A UNIFYING PLATFORM FOR WATER RESOURCES MANAGEMENT USING PHYSICALLY-BASED MODEL AND REMOTE SENSING DATA

A Dissertation

by

YONGCHUL SHIN

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

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Binayak P. Mohanty
Committee Members,	Yalchin Efendiev
	Patricia K. Smith
	Steven M. Quiring
Head of Department,	Stephen W. Searcy

December 2012

Major Subject: Biological and Agricultural Engineering

ABSTRACT

In recent years, physically-based hydrological models provided a robust approach to better understand the cause-effect relationships of effective hydraulic properties in soil hydrology. These have increased the flexibility of studying the behavior of a soil system under various environmental conditions. One disadvantage of physical models is their inability to model the vertical and horizontal heterogeneity of hydraulic properties in a soil system at the regional scale. In order to overcome this limitation, inverse modeling may be used. Near surface soil moisture, which has been collected routinely by remote sensing (RS) platforms, and evapotranspiration, that is also a pivotal key for water balance near the land surface can be used as alternatives for quantifying the effective soil hydraulic parameters through inverse modeling. However, the new approach suffers from not only the scale discrepancy between RS pixel resolution and model grid resolution, but also its application in complex terrains. Furthermore, hydrological models require a number of required input parameters. Hence, this dissertation focuses on developing a methodology for addressing these problems. The field-scale Soil-Water-Atmosphere-Plant model (SWAP) was extended to regional application, and then coupled with a Genetic Algorithm (GA), to operate as the core of the developed decision support system at the regional level. Also, various stochastic processes were developed and applied to the GA for improving the searching ability of optimization algorithms. The computational simulation-optimization approach was tested and evaluated under various synthetic and field validation experiments demonstrating that the methodology provided satisfactory results. In this dissertation, the

proposed methodologies analyzed the spatio-temporal root zone soil moisture with RS and *in-situ* soil moisture data at the multiple scales. Also, these approaches could provide better input parameters for hydro-climatic models, resulting in better understanding of the hydrologic cycle. Thus, a better understanding of water cycle would help us to be better prepared for efficient water resources management, agriculture, and devastating natural disasters in the real world.

DEDICATION

This dissertation is dedicated to the memory of my parent,

Hanbum Shin and Okja Yoo.

ACKNOWLEDGEMENTS

Ever since I joined the department of Biological & Agricultural Engineering at Texas A&M University as a graduate student, I looked forward to my graduation as time went by. Now, when I reflect upon the past, I see some very hard times. However, this experience has made me more mature mentally and academically.

I would like to thank my committee members, friends, and family who supported me throughout. I would like to convey my heartfelt gratitude to Dr. Binayak P. Mohanty, my doctoral committee chair, for introducing me to soil hydrology. His untiring guidance and encouragement lead me to this stage of my academic career. I greatly appreciate the trust he placed in me and the encouragement that he gave me while I was conducting my research. I would also like to thank my committee members Dr. Yalchin Efendiev, Dr. Patricia Smith, and Dr. Steven Quiring for their insightful guidance. I also express my gratitude to all teachers who encouraged and helped me during my undergraduate and graduate degrees. I wish to thank my father, mother, sisters, and brothers in law who helped and supported me with devoted and endless love. They always trusted and encouraged me to rebuild and move forward. I can never repay the special love they gave me. My dear friends, namely: Amor Ines, Nandita Gaur, Raghavendra Jana, Jonggun Kim, Champa Joshi, Narendra Das, Zhenlei Yang, Dipankar Dwivedi, Bhavna Arora, and Sayena Marandi are highly appreciated for the help that they provided. I also would like to acknowledge the financial support of NASA and NSF for funding my research.

Above all, I am grateful to God for his blessings and help.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

GENERAL INTRODUCTION

1.1 Problem Statement

Hydrological processes take place at field- to large-scales in the real world. Better understanding of the water cycle at different scales would be a key for sustainable water resources, agricultural production, and ecosystems health in the 21st century. Efficient agricultural/water resources management are required for sustainability of the growing global population. This warrants better predictive tools for root zone soil moisture at various scales. It will provide useful strategies for not only efficient water use but also for reducing potential risk due to agricultural drought and flood.

The soil hydraulic properties at the land surface and in the unsaturated zone are key variables for many hydro-climatic processes, because these parameters are essential inputs to large-scale hydrologic and hydro-climatic models. Soil hydraulic properties vary in the horizontal and vertical directions. Their variability is dependent on various factors like tillage practice, pore-size distribution due to structural cracks and root development and decay, textural layering, land surface conditions, etc. Thus, the understanding of how the land surface and subsurface information might affect vertical and horizontal moisture exchange is a challenging issue. In order to capture the vertical and horizontal variability of hydraulic parameters, field-scale measurements are required. However, obtaining *insitu* measurements of required soil hydraulic parameters at various spatial scales is practically impossible. Therefore, alternatives for quantifying the soil hydraulic parameters at the scales of interest are necessary.

1.2 Motivation

The influence of variable landscape heterogeneities (i.e., soil textures, vegetation covers, topography, etc.), atmospheric forcings (precipitation), complexity of the layered soil profile, initial and bottom boundary conditions, evapotranpiration, etc. on effective soil hydraulic properties is required to better understand soil hydrology at different scales. It will lead to the development of useful strategies for efficient water resources management and reducing potential risk of agricultural drought in the real world. Furthermore, new methodologies with respect to data assimilation, down-/up-scaling algorithms, non physically-based stochastic models, etc. need to be developed for overcoming limitations of current approaches in quantifying the soil hydraulic parameters across temporal and spatial domains.

This research focused on scaling (downscaling and upscaling) algorithms, data assimilation, inverse model development, and incorporation of stochastic evolutionary algorithms for addressing the connections between the environmental factors and soil hydraulic parameters at various scales across the land surface.

1.3 Research Objectives

The primary objectives of this research are to better understand hydro-climatic processes at the local and large scales and develop/improve efficient water resources management for agriculture, ecosystem dynamics and alteration, and hydrology.

The specific research objectives of this project are:

- To analyze the impacts of soil heterogeneity, vegetation covers, initial and bottom boundary conditions on prediction of near-/sub-surface water flow in a layered soil profile.
- To develop and test an improved inverse modeling algorithm for pixel-based effective soil hydraulic properties estimation at multiple scales using remotely sensed soil moisture and evapotranspiration data products.
- To develop a deterministic downscaling algorithm for pixel-based soil moisture and evapotranspiration to produce sub-pixel level soil moisture products from remote sensing pixel data at various resolutions.
- To develop and implement a scaling (joint downscaling and upscaling) algorithm for remotely sensed and *in-situ* soil moisture data.
- To develop a non-parametric evolutionary algorithm (using hidden Markov model genetic algorithms) for predicting long-term root zone soil moisture dynamics with multivariate time series of precipitation.
- To develop a drought severity assessment framework using remotely sensed soil moisture products with climate change scenarios under various regional hydro-climatic conditions.

In Chapter II, a layer-specific soil moisture assimilation scheme using a simulation-optimization framework, Soil-Water-Atmosphere-Plant model with genetic algorithm (SWAP-GA), was developed to analyze the impact of soil layering associated with various soil textural combinations in the profile. The new data assimilation scheme

quantifies soil hydraulic properties of different soil layers in the root zone (0-200 cm). This approach was tested under various synthetic and field validation experiments (i.e., initial and bottom boundary conditions, soil textures, vegetation covers, presence of ground water table depths, etc.) at the Lubbok site in Texas and Little Washita (LW) watershed in Oklahoma.

