PSR measurements for the model output with the calibrated parameters, and RMSE was further reduced on all pixels (Figure 4.14(b)). Applying the calibrated parameters in land surface modeling, the parameters could make up for the default parameters' weaknesses which include underestimating the soil moisture dynamics on the connected regions and overestimating on the unconnected regions.

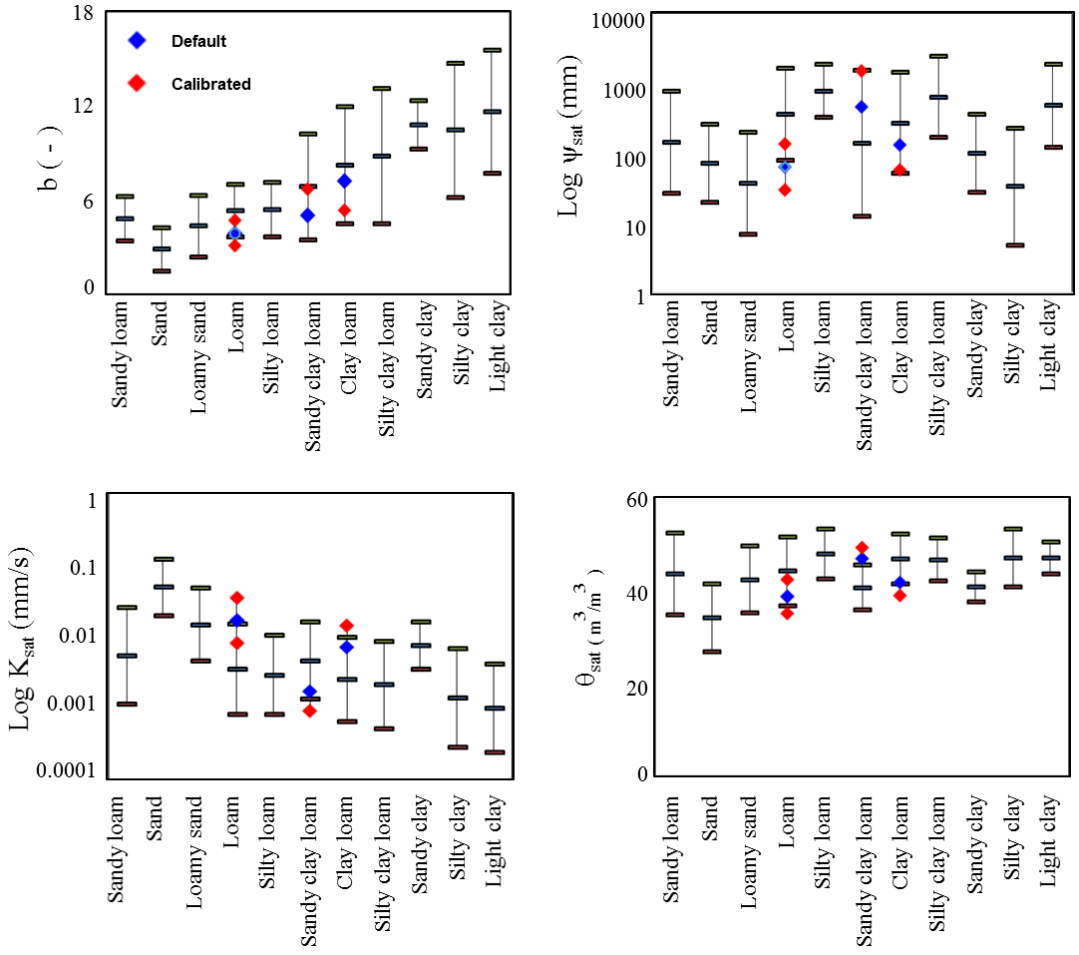

Figure 4. 13 Comparison of default parameters and calibrated parameters on 5 selected pixels for the USS watershed. The bar shows the ranges of parameters for 11 soil texture classes obtained from *Cosby et al.* [1984].

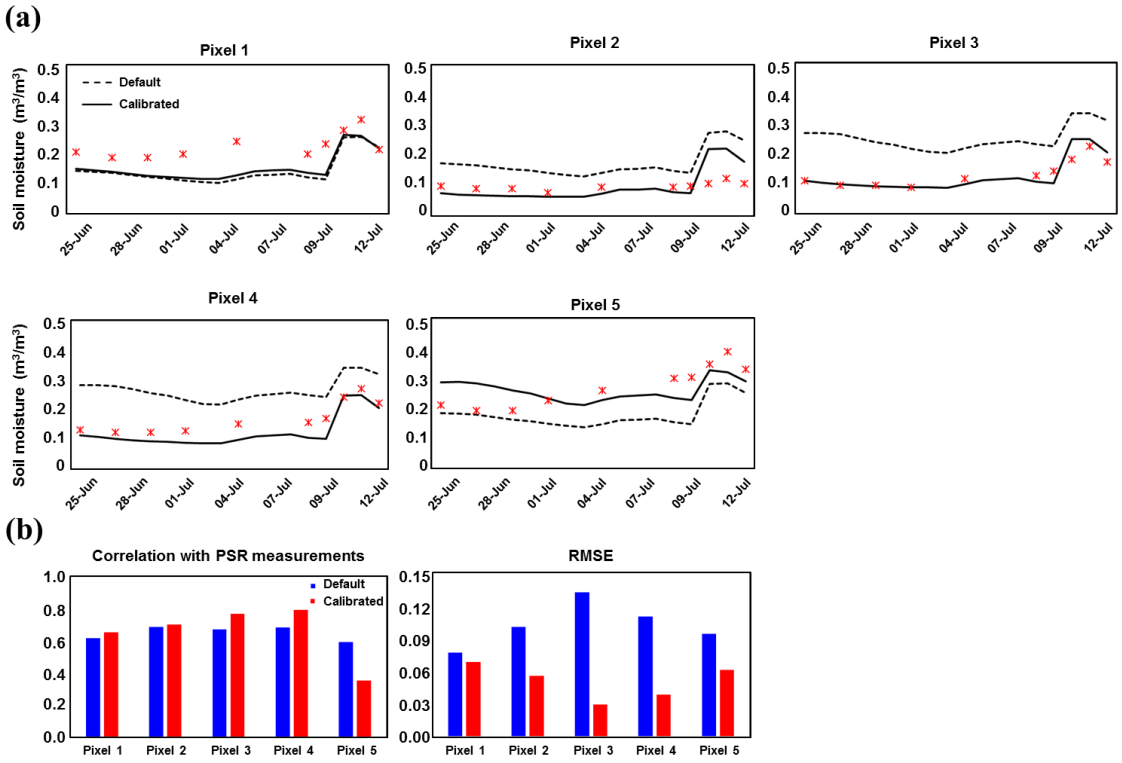

Figure 4. 14 (a) Comparisons of simulated and measured soil moisture dynamics on 5 pixels selected on connected and unconnected regions and (b) Correlation and RMSE with ESTAR measurements for 5 pixels in the USS watershed.

Figure 4.15 (a) and 4.16(a) show the comparisons of simulated near surface soil moisture and ESTAR measurement for the entire watersheds. The simulated soil moisture using the default parameters tends to be underestimated in wet regions (connected pixels) and overestimated in dry regions (unconnected pixels) that could not capture the variability of soil moisture due to the default soil parameters related to soil textural class alone in the watersheds. The calibrated model simulation matched well with the measurements (ESTAR and PSR) showing higher and lower soil water content on the connected and unconnected pixels, respectively. The spatial variability of soil

moisture prediction was compared at various extent scales. Table 3 shows the correlation coefficients and RMSE between measured soil moisture and model simulation (top 5cm) using the default and calibrated soil parameters for different extent scales within the watersheds. At all extent scales the calibrated model showed higher correlation coefficients (0.310 ~ 0.713 for LW and 0.400 ~ 0.712 for USS) and lower RMSE (0.016 ~ 0.048 for LW and 0.081 ~ 0.099 for USS) than that of the current model that represented improvements of model performance in space. Thus, the spatial variations of soil moisture can be properly described using soil parameters calibrated by physically-based hydrologic connectivity. Consequentially, these differences between the current and calibrated models can lead to different model outputs (e.g., root zone soil moisture, evapotranspiration, surface runoff, and water storage) as shown in Figure 15 ((b) – (e)) and 16((b) – (d)) that could have important effects not only on water cycle, but also on surface energy budgets.

Table 4. 3 Correlation coefficients and RMSE between soil moisture measurements and simulations (top 5cm) using default and calibrated soil parameters for the two study sites.

		LW				USS			
		2 × 7[a]	4 × 9[a]	8 × 13[a]	12 × 15[a]	24 × 4[a]	32 × 7[a]	36 × 11[a]	40 × 15[a]
R	Default	0.164	0.250	0.540	0.669	0.512	0.389	0.371	0.237
	Calibrated	0.310	0.452	0.674	0.713	0.712	0.587	0.488	0.400
RMSE	Default	0.048	0.047	0.051	0.050	0.128	0.123	0.115	0.102
	Calibrated	0.016	0.044	0.048	0.047	0.099	0.098	0.090	0.081

[a] Number of pixels (extent scale) – 1.6 × 1.6 km resolution, extent scales were determined by the shape of watersheds.

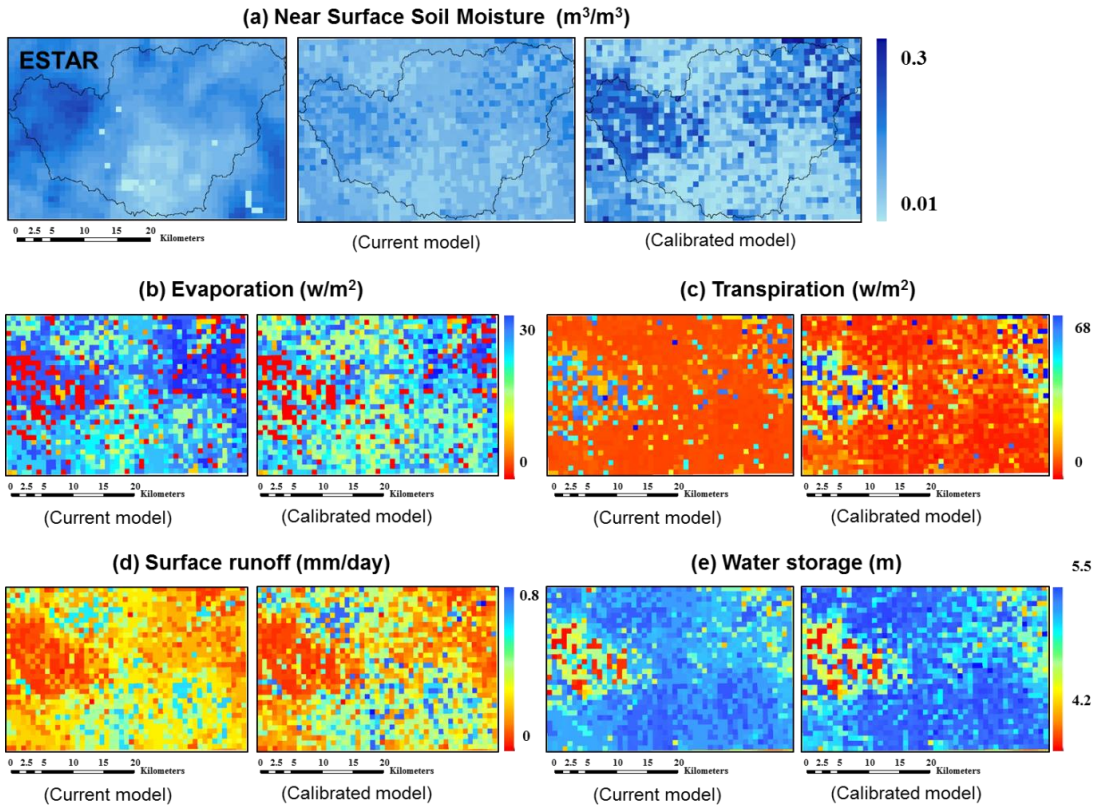

Figure 4. 15 (a) Comparisons of measured and simulated soil moisture and (b)-(e) model outputs using default and calibrated parameters (evaporation, transpiration, surface runoff, and water storage) for the LW watershed.