In Chapter III, a new inverse modeling algorithm by combining soil moisture based Noisy Monte Carlo Genetic Algorithm (NMCGA) and Surface Energy Balance Algorithm for Land (SEBAL) based ET products was developed. The coupled NMCGA and SEBAL approach uses remotely sensed (RS) soil moisture and evapotranspiration products for quantifying pixel-scale soil hydraulic properties. This approach was evaluated at multiple scales under different hydro-climates (Lubbock, Texas; Little Washita watershed, Oklahoma; and Walnut Creek watershed, Iowa).

In Chapter IV, a new deterministic downscaling algorithm (DDA) was developed for estimating finer-scale soil moisture with pixel-based remotely sensed (RS) soil moisture (SM) and evapotranspiration (ET) products. This approach was evaluated under various synthetic and field validation experiment (Little Washita 13 and 21, Oklahoma) conditions including homogeneous and heterogeneous land surface conditions composed of different soil textures and vegetation. Our algorithm is based on determining soil hydraulic properties for different sub-pixels and their specific locations within a RS pixel and estimating root zone soil moisture dynamics of each individual sub-pixel.

In Chapter V, a new scaling (joint downscaling and upscaling) algorithm was developed for improving the availability of remotely sensed (RS) and *in-situ* soil

moisture in hydrologic applications. This scheme can downscale RS soil moisture footprints as well as upscale *in-situ* measurements simultaneously across a topographically complex regional area. This algorithm adapted a near-surface soil moisture assimilation approach based on inverse modeling using a simulation-optimization framework, Soil-Water-Atmosphere-plant model with genetic algorithm (SWAP-GA). Also, various hydrological models such as Noah Land Surface Model (Noah LSM) and Community Land Model (CLM) were used for testing our proposed approach. The new algorithm used normalized digital elevation model (DEM_{norm} representing topography) and normalized difference vegetation index ($NDVI_{norm}$ representing vegetation covers) components at finer-scale (30 m$\times$30 m) for capturing the heterogeneity of land surface in scaling down and up soil moisture data. We selected the Little Washita watershed for testing this scheme at multiple scales.

In Chapter VI, a new genetic algorithm-based hidden Markov model (HMMGA) was developed for exploring root zone soil moisture dynamics at different soil depths using the multivariate time series of precipitation under two different hydro-climatic regions. It uses optimal state sequences derived by the HMMGA and statistics of historical soil moisture measurements for predicting root zone soil moisture in the future. Also, this approach reproduced rainfall occurrence probabilities and wet/dry spell lengths. To test the new approach, we selected two regions including Oklahoma (130 km $\times$ 130 km) and Illinois (300 km $\times$ 500 km) soil moisture fields during 1995 to 2009 and 1994 to 2010, respectively.

In Chapter VII, a drought severity assessment framework was explored using a modified grid-based disaggregation algorithm adapting a combined simulation-optimization algorithm, soil-water-atmosphere-plant (SWAP) model with a genetic algorithm (GA). Here, we developed a new evolutionary algorithm (called Ensemble Multiple Operators Genetic Algorithm, EMOGA) for improving limitations of GAs. A sub-grid analysis of root zone soil moisture using the SWAP-EMOGA linkage was performed for downscaling remotely sensed (RS) soil moisture products at the soil-vegetation level. Based on the historical data, pixel-scale hydraulic parameters at finer-scales were estimated from RS soil moisture using the SWAP-EMOGA. These estimated hydraulic parameters along with meteorological variables obtained from general circulation models (GCMs) were used to predict soil moisture using the SWAP model. Further, drought severity was calculated using a Soil Moisture Deficit Index (SMDI) based on the projected soil moisture obtained from the SWAP model. Then, we assessed potential risks (i.e., agricultural drought severity, water resources management, etc.) with forecasted root zone soil moisture dynamics at sub-grid scales for the next decade (2010-2020).

CHAPTER II

SOIL HYDRAULIC PROPERTIES IN ONE-DIMENSIONAL LAYERED SOIL PROFILE USING LAYER-SPECIFIC SOIL MOISTURE ASSIMILATION SCHEME*

2.1 Synopsis

We developed a layer-specific soil-moisture assimilation scheme using a simulation-optimization framework, Soil-Water-Atmosphere-Plant model with genetic algorithm (SWAP-GA). Here, we explored the quantification of the soil hydraulic properties in a layered soil column under various combinations of soil types, vegetation covers, bottom boundary conditions and soil layering using idealized (synthetic) numerical studies and actual field experiments. We demonstrated that soil layers and vertical heterogeneity (layering arrangements) could impact to the uncertainty of quantifying soil hydraulic parameters. We also found that, under layered soil system, when the sub-surface flows are dominated by upward fluxes, e.g., from a shallow water table, the solution to the inverse problem appears to be more elusive. However, when the soil profile is predominantly draining, the soil hydraulic parameters could be fairly estimated well across soil layers, corroborating the results of past studies on homogenous soil columns.

---

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

In the field experiments, the layer-specific assimilation scheme successfully matched soil moisture estimates with observations at the individual soil layers suggesting that this approach could be applied in real world conditions.

2.2 Introduction

Soil hydraulic parameters are significant components for many hydrological, meteorological, and general circulation models [*Hansen et al.,* 1999; *Mohanty et al.,* 2002; *Mohanty and Zhu,* 2007]. They are used to define the soil hydraulic properties in the vadose zone, characterizing the effective hydraulic behavior of the soil system [*Wood,* 1994; *Vrugt et al.,* 2004].

With the objective of exploring the utility of remote sensing of soil moisture for deriving soil hydraulic properties at aggregate scale, *Ines and Mohanty* [2008a,b, 2009] tested the hypothesis that near-surface soil moisture assimilation scheme can be used to quantify effective soil hydraulic properties of an "effective" soil column based on the inverse modeling. The effective soil column is a 'homogenous' conceptual representation of a real-world soil column (composed of soil horizons) characterized by effective soil hydraulic properties. The inverse method using near-surface soil moisture assumes that any perturbations made at the near-surface soil layer could influence the soil moisture dynamics at the sub-surface and hence can inform the estimations of sub-surface soil hydraulic properties. The effective soil hydraulic properties serve as 'average' properties of the system. *Ines and Mohanty* [2008b] however found that if the system is highly heterogeneous, the assumption of effective soil column could fail.

Understanding of how soil vertical layering might affect soil moisture exchange and soil hydraulic parameter estimations is therefore important. Significant efforts have been made to account for the impact of soil heterogeneity on field soil moisture contents. The soil hydraulic conductivity, moisture content, and soil hydraulic parameters are variable at the field scale [*Nielsen et al*., 1973; *Stockton and Warrick*, 1971; *Jana and Mohanty, 2012a,b*]. *Bosch* [1991] studied an analytical expression for forecasting (potential) errors by using point observation of the matric potential (h) to determine the average matric potential (h) in a heterogeneous column. The instantaneous profile method suggested by *Green et al.,* [1986] can be used to measure hydraulic conductivities (K(h)) at field-scales [*Rose et al*., 1965; *van Bavel* et al., 1968; *Nielsen et al*., 1973]. This method involves measurement of moisture content (θ) and matric potential (h) throughout the profile.