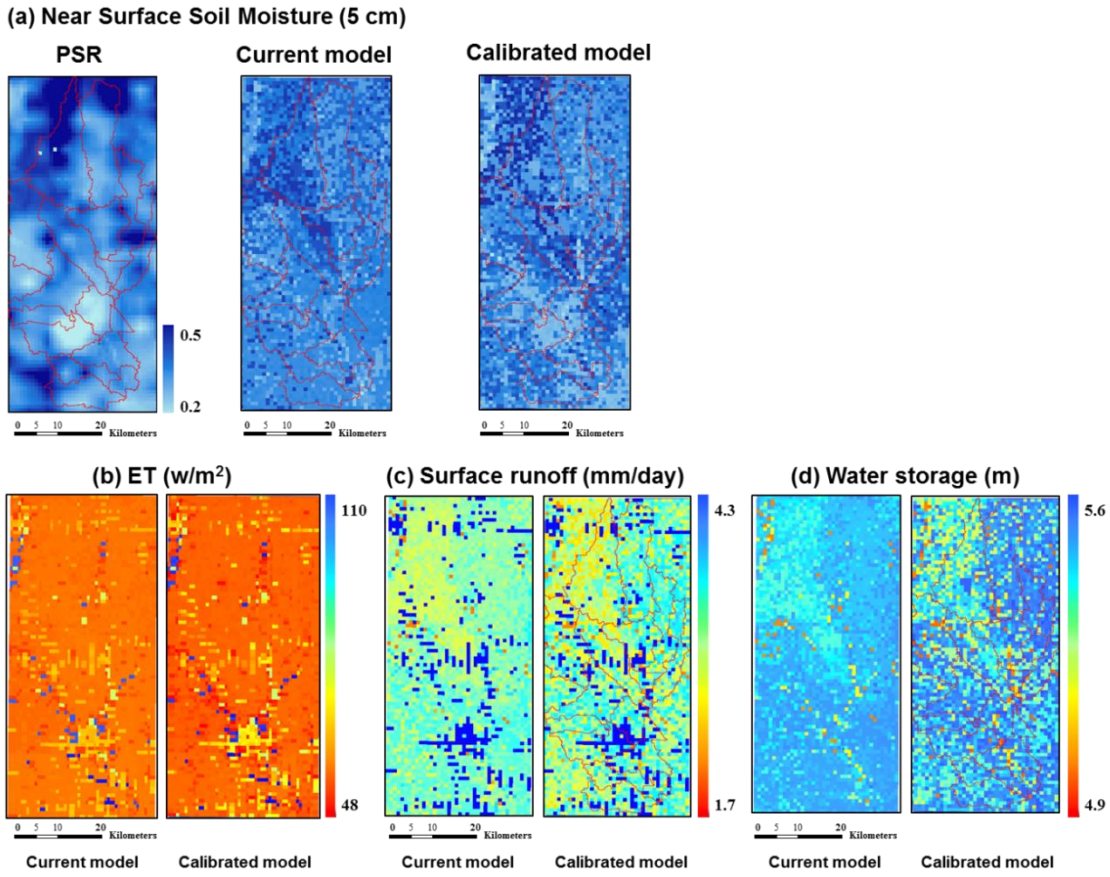

Figure 4. 16 (a) Comparisons of measured and simulated soil moisture and (b)-(d) model outputs using default and calibrated parameters (ET, surface runoff, and water storage) for the USS watershed.

Based on these findings, the physically-based hydrologic connectivity developed in this study helped to better understand the spatial variability of soil moisture in the unsaturated zone. Furthermore, the model performance using the calibrated soil hydraulic parameters based on the connectivity index was improved compared to the model predictions using the default parameters. It can be inferred that soil hydraulic

parameters calibrated with physically-based hydrologic connectivity can efficiently reflect the variations of soil moisture in space in land surface modeling at regional scales.

4.5. Conclusions

In this study, we developed a physically-based hydrologic connectivity algorithm to better understand catchment hydrologic characteristics and identify the soil moisture variability. To develop hydrologic connectivity based on dominant physical controls, the effects of mixed physical controls (e.g., topography, soil texture, vegetation) jointly on soil moisture spatial distribution were investigated at two different hydroclimate regions (sub-humid and humid climate). Using the Bayesian averaging scheme, the contributing ratios of physical controls to soil moisture distribution were derived to combine the controls for the two study sites. In the LW site, soil texture (%clay and %sand) and vegetation (NDVI) showed higher contributions and no significant contributions of interactions between the controls existed. On the other hand, soil texture and the interactions between vegetation and soil texture represented valid contributions to spatial patterns of soil moisture in the USS site. We found that the contributing ratios of physical controls could be site-specific depending on landscape characteristics, and the interaction terms of physical controls could also affect soil moisture distribution. Based on the contributing ratios, the dominant physical controls were combined and used for developing hydrologic connectivity using the integral connectivity scale technique. In order to identify the connectivity, we generated indicator maps using various thresholds selected from the connectivity functions. In turn, the physically-based hydrologic

connectivity index was developed by aggregating the indicator maps representing the connected (wet regions) and unconnected (dry regions) patterns across the watersheds, which can properly describe the soil moisture spatial variability.

The hydrologic connectivity index was applied in calibrating soil hydraulic properties (θ_{sat}, K_{sat}, ψ_{sat}, b) to improve the current parameterization in land surface modeling (CLM). When we compared the simulated soil moisture using the default and calibrated parameters to remote sensing measurements (ESTAR and PSR), the calibrated model simulation showed good agreement with the measurements. The simulated soil moisture dynamics on selected pixels were improved with the calibrated parameters indicating higher soil moisture prediction on the connected pixels and lower prediction on the unconnected pixels. Thus, using the physically-based hydrologic connectivity we could describe the spatial patterns of soil moisture and improve the current parameterization and model performance. Based on these results, the differences in model outputs using the default and calibrated soil parameters could have important effects not only on water cycle, but also on surface energy budgets. In general application, the physically-based hydrologic connectivity index can be applicable to other regions which have similar patterns of dominant physical controls for developing hydrologic connectivity using identical thresholds. For future work, since hydrologic connectivity patterns can vary with time, a dynamic connectivity index can be considered in the parameterization scheme to account for temporal variability of soil moisture in the unsaturated zone and improve model performance for long term simulation.

CHAPTER V

GENERAL CONCLUSIONS

The research in this dissertation was focused on improving hydrological model capabilities for better predictions of subsurface flow, especially soil water movement in the unsaturated zone at various spatial scales. Two main approaches (BMA-based multi-model simulation approach and physically-based hydrologic connectivity approach) were developed by bringing together advanced techniques and promising hydrologic concepts.

In past decade, many researchers have used a single hydrological model to simulate water and energy cycle. In Chapter II, strengths and weaknesses of hydrological models for near surface soil moisture predictions under various soil wetness conditions were investigated that caused by their inherent model parameterization and structures. Based on the finding in this research, hydrological modeling using a single model only is not enough to adequately describe real-world conditions (including all possible hydrological processes) that might cause uncertainties for a certain process and wetness condition. As shown in the results, the BMA-based multi-model simulation approach developed in this Chapter could reflect strengths of the models under various land surface wetness conditions and quantify the model structural uncertainties. Hence, the multi-model simulation approach can be useful to provide more robust and reliable model predictions than using a single model.

The physically-based hydrologic connectivity approach was introduced in Chapter III and IV to characterize the spatial variability of subsurface flow and improve subsurface flow processes and model parameterization in land surface modeling. In Chapter III, simplified processes of subsurface flow in land surface modeling were modified to include lateral subsurface flow processes based on complex topography and heterogeneous soil hydraulic properties. Hydrologic connectivity was used to derive spatially varying anisotropy for calculating lateral hydraulic conductivity that is the most challenging parameter in the lateral subsurface flow process. The modified land surface model was tested at field and sub-catchment scales. Results revealed that lateral subsurface flow plays an important role in redistributing soil water that might cause changes in the water and energy cycle, and hydrologic connectivity has the potential to effectively describe the lateral movement of subsurface flow. Furthermore, hydrologic connectivity can be derived from the combinations of dominant physical controls (e.g., soil properties, vegetation, topography), and it can be useful for characterizing the subsurface flow variability in the unsaturated zone. Based on these finding, a limitation of 1-dimentional land surface modeling can be overcome by considering the lateral subsurface flow component derived from the wetness- and physically-based hydrologic connectivity algorithm.