Zhu and Mohanty [2002] reported various hydraulic parameter averaging schemes and the mean hydraulic conductivity for predicting the mean fluxes in the horizontal heterogeneous blocks under steady-state of infiltration and evaporation using Gardner-Russo exponential model [*Gardner* 1958] and the Brooks-Corey model [*Brooks and Corey* 1964]. The effective hydraulic parameter estimations were related to areal soil heterogeneity and land surface conditions such as root distribution and surface ponding depth [*Zhu and Mohanty* 2003, 2004, 2006; *Zhu et al*. 2004, 2006]. The soil hydraulic properties were also influenced by vertical heterogeneity (e.g., tillage practice, pore-size distribution due to structural cracks and root development and decay, textural layering and geology), and parameter estimations could vary in the vertical direction

(*Mohanty et al*., 1994, *Mallants et al., 1996*). Although *Mohanty and Zhu* [2007] investigated effective soil hydraulic parameter averaging schemes for steady-state flow in heterogeneous shallow subsurface useful to land-atmosphere interaction modeling, not many studies have been carried out to explore issues for vertical subsurface heterogeneity associated with various soil types.

In this study, we adopted a layer-specific soil moisture assimilation scheme for determining the soil hydraulic parameters in layered soil profiles. The main objective is to analyze the impact of soil layering associated with various soil textural combinations in the profile and to quantify the one-dimensional soil hydraulic properties of different soil layers in the root zone (0-200 cm) based on the layer-specific soil moisture assimilation scheme. This work could be useful to characterize hydrologic systems that are instrumented to measure root zone soil moisture. Additionally, this approach may serve as a basis for developing futuristic analytical platforms to characterize vadose zone systems at regional and global scales by synthesizing profile soil moisture data collected using various ground-, air-, and space-based sensors of different spectral frequencies and penetrating depths.

2.3 Materials and Methods

2.3.1 1D Layer-Specific Soil Moisture Assimilation

2.3.1.1 Conceptual Framework

The aim of the layer-specific soil moisture assimilation scheme is to estimate the soil moisture retention $\theta(h)$ and hydraulic conductivity $K(h)$ curves in a layered soil column (e.g., 1^{st}: 0-10 cm, 2^{nd}: 10-60 cm, and 3^{rd}: 60-200 cm) by optimizing the

effective soil hydraulic parameters for each layer based on a simulation-optimization [*Ines and Droogers*, 2002]. As depicted in Fig. 2.1, the approach uses the soil moisture in the layers (here, we set up at 5 cm depth for the 1st and 2nd layers and at 10 cm depth for the 3rd layer) to estimate the layer specific soil hydraulic properties. The choice on locations of soil moisture measurements in the soil profile to be used in the simulation-optimization can be established in a more systematic way. Mathematically, the soil hydraulic parameters are obtained by finding a set of soil hydraulic parameter **p** such that the differences between observed $q_i(t)$ and simulated $\theta_i(t; p_i)$ soil moisture at soil layers i, are minimized; where $\mathbf{p} = p_{i=1,\ldots,M}$; and p_i is the corresponding soil hydraulic parameters in the individual soil layers. The choice of objective function is critical in inverse modeling; from sensitivity analysis (see Section 3.3) we selected the additive absolute form (Eq. 2.1) as it produced better results than other forms considered in this study (e.g., multiplicative and additive squared delta); $Z(\mathbf{p})$ is the objective function, M is the number of soil layers, N is the time domain, and t is the index for time.

$$Z(\mathbf{p}) = \min\left\{\frac{1}{N}\frac{1}{M}\sum_{t=1}^{N}\sum_{i=1}^{M}|\theta_i(t) - \theta_i(t;p_i)|\right\} \qquad (2.1)$$

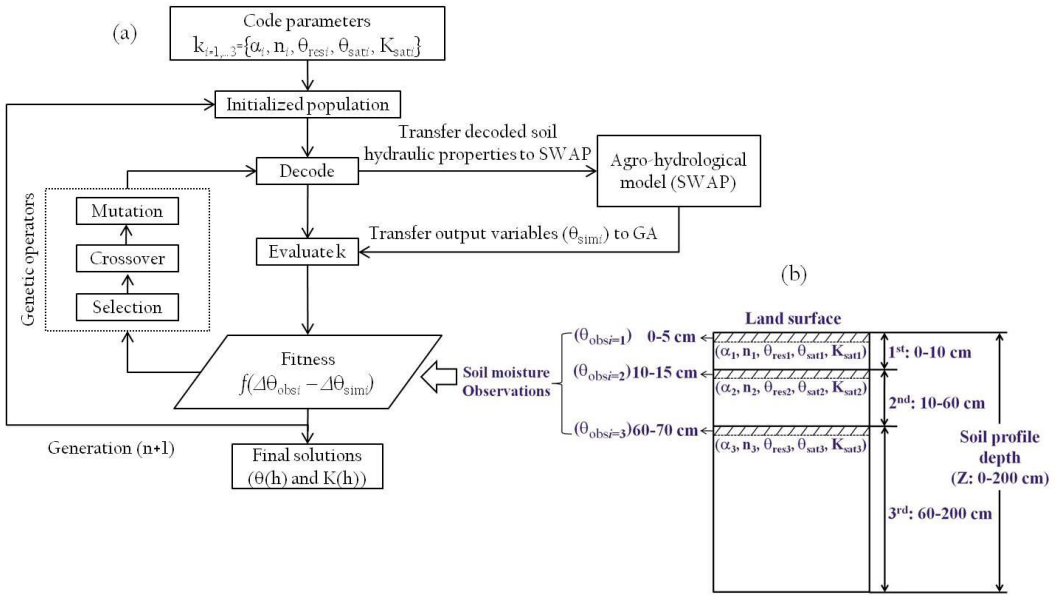

Figure 2.1: (a) Schematic diagram of layer-specific soil moisture assimilation scheme (SWAP-GA linkage) based on inverse modeling, (b) layered soil column

2.3.1.2 Description of the SWAP Model

SWAP is a physically-based model that simulates the processes of the soil-water-atmosphere-plant system [*van Dam et al.,* 1997]. The soil moisture dynamics in the soil column can be described using the one-dimensional Richards' equation (Eq. 2.2). SWAP model solves Eq. (2.3) numerically using the implicit finite difference scheme of *Belmans et al*. (1983),

$$\frac{\partial\theta(z,t)}{\partial t} = C(h(z,t))\frac{\partial h(z,t)}{\partial t} = \frac{\partial[K(h(z,t))(\frac{\partial h(z,t)}{\partial z}+1)]}{\partial z} - S(h,z,t) \qquad (2.2)$$

where K is the hydraulic conductivity (cm d^{-1}), h is the soil water pressure head (-cm), z is the vertical soil depth (cm) taken positively upward, C is the differential water capacity (cm^{-1}), and S(h,z,t) is the actual soil moisture extraction rate by plant roots (cm^3 cm^{-3} d^{-1}) defined as Eq. (2.3).

$$S(h,z,t) = \begin{cases} \alpha_w(h,z,t)\dfrac{T_{pot}(t)}{Z_r}; z < Z_r \\ 0; z > Z_r \end{cases} \qquad (2.3)$$

where T_{pot} is the potential plant transpiration (cm d^{-1}), Z_r is the rooting depth (cm), and α_w is a reduction factor as function of h (at depth z and time t) and accounts for water deficit and oxygen stress [*Feddes et al.,* 1978].