A limitation of the above study is that dominant physical variables equally contribute to soil moisture variability for developing hydrologic connectivity. To overcome it, the impacts of mixed physical controls (based on their contribution ratios) on soil moisture variability were explored in complex landscapes in Chapter IV. The

physical controls and interaction terms between them differently contributed to describing the spatially distributed soil moisture according to watershed characteristics. Physically-based hydrologic connectivity was developed based on the contributing ratios of variables in two watersheds which have different hydro-climatic conditions and watershed characteristics. The hydrologic connectivity patterns could effectively characterize the spatial variability of soil moisture in complex heterogeneous landscapes. Furthermore, the approach was successfully applied in calibrating soil hydraulic parameters in land surface modeling to improve the current model parameterization and eventually to enhance the land surface model capability. Currently, calibrating distributed model parameters in large watersheds is still challenging due to its insufficient parameterization based on a pedo-transfer function. As finding in this study, the distributed model parameters can be successfully calibrated through the physically-based hydrologic connectivity approach reflecting complex heterogeneous landscapes effectively.

Currently, various hydrological models have been widely used in agricultural management practices, flood/drought prediction, and climate forecast modeling, among others with critical deficiencies remaining in simulating land surface hydrology. Based on the knowledge and findings of this study, the approaches proposed in this study can be helpful to develop better hydrologic understanding and modeling capability in complex landscapes. Although the modeling capability was enhanced in this research, some limitations still remain in land surface modeling such as parameter scaling issue and the impacts of lateral subsurface flow at large scales (e.g., regional or global). This

research was focused on improving modeling capability at field, sub-catchment, and watershed scales. In future work, the current study can be extended by understanding the impacts of lateral flow at large scales through exploring soil moisture dynamics obtained from satellite remote sensing products. Furthermore, the current model parameterization developed from a pedo-transfer function using soil samples collected at a local scale can be improved to be applied in large scale land surface modeling by using an up-scaling approach based on a similarity concept. This similarity concept has been widely applied to characterize the spatial variability of model parameters that is useful to upscale soil hydraulic parameters by understanding hydrologic similarity at different spatial scales. The physically-based hydrologic connectivity algorithm based on various physical controls (e.g., soil type, vegetation, topographic configuration) developed in this research can be used for deriving hydrologic similarity in complex heterogeneous landscapes with scales. Hence, land surface modeling capacity will be enhanced by characterizing the spatial variability of model parameters and calibrating model parameters using the promising concept at various spatial scales.

REFERENCES

Agnese, C., F. Blanda, A. Drago, M. Iovino, M. Minacapilli, G. Provenzano, G. Rallo, and M. Sciortino (2007), Assessing the agro-hydrological SWAP model to simulate soil water balance in typical Mediterranean crops, *Geophys. Res. Abstr., 9*, 08146.

Ajami, N. K., Q. Y. Duan, and S. Sorooshian (2007), An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction, *Water Resour. Res.*, 43, W01403, doi:10.1029/2005WR004745.

Ali, G. A. and A. G. Roy (2009), Revisiting hydrologic sampling strategies for an accurate assessment of hydrologic connectivity in humid temperate systems, *Geogr. Compass, 3(1)*, 350–374, 10.1111/j.1749-8198.2008.00180.x.

Ali, G. A., and A. G. Roy (2010), Shopping for hydrologically representative connectivity metrics in a humid temperate forested catchment, *Water Resour. Res., 46*, W12544, doi:10.1029/2010WR009442.

Amoros, C., and G. Bornette (2002), Connectivity and biocomplexity in waterbodies of riverine floodplains, *Freshwater Biology 47 (4)*, pp. 761–776.

Assouline, S., and D. Or (2006), Anisotropy factor of saturated and unsaturated soils, *Water Resour. Res., 42*, W12403, doi:10.1029/2006WR005001.

Baroni, G. and S. Tarantola (2012), A general probabilistic framework for uncertainty and sensitivity analysis of deterministic models, International Environmental Modelling and Software Society (iEMSs), 2012 International Congress on

Environmental Modelling and Software Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany.

Benke K. K., K. E., Lowell, and A. J. Hamilton (2008), Parameter uncertainty, sensitivity analysis and prediction error in a water-balance hydrological model, *Math. Comput. Model.,* 47, 1134-1149.

Bindlish, R. and T. J. Jackson (2002), SMEX02 Aircraft Polarimetric Scanning Radiometer (PSR) Data. [indicate subset used]. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: 10.5067/4N9BFV091ZXH.

Bonan, G. B. (1996), A land surface model (LSM version 1.0) for ecological, hydrological, and atmospheric studies: Technical description and user's guide. NCAR Tech. Note NCAR/TN-417+STR, pp. 150.

Brocca, L., F. Melone, and T. Moramarco (2008), On the estimation of antecedent wetness conditions in rainfall-runoff modelling, *Hydrol. Process. 22*, 629-642, doi: 10.1002/hyp.6629.

Brocca, L., F. Melone, T. Moramarco, and R. Morbidelli (2010), Spatial-temporal variability of soil moisture and its estimation across scales, *Water Resour. Res., 46*, W02516, doi:10.1029/2009WR008016.

Carroll, D. L. (1998), GA Fortran Driver version 1.7, Available from: http://www.cuaerospace.com/carroll/ga.html.

Caswell, H. (2000), Prospective and retrospective perturbation analyses: their roles in conservation biology, *Ecology 81,* 619-627.

Chen, J. and P. Kumar (2001), Topographic influence on the seasonal and inter-annual variation of water and energy balance of basins in North America. *J. Clim, 14,* 1989-2014.

Choi, H. I., P. Kumar, and X.-Z. Liang (2007), Three-dimensional volume-averaged soil moisture transport model with a scalable parameterization of subgrid topographic variability, *Water Resour. Res., 43,* W04414, doi:10.1029/2006WR005134.

Clapp, R. B. and G. M. Hornberger (1978), Empirical equations for some soil hydraulic properties, *Water Resour. Res., 14*(4), 601–604, doi:10.1029/WR014i004p00601.

Coccia, G. & E. Todini (2011), Recent developments in predictive uncertainty assessment based on the model conditional processor approach. *Hydrology and Earth System Sciences,* 15(10): 3253-3274.

Collins, W. D., P. J. Rasch, B. A. Boville, J. J. Hack, J. R. McCaa, D. L. Williamson, B. P. Briegleb, C. M. Bitz, S.-J. Lin, and M. Zhang (2006), The Formulation and Atmospheric Simulation of the Community Atmosphere Model Version 3 (CAM3). *J. Clim. 19,* 2144-2161.

Cosby, B. J., G. M. Hornberger, R. B. Clapp, and T. R. Ginn (1984), A statistical exploration of the relationship of soil moisture characteristics to the physical properties of soils, *Water Resour. Res., 20(6),* 682-690.

Dabney, S. M. and H. M. Selim. (1987), Anisotropy of a fragipan soil: vertical vs. horizontal hydraulic conductivity. *Soil Sci. Soc. Am. J.* 51:3-6.

de Kroon, H., A. Plaisier, J. van Groenendael, and H. Caswell (1986), Elasticity: the relative contribution of demographic parameters to population growth rate, *Ecology 67*, 1427-1431.

Dickinson, R. E., A. Henderson-Sellers, and P. J. Kennedy (1993), Biosphere atmosphere transfer scheme (BATS) version 1e as coupled to the NCAR Community Climate Model, *NCAR Tech. Note, NCAR/TN-378+STR*, Natl. Cent. For Atmos. Res., Boulder, Colo..

D'Odorico, P., K. Caylor, G. S. Okin, and T. M. Scanlon (2007), On soil moisture-vegetation feedbacks and their possible effects on the dynamics of dryland ecosystem, *J. Geophys. Res., 112*, G04010, doi:10.1029/2006JG000379.

Duan, Q., N. K. Ajami, X. Gao, and S. Sorooshian (2007), Multi-model ensemble hydrologic prediction using Bayesian model averaging, *Adv. Water Resour., 30*, 1371-1386.

Duan, Q., S. Sorooshian, and V. Gupta (1992), Effective and efficient global optimization for conceptual Rainfall-Runoff models, *Water Resour. Res., 28*(4), 265–284, doi: 10.1029/91WR02985

Duan, Q. and T. J. Phillips (2010), Bayesian estimation of local signal and noise in multimodel simulations of climate change, *J. Geophys. Res.*, 115, D18123, doi:10.1029/2009JD013654.

Ek, M., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advances in the

National Centers for Environmental Prediction operational mesoscale Eta Model, *J. Geophys. Res., 108*(D22), 8851, doi:10.1029/2002JD003296.

Ek, M., Y. Xia, E. F. Wood, J. Sheffield, B. Cosgrove, and K. Mo (2011), The North American Land Data Assimilation System: Application to Drought over CONUS, *AGU*, Fall Meet., Abstract GC31A-1015.

Emanuel, R. E., A. G. Hazen, B. L. McGlynn, and K. G. Jencso (2014), Vegetation and topographic influences on the connectivity of shallow groundwater between hillslopes and streams, *Ecohydrol., 7*, doi:10.1002/eco.1409.

Famiglietti, J. S., and E. F. Wood (1994), Multiscale modeling of spatially variable water and energy balance processes, *Water Resour. Res., 30*(11), 3061-3078.

Famiglietti, J., J. Devereaux, C. A. Laymon, T. Tsegaye, P. R. Houser, T. J. Jackson, S. T. Graham, M. Rodell, and P. J. van Oevelen (1999), Ground-based investigation of soil moisture variability within remote sensing footprints during the Southern Great Plains 1997 (SGP 97) Hydrology Experiment, *Water Resour. Res., 35(6)*, 1839–1851.

Fan, Y., G. Miguez-Macho, C. P. Weaver, R. Walko, and A. Robock (2007), Incorporating water table dynamics in climate modeling: 1. Water table observations and equilibrium water table simulations, *J. Geophys. Res., 112*, D10125, doi:10.1029/2006JD008111.

Flügel, W. A., (1995), Delineating Hydrological Response Units (HRUs) by GIS analysis regional hydrological modelling using PRMS/ MMS in the drainage basin of the River Bröl, Germany, *Hydrological Processes, 9*, pp. 423-436.

Gaur, N., and B. P. Mohanty (2013), Evolution of physical controls for soil moisture in humid and subhumid watersheds, *Water Resour. Res., 49*, doi:10.1002/wrcr.20069.

Georgakakos, K. P., D. J. Seo, H. Gupta, J. Schake, and M .B. Butts (2004), Characterizing streamflow simulation uncertainty through multimodel ensembles. *J. Hydrol., 298(1-4)*, 222–41.

Gneiting, T., A. E. Raftery, A. H. Westveld, and T. Goldman (2005), Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Monthly Weather Review* 133, 1098-1118.

Gochis, D. J. and F. Chen (2003), Hydrological enhancements to the community Noah land surface model. NCAR Technical Note, NCAR/TN-454+STR, 68 pgs.

Hatton, T. J. (1998), *The Basics of Recharge and Discharge*, part 4, *Catchment Scale Recharge Modeling*, Commonw. Sci. and Ind. Res. Organ., Collingwood, Victoria, Australia.

Hoetting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky (1999), Bayesian modeling averaging: A tutorial, *Stat. Sci.*, 14(4), 382-417.

Holland, J. H. (1975), *Adaptation in natural and artificial systems*, 183 pp., Univ. of Michigan Press, Ann Arbor, MI.

Hong, S., V. Lakshmi, E. E. Small, F. Chen, M. Tewari, and K. W. Manning (2009), Effects of vegetation and soil moisture on the simulated land surface processes from the coupled WRF/Noah model, *J. Geophys. Res., 114*, D18118, doi:10.1029/2008JD011249.

Hsu, K. L., H. Moradkhani, and S. Sorooshian (2009), A sequential Bayesian approach for hydrologic model selection and prediction, *Water Resour. Res., 45*, W00B12, doi:10.1029/2008WR006824.

Hupet, F., S. Lambot, M. Javaux, and M. Vanclooster (2002), On the identification of macroscopic root water uptake parameters from soil water content observations, *Water Resour. Res., 38(12),* 1300, doi:10.1029/2002WR001556.

Huang, M., Z. Hou, L. R. Leung, Y. Ke, Y. Liu, Z. Fang, and Y. Sun (2013), Uncertainty Analysis of Runoff Simulations and Parameter Identifiability in the Community Land Model: Evidence from MOPEX Basins. *J. Hydrometeor, 14*, 1754–1772. doi: http://dx.doi.org/10.1175/JHM-D-12-0138.1.

Hwang, T., L. Band, and T. C. Hales (2009), Ecosystem processes at the watershed scale: Extending optimality theory from plot to catchment, *Water Resour. Res., 45,* W11425, doi:10.1029/2009WR007775.

Hwang, T., L. E. Band, J. M. Vose, and C. Tague (2012), Ecosystem processes at the watershed scale: Hydrologic vegetation gradient as an indicator for lateral hydrologic connectivity of headwater catchments, *Water Resour. Res., 48,* W06514, doi:10.1029/2011WR011301.

Ines, A. V. M. and B. P. Mohanty (2008), Near-surface soil moisture assimilation for quantifying effective soil hydraulic properties using genetic algorithm: I. Conceptual modeling, *Water Resour. Res., 44*, W06422, doi:10.1029/2007WR005990.

Ines, A. V. M. and B. P. Mohanty (2009), Near-surface soil moisture assimilation for quantifying effective soil hydraulic properties using genetic algorithm (II. Using airborne remote sensing during SGP97 and SMEX02), *Water Resour. Res., 45*, W01408, doi:10.1029/2008WR007022.

Ines, A. V. M. and P. Droogers (2002), Inverse modeling in estimating soil hydraulic functions: a genetic algorithm approach, *Hydrol. Earth Sys. Sci. 6* (1): 49–65.

Jackson, T. J., D. M. Le Vine, A. Y. Hsu, A. Oldak, P. J. Starks, C. T. Swift, J. D. Isham, and M. Haken (1999), Soil moisture mapping at regional scales using microwave radiometry: The Southern Great Plains hydrology experiment, *IEEE Trans. Geosci. Remote Sens., 37*, 2136-2151.