The soil hydraulic functions are described by analytical expressions of *van Genuchten* [1980] and *Mualem* [1976].

$$S_e(h,z,t) = \frac{\theta(h(z,t)) - \theta_{res}}{\theta_{sat} - \theta_{res}} = [\frac{1}{1+|\alpha h(z,t)|^n}]^m \qquad (2.4)$$

$$K(h,z,t) = K_{sat}S_e(h,z,t)^{\lambda}[1-(1-S_e(h,z,t)^{1/m})^m]^2 \qquad (2.5)$$

where S_e is the relative saturation (-), θ_{res} and θ_{sat} are the residual and saturated water contents ($cm^3\ cm^{-3}$), α (cm^{-1}), n (-), m (-), and λ (-) are shape parameters of the retention and the conductivity functions, K_{sat} is the saturated hydraulic conductivity (cm d^{-1}), and m=1-1/n.

The SWAP model considers for several combinations of the top (atmospheric) and bottom boundary conditions [*van Dam et al.,* 1997]. Moreover, it contains water management modules for irrigation and drainage modules as well as process-based crop growth models including WOFOST for simulating the impacts of weather, soil type, plant type, and water management on the crop growth [*van Dam et al.,* 1997, *van Dam,* 2000]. The SWAP model calculates the potential evapotranspiration (ET_{pot}) according to the Penman-Monteith equation using daily meteorological data. The partition of ET_{pot} rate into potential transpiration rate (T_{pot}) and potential evaporation rate (E_{pot}) is determined by the leaf area index or the soil cover fraction. The potential evapotranspiration (ET_{pot}) is calculated using the minimum value of canopy resistance and actual resistance. Then the actual evapotranspiration (ET_{act}) is calculated by the root water uptake reduction due to water and salinity stress.

2.3.1.3 Genetic Algorithm

Genetic algorithms (GAs) are powerful search algorithms based on the precept of natural selection [*Holland,* 1975; *Goldberg,* 1989]. The unknown parameters in a search problem are represented by genes, which are arranged in an array called chromosome. The search starts by initializing a population of chromosomes becoming the starting points in the search across the search surface. The suitability of a chromosome is

evaluated using a fitness function. Based on their fitness, they are selected to the mating pool, reproduce through the process of crossover, and allowed to mutate. The solution of the search problem would be the fittest chromosome that survives after many generations. In this study, a modified-*micro*GA was used to search for the parameter set (**p**) by minimizing the error between the simulated and observed soil moisture in the layered soil column. The modified-*micro*GA is a GA variant that uses a micro population to search for the solution of the inverse problem. The uniqueness in the modified-*micro*GA is the ability to restart when the chromosomes of the micro-population are nearly 90% similar in structure, allowing more micro-population restarts [*Ines and Droogers*, 2002; *Carroll*, 1996; *Goldberg*, 2002; *Krishnakumar*, 1989; D. L. Carroll, Fortran genetic algorithm (GA) driver, available at www.cuaerospace.com/carroll/ga.html]. The modified-*micro*GA allows a creep mutation (at base 10). *Ines and Mohanty* (2008a) added an intermittent jump mutation to further introduce new genetic materials during the search. A time saving mechanism was designed by remembering not only the elite chromosome of the previous generation (g-1) but also its remaining chromosomes such that when they are generated in the next generation, there is no need to run them anymore in the SWAP model, saving computational time [*Ines and Honda*, 2005]. The elite chromosome is always reproduced in the next generation. The modified-*micro*GA was applied to the inverse modeling (IM)-based layer-specific soil moisture assimilation scheme [*Ines and Droogers,* 2002, *Ines and Mohanty*, 2008a,b]. The search spaces for each Mualem-Van Genuchten parameters in the multi-layered soil system as used in this study are shown in Table 2.1.

Table 2.1: Summary of the parameter constraints in the genetic algorithm

Case studies	Parameter	Search space		No. of bit (L)	Binary(2^L)
		Min. values	Max. values		
Case 1 to 2	α	0.006	0.033	5	2^5 (32)
	n	1.200	1.610	6	2^6 (64)
	θ_{res}	0.061	0.163	7	2^7 (128)
	θ_{sat}	0.370	0.550	5	2^5 (32)
	K_{sat}	1.840	55.700	10	2^{10} (1024)
Case 3 (LW 02 and 11)	α	0.006	0.033	5	2^5 (32)
	n	1.200	2.200	6	2^6 (64)
	θ_{res}	0.040	0.163	7	2^7 (128)
	θ_{sat}	0.340	0.550	5	2^5 (32)
	K_{sat}	1.840	250.000	10	2^{10} (1024)
Case 3 (LW 07)	α	0.006	0.033	5	2^5 (32)
	n	1.200	2.200	6	2^6 (64)
	θ_{res}	0.040	0.163	7	2^7 (128)
	θ_{sat}	0.340	0.550	5	2^5 (32)
	K_{sat}	1.840	130.000	10	2^{10} (1024)

Note: Total search space = 32*64*128*32*1024 = 8,589,934,592
Example of **p** = {α, n, θ_{res}, θ_{sat}, K_{sat} } = {00101, 110010, 0001111, 00001, 0101000101}
Prob. of crossover = 0.5
Prob. of creep mutation = 0.5
Prob. of intermittent jump mutation = 0.05
Population = 10 chromosomes
No. of multi-population = 3
Max. generation = 500

2.3.1.4 Parameter Uncertainty

When an elitist modified-*micro*GA converges to the solution, all of the chromosomes in a population are almost similar. To create some sort of uncertainty bounds to the solution, a multi-population generated by various random number generator seeds (e.g., -1000, -950, and -750) were run concurrently. After many generations, the average fitness of all the chromosomes from the multi-populations is calculated and classified as above or below average. The above average solutions are considered as the most probable solutions. The 95 percent confidence interval (95PCI) of the most probable solutions was calculated as,

$$\text{Range}_{p,s,t,i} = 95\text{PCI}_{p,s,t,i+} - 95\text{PCI}_{p,s,t,i-} \quad (2.6)$$

where $95\text{PCI}_{p,s,t,i+}$ and $95\text{PCI}_{p,s,t,i-}$ are the upper and lower boundary of the 95PCI, p is the soil hydraulic parameter, s is the index of soil type, t is the time (running) index, and i is the soil layers.