Jackson, T. J., M. H. Cosh, J. Famiglietti, and R. Bindlish (2009), SMEX04 Regional Soil Moisture Data: Arizona, Digital media, Natl. Snow and Ice Data Cent., Boulder, Colo..

Jana, R. B., and B. P. Mohanty (2012a), On topographic controls of soil hydraulic parameter scaling at hillslope scales, *Water Resour. Res., 48*, W02518, doi:10.1029/2011WR011204.

Jana, R. B., and B. P. Mohanty (2012b), A topography-based scaling algorithm for soil hydraulic parameters at hillslope scales: Field testing, *Water Resour. Res., 48*, W02519, doi:10.1029/2011WR011205.

Jana, R. B., and B. P. Mohanty (2012c), A comparative study of multiple approaches to soil hydraulic parameter scaling applied at the hillslope scale, *Water Resour. Res., 48*, W02520, doi:10.1029/2010WR010185.

Jencso, K. G., B. L. McGlynn, M. N. Gooseff, S. M. Wondzell, K. E. Bencala, and L. A. Marshall (2009), Hydrologic connectivity between landscapes and streams: Transferring reach- and plot-scale understanding to the catchment scale, *Water Resour. Res., 45,* W04428, doi:10.1029/2008WR007225.

Jencso K. G., B. L. McGlynn, M. N. Gooseff, K. E. Bencala, and S. M. Wondzell (2010), Hillslope hydrologic connectivity controls riparian groundwater turnover: implications of catchment structure for riparian buffering and stream water sources, *Water Resour. Res., 46(11),* W10524. DOI:10.1029/2009WR008818.

Jencso, K. G., and B. L. McGlynn (2011), Hierarchical controls on runoff generation: Topographically driven hydrologic connectivity, geology, and vegetation, *Water Resour. Res., 47*, W11527, doi:10.1029/2011WR010666.

Joshi, C., and B. P. Mohanty (2010), Physical controls of near surface soil moisture across varying spatial scales in an agricultural landscape during SMEX02, *Water Resour. Res., 46,* W12503, doi:10.1029/2010WR009152.

Katata, G., H. Nagai, H. Ueda, N. Agam, and P. R. Berliner (2007), Development of a land surface model including evaporation and adsorption processes in the soil for the land-air exchange in arid regions, *J. Hydromet., 8,* doi:10.1175/2007JHM829.1.

Keefer, T., D. Goodrich, and S. Moran (2009), SMEX04 Meteorological Network Data: Arizona. Boulder, Colorado USA: NASA DAAC at the National Snow and Ice Data Center.

Kennedy J. and R. C. Eberhart (2001), *Swarm Intelligence*, Morgan Kaufmann, San Mateo, CA.

Kim, J., and B. P. Mohanty (2016), Influence of lateral subsurface flow and connectivity on soil water storage in land surface modeling, *J. Geophys. Res. Atmos., 121*, doi:10.1002/2015JD024067.

Koren, V., J. Schaake, K. Mitchell, Q. Y. Duan, F. Chen, and J. M. Baker (1999), A parameterization of snowpack and frozen ground intended for NCEP weather and climate models, *J. Geophys. Res., 104*, 19 569-19 585.

Koster, R. and M. Suarez (1996), Energy and water balance calculations in the Mosaic LSM, *NASA Tech. Memo. 9*, 104606.

Krishnakumar, K. (1989), *Micro-Genetic Algorithms for stationary and non-stationary function optimization*, SPIE: Intelligent Control and Adaptive Systems, Vol. 1196, pp. 289-296, Philadelphia, PA.

Kroes, J. G., J. G. Van Dam, J. Huygen, and R. W. Vervoort (1998), User's Guide of SWAP version 2.0, *Tech. Doc. 48*, DLO Winand Staring Cent., Wageningen Agric. Univ., Wageningen, Netherlands.

Krzysztofowicz, R. (1999), Bayesian Theory of Probabilistic Forecasting via Deterministic Hydrologic Model, *Water Resour. Res.*, 35, 2739–2750.

Krzysztofowicz, R. and K. S. Kelly (2000), Hydrologic uncertainty processor for probabilistic river stage forecasting. *Water Resour. Res.*, 36(11), 3265-3277

Kumar, P (2004), Layer averaged Richard's equation with lateral flow, *Adv. Water Resour., 27(5)*, 521-531.

Larsen, L. G., J. Choi, M. K. Nungesser, and J. W. Harvey (2012), Direction connectivity in hydrology and ecology. *Ecological Applications, 22(8)*, 2204–2220.

Leij, F. J., W. J. Alves, M. T. van Genuchten, and J. R. Williams (1999), The UNSODA unsaturated soil hydraulic database, in *Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media*, edited by M. T. van Genuchten et al., pp. 1269-1281, Univ. of Calif., Riverside, CA.

Leung, L. R., M. Huang, Y. Qian, and X. Liang (2011), Climate-soil-vegetation control on groundwater table dynamics and its feedbacks in a climate model, *J. Clim., 36*: 57-81.

Li, H., M. Huang, M. S. Wigmosta, Y. Ke, A. M. Coleman, L. R. Leung, A. Wang, and D. M. Ricciuto (2011), Evaluating runoff simulations from the Community Land Model 4.0 using observations from flux towers and a mountainous watershed, *J. Geophys. Res., 116,* D24120, doi:10.1029/2011JD016276.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land surface water, energy fluxes for general circulation models, *J. Geophys. Res., 99* (D7), 14415– 14428, doi:10.1029/94JD00483.

Liu, Y., H. V. Gupta, S. Sorooshian, L. A. Bastidas, and W. J. Shuttleworth (2004), Exploring parameter sensitivities of the land surface using a locally coupled land-atmosphere model, *J. Geophys. Res., 109*, D21101, doi:10.1029/2004JD004730.

Liu, Y. Y., M. F. McCabe, J. P. Evans, A. I. J. M. van Dijk, R. A. M. de Jeu, and H. Su (2009), Comparison of soil moisture in GLDAS model simulations and satellite observations over the Murray Darling Basin in *18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*, edited by R. S.

Anderssen, et al., pp. 2798-2804, Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, Cairns.

Lu, N., B. S. Kaya, and J. W. Godt (2011), Direction of unsaturated flow in a homogeneous and isotropic hillslope, *Water Resour. Res., 47*, W02519, doi:10.1029/2010WR010003.

MacQueen, J. (1967), Some methods for classification and analysis of multivariate observations in *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, pp. 281-297, Univ. of Calif. Press, Berkeley, CA.

Mayor, Á . G., S. Bautista, E. E. Small, M. Dixon, and J. Bellot (2008), Measurement of the connectivity of runoff source areas as determined by vegetation pattern and topography: A tool for assessing potential water and soil losses in drylands, *Water Resour. Res., 44*, W10423, doi:10.1029/2007WR006367.

Maxwell, R. M., and N. L. Miller (2005), Development of a coupled land surface and groundwater model, *J. Hydrometeorol.*, 6(3), 233-247.