Pearson's correlation (R^2) and uncertainty analysis (Mean Absolute Error-MAE, Mean Bias Error-MBE, and Root Mean Square Error-RMSE) between observed and simulated data are also used to assess the performance of the modified-*micro*GA,

$$R_i^2 = \frac{\sum_{t=1}^{n}(\hat{\theta}_{sim,i,t} - \bar{\theta}_{sim,i})(\theta_{obs,i,t} - \bar{\theta}_{obs,i})}{\sqrt{\sum_{t=1}^{n}(\hat{\theta}_{sim,i,t} - \bar{\theta}_{sim,i})^2 \sum_{t=1}^{n}(\theta_{obs,i,t} - \bar{\theta}_{obs,i})^2}} \quad (2.7)$$

$$\left.\begin{aligned} \mathrm{MAE}_i &= \frac{1}{n}\sum_{t=1}^{n}|(\theta_{obs,i,t}-\hat{\theta}_{sim,i,t})| \\ \mathrm{MBE}_i &= \frac{1}{n}\sum_{t=1}^{n}(\theta_{obs,i,t}-\hat{\theta}_{sim,i,t}) \\ \mathrm{RMSE}_i &= \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\theta_{obs,i,t}-\hat{\theta}_{sim,i,t})^2} \end{aligned}\right\} \quad (2.8)$$

where, $\hat{\theta}_{sim,i,t}$ is the average soil moisture of different populations with the time index (t), $\bar{\theta}_{sim,i}$ is the average of $\hat{\theta}_{sim,i,t}$, $\theta_{obs,i,t}$ is the observed soil moisture for the time index (t), and $\bar{\boldsymbol{\theta}}_{\boldsymbol{obs,i}}$ is the average of $\theta_{obs,i,t}$, respectively. Note that the MBE and RMSE were tested only for the field experiments.

2.3.2 Numerical Experiments

This study estimates the effective soil hydraulic parameters in a layered soil column adopting the layer-specific soil moisture assimilation scheme based on the inverse modeling approach [*Ines and Mohanty,* 2008a,b, 2009]. The numerical experiments were conducted for three cases: i) Case 1: layered soil column with free drainage, ii) Case 2: layered soil column with varying water table depths (i.e., -200, -150, and -100 cm from the soil surface), and iii) Case 3: field experiments.

The soil profile layering is given as follows: the 1st (top 0-10 cm), 2nd (10-60 cm), and 3rd (60-200 cm) soil layers (Fig. 2.2). The top soil moisture (1st: 0-5 cm, 2nd: 10-15 cm, 3rd: 60-70) below the soil interfaces were only extracted and used for quantifying the soil hydraulic parameters in the soil layers for Cases 1 to 3, respectively.

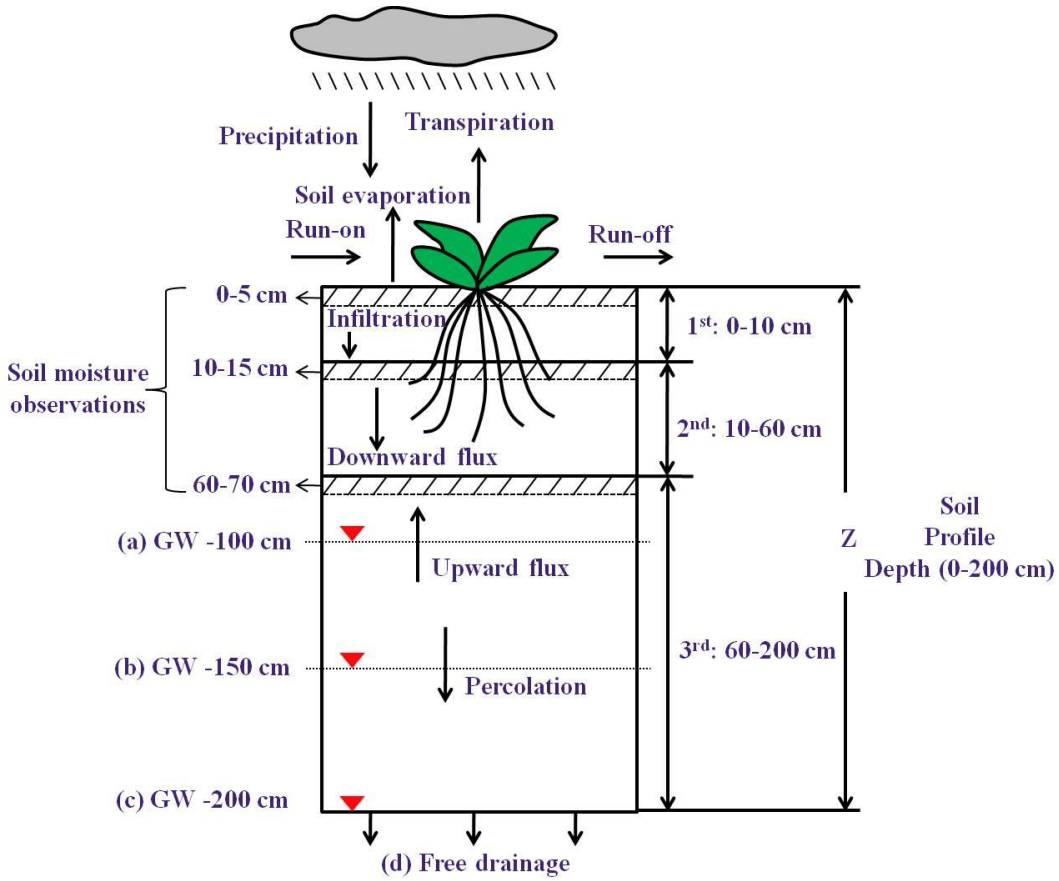

Figure 2.2: Layered soil columns used in the numerical experiments with free drainage and various ground water (GW) table depths; (a) GW -100 cm, (b) GW -150 cm, (c) GW -200 cm, (d) free drainage

In real-world conditions, soil profiles are irregularly layered, thus the decision for selecting the layer depths where soil moisture observations will be compared with simulations should be based upon the available data. For all the simulations, the soil column was discretized into 33 computational layers. The first soil layer was finely discretized at intervals of 1 cm. The second and third soil layers were discretized at intervals of 5 cm and 10 cm (except 33rd layer with 20 cm discretization), respectively. For the free-draining case, the initial soil water pressure head distribution in the soil profile was prescribed uniformly at -150 cm. For the cases with groundwater table bottom boundaries, they are prescribed with initial soil water pressure head distribution in hydrostatic equilibrium with the initial water table depths. Various land covers (bare soil, grass, and wheat) representative of annual crops in the study area (Little Washita Watershed, Oklahoma) were considered for the numerical experiments.

In the hypothetical cases, we used the soil hydraulic parameter values from the UNSODA database as reference soil hydraulic data for the given soils in each soil layers. Using the weather data at the Little Washita (ARS 134) site in 1997, we generated synthetic daily soil moisture datasets using SWAP in a forward mode. These daily soil moisture data were then used to estimate back the soil hydraulic parameters for the layered system. Several field sites were selected to evaluate the applicability of the layer-specific soil moisture assimilation scheme under actual field condition in Case 3 (Fig. 2.3). Some details of the different cases are given below.

2.3.2.1 Case 1: Layered Soil Column under Free Drainage Condition

Some of the uncertainties in the estimation of soil hydraulic parameters in the soil system can be associated with various environmental factors (e.g., root density, rooting depth, soil layers, different combinations of soil types, and profile arrangement, etc.). For this reason, we conducted nine inverse modeling scenarios for Case 1 comprising of various soil types, soil layers, and vegetation combinations. As base case scenarios, the six scenarios were comprised of layer combinations of sandy loam, silt loam, and clay loam along the soil profile with grass cover. These scenarios aimed at assessing the effects of soil layering and heterogeneity in the sub-surface (Table 2.2, where CB1 to 6 denote soil layering combinations).

The other three scenarios included varying the vegetation covers, e.g., bare soil, grass, and wheat to evaluate the impact of varying vegetations in the layer specific data assimilation procedure using only the CB 5 case (Table 2.2). Also analyzed are the interactions between water stress by crops (T_{act}/T_{pot}) and near-surface (0-5 cm) soil moisture changes near the land surface using the CB 5. This study considered only rain-fed conditions for the numerical cases.

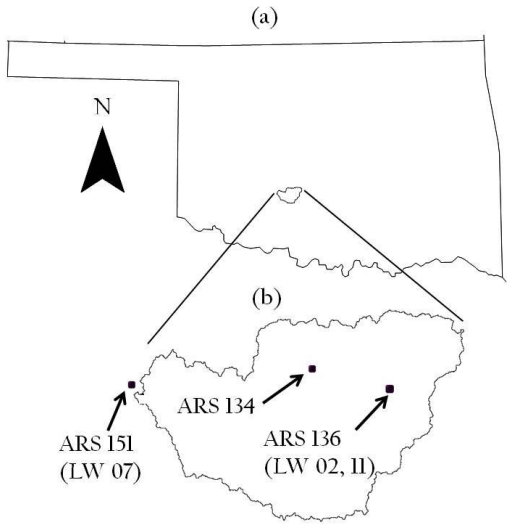

Figure 2.3: The study area; (a) Oklahoma, (b) the Little Washita (LW 02, 07, and 11) watershed

Table 2.2: Combinations of three soil types for Case 1

Depth	Combinations (CB) of three soil types					
	CB 1	**CB 2**	**CB 3**	**CB 4**	**CB 5**	**CB 6**
1st (0-10 cm)	Sandy loam	Sandy loam	Silt loam	Silt loam	Clay loam	Clay loam
2nd (10-60 cm)	Silt loam	Clay loam	Sandy loam	Clay loam	Sandy loam	Silt loam
3rd (60-200 cm)	Clay loam	Silt loam	Clay loam	Sandy loam	Silt loam	Sandy loam

2.3.2.2 Case 2: Layered Soil Column under Various Ground Water Table Conditions

Under cases when soil moisture dynamics in the unsaturated zone is governed significantly by shallow water table, additional experiments with various water table depths (-200, -150, and -100 cm from the soil surface) were conducted. This case aimed to assess the effects of groundwater on the estimates of effective soil hydraulic

properties (only for CB 5 in Case 1 - grass cover) using the layer-specific soil moisture assimilation approach.

2.3.2.3 Case 3: Field Experiments

The layer-specific soil moisture assimilation scheme was applied to the several field monitoring sites within the Little Washita (LW 02, 07, and 11) watershed in Oklahoma using datasets from the Southern Great Plains Hydrology Experiment 1997 (SGP97) [*Mohanty et al.*, 2002; *Heathman et al.*, 2003, *Das and Mohanty, 2006*].

Daily weather datasets (e.g., precipitation, solar radiation, relative humidity, minimum and maximum temperature, and wind speed) were collected at the USDA Agricultural Research Service (ARS 136 and 151) micronet and the Oklahoma Mesonet weather stations from Jan. 1 – Dec. 31, 1997. The LW 02, 07, and 11 sites are characterized by a mixture of loam, sandy loam, and sandy loam with grass covers, with a rooting depth of (up to) 100-120 cm [*Mohanty et al.*, 2002; Table 2.3]. The bottom boundary condition was unknown at the field sites. Therefore, we tested free-drainage conditions and several ground water table depths (-100, -150, and -200 cm) as bottom boundaries and selected the bottom boundary condition (free-drainage conditions for the LW 02, 07, and 11 were selected), with the best performance (fitness) obtained by the genetic algorithm.

Table 2.3: Field-scale soil texture and soil hydraulic properties in the layered soil column at the LW 02, 07, and 11 sites

Sites	No. of soil layers	Depth increment[a]	Sand[a] (%)	Silt[a] (%)	Clay[a] (%)	Soil texture[a]	α^a	n^a	$\theta_{res}{}^a$	$\theta_{sat}{}^a$	$K_{sat}{}^a$	Vegetation rooting depth(cm)
	1st layer	0-30 cm	40.47	43.15	16.37	Loam (L)	0.012	1.679	0.127	0.397	114.650	
LW 02	2nd layer	30-60 cm	40.47	41.38	18.14	Loam (L)	0.013	1.505	0.091	0.397	203.560	Up to 120
	3rd layer	60-90 cm	35.66	45.59	18.75	Loam (L)	0.027	1.616	0.102	0.482	238.120	
	1st layer	0-20 cm	83.89	8.61	7.50	Loam sand (LS)	0.011	2.112	0.061	0.348	53.533	
LW 07	2nd layer	20-40 cm	65.86	20.39	13.75	Sandy loam (SL)	0.016	1.736	0.048	0.345	65.837	Up to 100
	3rd layer	40-60 cm	61.82	24.43	13.75	Sandy loam (SL)	0.021	1.711	0.091	0.387	120.100	
LW 11	1st layer	0-20 cm	59.13	21.40	19.47	Sandy loam (SL)	0.019	1.460	0.046	0.416	186.190	Up to 100

[a]Field observations [*Mohanty et al.*, 2002]

In this study, the soil core samples extracted from the different soil depths (1st: 0-5 cm, 2nd: 30-35 cm, 3rd: 60-65 cm for LW 02 and 1st: 0-5 cm, 2nd: 20-25 cm, 3rd: 40-45 cm for LW 07) collected during the SGP97 (June 18 to July 18, 1997) were analyzed to obtain the soil hydraulic parameters in the laboratory experiment. Using the laboratory derived soil hydraulic properties, we estimated the soil moisture in the soil layers for the field sites with the hydrological (SWAP) model in a forward modeling mode. These soil moisture data were used to estimate back the heterogeneous soil hydraulic properties across the soil profile to test if the layer specific assimilation method could successfully match the laboratory derived soil hydraulic parameters (a.k.a., forward-backward modeling under actual field condition setting). This part of the study does not serve as a validation of the method because there were no measured soil moisture profile data available at the sites to derive independently the layer soil hydraulic parameters, but aimed to ascertain the utility of the approach under field conditions. We compared the derived soil hydraulic properties by inverse modeling, with UNSODA soil hydraulic data [*Leij et al.*, 1999] and the laboratory-derived data.

The case of LW 11 is more of calibration-validation study. The daily (*in-situ*) soil moisture observations (21 days; DOY: 169-181) measured by the time domain reflectometer (TDR) probe in the soil layers (1st: 0-5 cm, 2nd: 20-25 cm, 3rd: 40-45 cm) were used for calibration, then validation runs were done for DOY: 182-197. Validation here means that we used the derived layered soil hydraulic parameters to simulate soil moisture for the remaining days. The modeling soil column was composed of three layers (1st: 0-20 cm, 2nd: 20-40 cm, 3rd: 40-60 cm for LW 11) determined by the depths

at where the soil moisture were measured. The derived soil hydraulic parameters were compared with laboratory-derived parameters (no data at deeper depths from our SGP97 hydraulic property database near-by).

2.4 Results and Discussion

In this study, various combinations of soil layers, soil types, vegetations, and ground water table depths are used for studying their impacts on estimation of soil hydraulic parameters in a layered soil domain. The following sections present the results of the inverse modeling experiments.