Maxwell, R. M., and S. J. Kollet (2008), Quantifying the effects of three-dimensional subsurface heterogeneity on Hortonian runoff processes using a coupled numerical, stochastic approach, *Adv. Water Resour., 31*, 807-817, doi:10.1016/j.advwatres.2008.01.020.

McCord, J. T., and D. B. Stephens (1987), Lateral moisture flow beneath a sandy hillslope without an apparent impeding layer, *Hydrol. Processes, 1*, 225-238.

McDonnell, J. J., et al. (2007), Moving beyond heterogeneity and process complexity: a new vision for watershed hydrology, *Water Resour. Res., 43 (7)*; doi: 10.1029/2006WR005467.

McGarigal, K., S. A. Cushman, M. C. Neel, and E. Ene (2002), FRAGSTATS: spatial pattern analysis program for categorical maps, University of Massachusetts, Amherst, Massachusetts, USA. http://www.umass.edu/landeco/research/fragstats/fragstats.html.

Mohanty, B. P., P. J. Shouse, D. A. Miller, and M. T. Van Genuchten (2002), Soil property database: Southern Great Plains 1997 hydrology experiment, *Water Resour. Res., 38(5),* 1047, doi:10.1029/2000WR000076.

Mohanty, B. P. and T. H. Skaggs (2001), Spatio-temporal evolution and time-stable characteristics of soil moisture within remote sensing footprints with varying soil, slope, and vegetation. *Adv. Water Resour.* 24: 1051-1067.

Mohanty, B. P., D. A. Miller, and M. T. van Genuchten (2002), Soil property database: Southern Great Plains 1997 Hydrology Experiment, *Water Resour. Res., 38*(5): 10.1029/2000WR000076.

Mualem, Y., (1976), A new model for predicting the hydraulic conductivity of unsaturated porous media, *Water Resour. Res., 12*: 513-522.

Mueller, E. N., J. Wainwright, and A. J. Parsons (2007), Impact of connectivity on the modeling of overland flow within semiarid shrubland environments, *Water Resour. Res., 43,* W09412, doi:10.1029/2006WR005006.

Musters, P. A. D., W. Bouten, and J. M. Verstraten (2000), Potentials and limitations of modelling vertical distributions of root water uptake of an Austrian pine forest on a sandy soil, *Hydrol. Process. 14,* 103-115.

Narasimhan, B., R. Srinivasan, J. G. Arnold, and M. Di Luzio (2005), Estimation of long-term soil moisture using a distributed parameter hydrologic model and verification using remotely sensed data, *Transactions of the ASAE 48 (3),* 1101-1113.

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, and L. E. Gulden (2005), A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in GCMs, *J. Geophys. Res., 110,* D21106, doi:10.1029/2005JD006111.

Niu, G. Y., Z. L. Yang, R. E. Dickinson, L. E. Gulden, and H. Su (2007), Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, *J. Geophys. Res., 112*:D07103. DOI:10.1029/2006JD007522.

Niu, G.-Y., and Z. -L. Yang (2007), An observation-based formulation of snow cover fraction and its evaluation over large North American river basins. *J. Geophys. Res. 112,* D21101. DOI:10.1029/2007JD008674.

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, L. E. Gulden, and H. Su (2007), Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, *J. Geophys. Res., 112*:D07103. DOI:10.1029/2006JD007522.

Oleson., K. W., D. M. Lawrence, G. B. Bonan, M. G. Flanner, E. Kluzek, P. J. Lawrence, S. Levis, S. C. Swenson, and P. E. Thornton (2010), Technical Description of version 4.0 of the Community Land Model (CLM), *NCAR Tech. Notes (NCAR/TN-478+STR)*, 257 pp.

Oleson, K. W., G.-Y. Niu, Z.-L. Yang, D. M. Lawrence, P. E. Thornton, P. J. Lawrence, R. Stockli, R. E. Dickinson, G. B. Bonan, S. Levis, A. Dai, and T. Qian (2008), Improvements to the Community Land Model and their impact on the hydrological cycle, *J. Geophys. Res., 113*, G01021, doi:10.1029/2007JG000563

Philip, J. R. and D. de Vries (1957), Moisture movement in porous materials under temperature gradients, *Trans. Am. Geophys. Union, 38*: 222-232.

Qian, T., A. Dai, K. E. Trenberth, and K. W. Oleson (2006), Simulation of global land surface conditions from 1948 to 2004: Part I: Forcing data and evaluations. *J. Hydrometeor. 7*, 953-975.

Radell, D. B. and C. M. Rowe (2008), An observational analysis and evaluation of land surface model accuracy in the Nebraska Sand Hills, *J. Hydromet., 9*, DOI: 10.1175/2007JHM913.1.

Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski (2005), Using Bayesian model averaging to calibrate forecast ensembles, *Mon. Weather Rev., 133*:1155–1174.

Reed, P., B. Minsker, and D. E. Goldberg (2000), Designing a competent simple genetic algorithm for search and optimization, *Water Resour. Res., 36*(12), 3757–3761, doi:10.1029/2000WR900231.

Ricotta, C., A. Stanisci, G. C. Avena, and C. Blasi (2000), Quantifying the network connectivity of landscape mosaics: a graph-theoretical approach, *Community Ecology 1*: 89–94.

Ritter, A., F. Hupet, R. Munoz-Carpena, S. Lambot, and M. Vanclooster (2003), Using inverse methods for estimating soil hydraulic properties from field data as an alternative to direct methods, *Agricultural Water Management, 59,* 77-96.

Rojas, R., L. Feyen, and A. Dassargues (2008), Conceptual model uncertainty in groundwater modeling: Combining generalized likelihood uncertainty estimation and Bayesian model averaging, *Water Resour. Res.*, 44, W12418, doi:10.1029/2008WR006908.

Rosero, E., Z.-L. Yang, T. Wagener, L. E. Gulden, S. Yatheendradas, and G.-Y. Niu (2010), Quantifying parameter sensitivity, interaction, and transferability in hydrologically enhanced versions of the Noah land surface model over transition zones during the warm season, *J. Geophys. Res.*, *115*, D03106, doi:10.1029/2009JD012035.

Sahoo, A. K., P. A. Dirmeyer, P. R. Houser, and M. Kafatos (2008), A study of land surface processes using land surface models over the Little River Experimental Watershed, Georgia, *J. Geophys. Res., 113*, D20121, doi:10.1029/2007JD009671.

Schaap, M. G., F. J. Leij, and M. T. van Genuchten (1999), A bootstrap-neural network approach to predict soil hydraulic parameters, in Proc. Int. Workshop, *Characterization and Measurements of the Hydraulic Properties of Unsaturated*

Porous Media, edited by M. T. van Genuchten, F. J. Leij, and L. Wu, pp 1237-1250, Univ. of Calif., Riverside, CA.

Shen, C., J. Niu, and M. S. Phanikumar (2013), Evaluating controls on coupled hydrologic and vegetation dynamics in a humid continental climate watershed using a subsurface – land surface processes model, *Water Resour. Res., 49,* 2552-2572, doi:10.1002/wrcr.20189.