2.4.1 Case 1: Layered Soil Column under Free Drainage Condition

We estimated the θ(h) and K(h) curves in the layered soil profile using the combinations (CB 1 to 6) of three soil types in Fig. 2.4. The estimated θ(h) in CB 1 to 6 corresponded well with the reference curves although the uncertainty bounds showed increasing trends with soil depths. On the other hand, the estimated K(h) in the layered soil profile is more uncertain than θ(h), suggesting that K(h) is more difficult to estimate than θ(h) in a layered system with soil moisture information only being used in the inverse modeling. We observed that soil hydraulic parameter estimation is influenced not only by soil layering, but also the order/sequence of vertical heterogeneity in the soil profile. CB 6 for example, although clay loam and silt loam soils were located in the 1^{st} and 2^{nd} layers, their θ(h) and K(h) uncertainty bounds have broader range (more uncertainty) than those at the 3^{rd} layer, while when they are located in other arrangements, they are better identified.

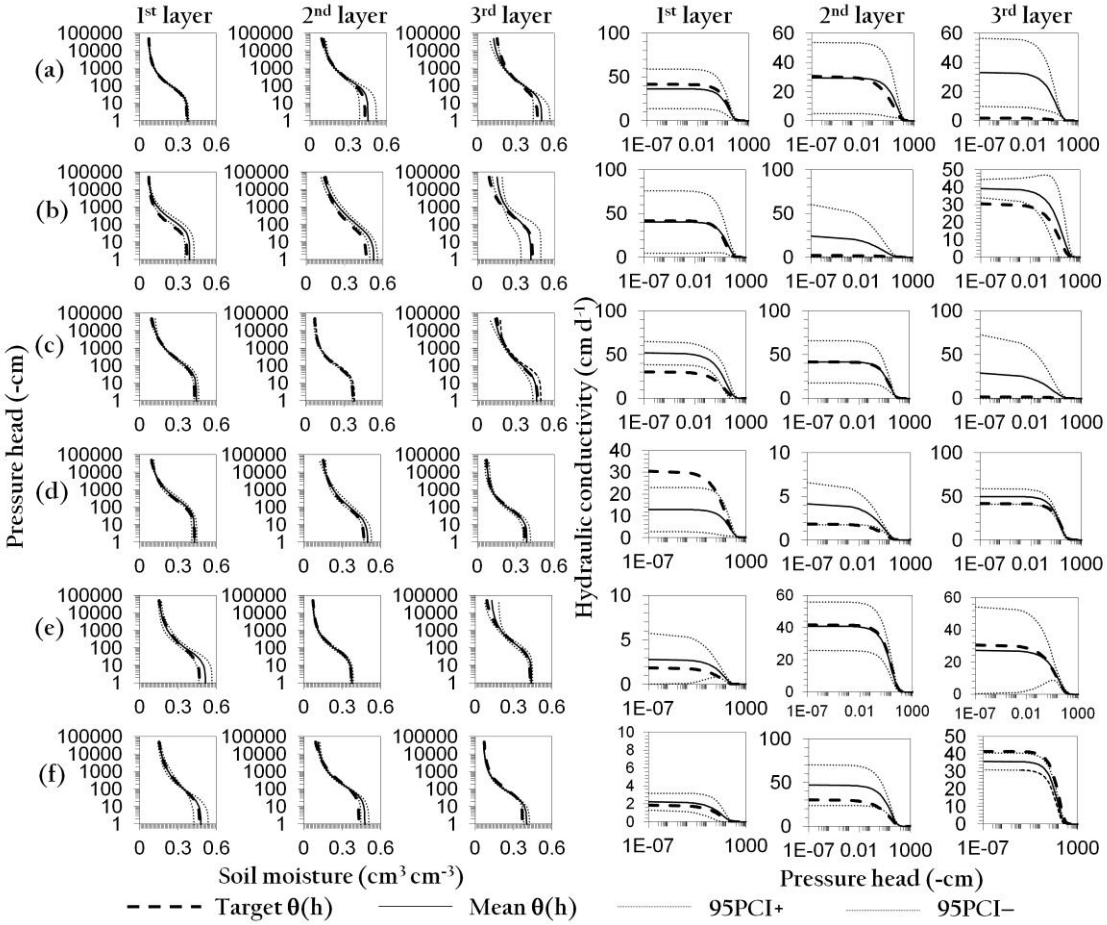

Figure 2.4: Derived θ(h) and K(h) functions of the layered soil column with grass for Case 1 using the layer-specific soil moisture assimilation scheme; (a) CB1: 1st sandy loam, 2nd silt loam, 3rd clay loam, (b) CB2: 1st sandy loam, 2nd clay loam, 3rd silt loam, (c) CB3: 1st silt loam, 2nd sandy loam, 3rd clay loam, (d) CB4: 1st silt loam, 2nd clay loam, 3rd sandy loam, (e) CB5: 1st clay loam, 2nd sandy loam, 3rd silt loam, (f) CB6: 1st clay loam, 2nd silt loam, 3rd sandy loam

Table 2.4 presents the correlations (R^2) and uncertainties (MAE) of observed and simulated soil moisture dynamics in the top portion (near the soil layer interfaces) of the 1st (between 0-5 cm), 2nd (between 10-15 cm), and 3rd (between 60-70 cm) soil layers for the 6 combinations involving three different soil types at the ARS 134 site. Mostly, the simulated soil moisture estimates in the soil layers matched well with the observations in the range of R^2 (1st: 0.974-0.999, 2nd: 0.978-0.998, 3rd: 0.980-0.997) and MAE (1st: 0.004-0.016, 2nd: 0.004-0.020, 3rd: 0.001-0.012) as shown in Fig. 2.4.

Figure 2.5 shows the daily precipitation, water stress (T_{act}/T_{pot}), and soil moisture changes for CB 5 (only shown here for CB5 case). Under the rain-fed condition, the water stress by the crop (grass) corresponded to the weather condition. As the near-surface soil moisture becomes dry, the water stress level for the CB 5 increased considerably during the dry periods indicating that the plant activities were affected by the dry condition near the land surface. The soil moisture estimates in all the layers were identified well with the target values and the uncertainties in the 1st layer (between 0-5 cm depth) are higher than those in the 2nd (between 10-15 cm depth) and 3rd (between 60-70 cm depth) layers. It is evident that there are uncertainties involved in $\theta(h)$ and $K(h)$ estimates, because the soil moisture estimates reflect uncertainties associated with various conditions (e.g. vegetation covers, soil hydraulic properties or meteorological data, as well as functional errors of the hydrological model itself for estimating the soil moisture). The correlation (R^2) and MAE values of the simulated and observed (hypothetical) soil moisture estimates in the numerical experiments range from 0.990 to 0.996 and 0.001 to 0.013, respectively.