Shin, Y., and B. P. Mohanty (2013a), Development of a deterministic downscaling algorithm for remote sensing soil moisture footprint using soil and vegetation classifications, *Water Resour. Res.,* doi: 10.1002/wrcr.20495.

Shin, Y., B. P. Mohanty, and A. V. M. Ines (2013b), Estimating effective soil hydraulic properties using spatially distributed soil moisture and evapotranspiration products at multiple scales, *Vadose Zone J., 12,* doi:10.2136/vzj2012.0074.

Shin Y., B. P. Mohanty, and A. V. M. Ines (2012), Soil hydraulic properties in one-dimensional layered soil profile using layer-specific soil moisture assimilation scheme, *Water Resour. Res., 48*, W06529, doi:10.1029/2010WR009581.

Sivapalan, M. (2005), Pattern, process and function: elements of a unified theory of hydrology at the catchment scale. In: Anderson, M. G. (ed.) Encyclopedia of hydrological sciences. Chichester, UK: John Wiley & Sons, pp. 193–220.

Sloughter, J. M., A. E. Raftery, and T. Gneiting (2006), Probabilistic quantitative precipitation forecasting using Bayesian model averaging, *Tech. Rep. 496,* Dep. of Stat., Univ. of Wash., Seattle.

Smith, M. W., L. J. Bracken, and N. J. Cox (2010), Toward a dynamic representation of hydrological connectivity at the hillslope scale in semiarid areas, *Water Resour. Res., 46*, W12540, doi:10.1029/2009WR008496.

Smith, T., L. Marshall, B. McGlynn, and K. Jencso (2013), Using field data to inform and evaluate a new model of catchment hydrologic connectivity, *Water Resour. Res., 49*, doi:10.1002/wrcr.20546.

Sridhar, V., R. L. Elliott, F. Chen, and J. A. Brotzge (2002), Validation of the NOAH-OSU land surface model using surface flux measurements in Oklahoma, *J. Geophys. Res.*, 107(D20), 4418, doi:10.1029/2001JD001306.

Todini, E., (2008), A model conditional processor to assess predictive uncertainty in flood forecasting, *Intl. J. River Basin Management*, 6(2): 123–137.

Tsai, F. T.-C. and X. Li (2008), Groundwater inverse modeling for hydraulic conductivity estimation using Bayesian model averaging and variance window, *Water Resour. Res.*, 44, W09434, doi:10.1029/2007WR006576.

Van Dam, J. C., J. Huygen, J. G. Wesseling, R. A. Feddes, P. Kabat, P. E. V. Van Waslum, P. Groenendjik, and C. A. van Diepen (1997), Theory of SWAP version 2.0: Simulation of water flow and plant growth in the soilwater-atmosphere-plant environment, *Tech. Doc. 45*, DLO Winand Staring Cent., Wageningen Agric. Univ., Wageningen, Netherlands.

Van Genuchten, M. T. (1980), A closed-form equation foe predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J., 44*:892-898.

Vrugt, J. A., C. J. F. ter Braak, C. G. H. Diks, B. A. Robinson, J. M. Hyman, and D. Higdon (2009), Accelerating markov chain monte carlo simulation by differential evolution with self-adaptive randomized subspace sampling, *Int. J. Nonlinear Sci. Numer. Simul., 10*(3): 271-288.

Vrugt, J. A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M., and Robinson, B. A. (2008), Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, *Water Resour. Res., 44*, W00B09, doi:10.1029/2007WR006720.

Vrugt, J. A. and B. Robinson (2007), Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging, *Water Resour. Res.,* 43, W01411, doi:10.1029/2005 WR004838.

Wang, X.-S., X.-W. Jiang, L. Wan, S. Ge, and H. Li (2011), A new analytical solution of topography-driven flow in a drainage basin with depth-dependent anisotropy of permeability, *Water Resour. Res., 47,* W09603, doi:10.1029/2011WR010507.

Western, A. W., G. Bloschl, and R. B. Grayson (1998), How well do indicator variograms capture the spatial connectivity of soil moisture? *Hydrol. Process. 12*, 1851–1868.

Western, A. W., G. Bloschl, and R. B. Grayson (2001), Toward capturing hydrologically significant connectivity in spatial patterns, *Water Resour. Res., 37*, 83–97.

Western, A. W., S. Zhou, R. B. Grayson, T. A. McMahon, G. Bloschl, and D. J. Wilson (2004), Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes, *J. Hydrology 286*, 113-134.

Wöhling, T. and J. A. Vrugt (2008), Combining multiobjective optimization and Bayesian model averaging to calibrate forecast ensembles of soil hydraulic models, *Water Resour. Res.*, *44*, W12432, doi:10.1029/2008WR007154.

Wood, E. F., et al. (2011), Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, *Water Resour. Res.*, *47*, W05301, doi:10.1029/2010WR010090.

Wösten, J. H. M., G. H. Veerman, and J. Stolte (1994), Water retention and hydraulic conductivity functions of top- and subsoils in The Netherlands: The Staring-series, *Tech. Doc. 18*, Winand Staring Centre, Wageningen, The Netherlands. 66.

Wu, H., X. Zhang, S. Liang, H. Yang, and G. Zhou (2012), Estimation of clear-sky land surface longwave radiation from MODIS data products by merging multiple models, *J. Geophys. Res.,* 117, D22107, doi:10.1029/2012JD017567.

Yang, D. and Musiake K. (2003), A continental scale hydrological model using the distributed approach and its application to Asia, *Hydrol. Process. 17,* 2855-2869.

Ying, M., S. Feng, Z. Huo, and X. Song (2011), Application of the SWAP model to simulate the field water cycle under deficit irrigation in Beijing, China, *Math. Comput. Model., 54*, 1044-1052.

Zaslavsky, D., and G. Sinai (1981), Surface hydrology: I. Explanation of phenomena, *J. Hydraul. Div. Am. Soc. Civ. Eng., 107,* 1-16.

Zeng, X. B. and M. Decker (2009), Improving the Numerical Solution of Soil Moisture-Based Richards Equation for Land Models with a Deep or Shallow Water Table, *J. Hydromet., 10*, 308-319

Zhang, L., W. R. Dawes, T. J. Hatton, P. H. Reece, G. T. H. Beale, and I. Packer (1999), Estimation of soil moisture and groundwater recharge using the TOPOG_IRM model, *Water Resour. Res., 35,* 149-161.

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

Zhang, X., R. Srinivasan, K. Zhao, and M. V. Liew (2008), Evaluation of global optimization algorithms for parameter calibration of a computationally intensive hydrologic model, *Hydrolo. Process.*, doi: 10.1002/hyp.7152.

Zhu, J. and B. P. Mohanty (2006), Effective scaling factor for transient infiltration in heterogeneous soils, *J. Hydrol., 319,* 96-108.

Zhu, Q. and H. S. Lin (2009), Simulation and validation of concentrated subsurface lateral flow paths in an agricultural landscape, *Hydrol. Earth Syst. Sci.*, 13, 1503-1518.