Table 2.4: Correlations (R^2) and mean absolute error (MAE) of soil moisture dynamics at 0-5 cm, 10-15 cm, and 60-70 cm depths in the layered soil column using the θ(h) and K(h) functions derived by the layer-specific soil moisture assimilation scheme at the ARS 134 site for Case 1 (CB 1 to 6: free drainage with grass)

Depth	CB1		CB2		CB3		CB4		CB5		CB6	
	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE
1st (0-5 cm)	0.989	0.007	0.974	0.016	0.998	0.006	0.999	0.004	0.996	0.013	0.998	0.013
2nd (10-15 cm)	0.992	0.007	0.978	0.020	0.998	0.004	0.998	0.004	0.990	0.004	0.995	0.006
3rd (60-70 cm)	0.980	0.006	0.997	0.012	0.990	0.009	0.997	0.001	0.996	0.001	0.994	0.005

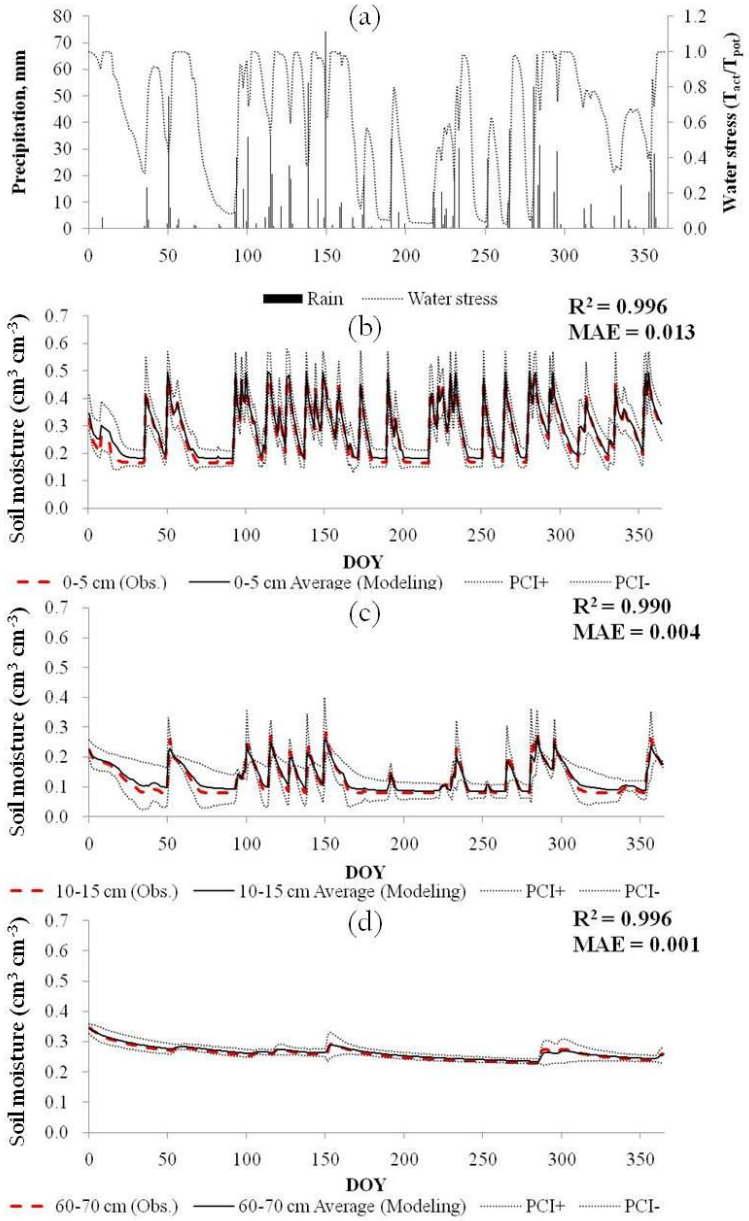

Figure 2.5: (a) Daily precipitation (mm) and water stress (T_{act}/T_{pot}) and (b-d) root zone soil moisture dynamics (cm^3 cm^{-3}) at 0-5 cm, 10-15 cm, and 60-70 cm depths in the layered soil column using the θ(h) and K(h) functions derived by the layer-specific soil moisture assimilation scheme at the ARS 134 site (Case 1- CB 5: free drainage with grass)

Various land covers (e.g., bare soil, grass, and wheat) were applied to CB 5 as shown in Table 2.5. The soil hydraulic parameters with the bare soil cover were better identified with the target values than those with grass and wheat, although the estimates in the 2nd and 3rd layers have uncertainties, especially for K_{sat}. In the cases of grass and wheat, the parameter uncertainties in the 1st and 3rd layers were considerably higher than those in the 2nd layer. This is more evident when compared with the results of bare soil, which indicates that complexities incurred by plant root activities to soil moisture dynamics in the root zone. The parameters in the 3rd layer have more uncertainties than those in the 1st and 2nd layers with all vegetation covers as shown in Table 2.5. There are no apparent differences between grass and wheat vegetations, although the K_{sat} values in the 2nd and 3rd layers with grass are better identified than those with wheat.

Table 2.5: Solutions of the layer-specific soil moisture assimilation scheme of CB 5 with bare soil, grass, and wheat vegetations (1st: clay loam, 2st: sandy loam, 3rd: silt loam) for Case 1 – layered soil column

Soil Layers	Parameter	Target Values*	Bare soil		Grass		Wheat	
			Mean	PCI	Mean	PCI	Mean	PCI
1st	α	0.030	0.031	0.028-0.033	0.028	0.020-0.036	0.031	0.029-0.032
	n	1.370	1.393	1.292-1.493	1.492	1.269-1.715	1.423	1.332-1.514
	θ_{res}	0.129	0.132	0.088-0.177	0.141	0.106-0.176	0.141	0.114-0.169
	θ_{sat}	0.470	0.474	0.457-0.490	0.515	0.463-0.566	0.480	0.465-0.495
	K_{sat}	1.840	2.397	0.765-4.029	2.817	-0.140-5.775	2.004	1.808-2.200
2nd	α	0.021	0.027	0.018-0.035	0.023	0.019-0.028	0.022	0.021-0.024
	n	1.610	1.577	1.562-1.592	1.605	1.596-1.614	1.597	1.577-1.617
	θ_{res}	0.067	0.062	0.060-0.064	0.065	0.061-0.068	0.065	0.061-0.069
	θ_{sat}	0.370	0.376	0.371-0.381	0.375	0.365-0.384	0.373	0.364-0.381
	K_{sat}	41.600	52.466	43.918-61.015	39.984	23.391-56.578	31.019	17.579-44.459
3rd	α	0.012	0.010	0.005-0.014	0.014	0.005-0.024	0.014	0.010-0.017
	n	1.390	1.384	1.218-1.550	1.530	1.368-1.692	1.446	1.320-1.572
	θ_{res}	0.061	0.063	0.059-0.066	0.117	0.050-0.184	0.119	0.021-0.217
	θ_{sat}	0.430	0.404	0.380-0.429	0.441	0.432-0.450	0.429	0.421-0.437
	K_{sat}	30.500	16.446	8.964-23.927	27.237	0.280-54.193	26.960	24.898-29.021

*UNSODA database [*Leij et al.,* 1999]

2.4.2 Case 2: Layered Soil Column under Varying Ground Water Table Conditions

This analysis is done only for the CB 5 soil-layering scenario. Table 2.6 shows the summary of results of the layered soil column with the ground water table depths of -200, -150, and -100 cm from the soil surface. We can see a visible trend which indicates that the soil hydraulic parameters in the 1st, 2nd, and 3rd layers with the presence of deeper ground water (GW) table depth of -200 cm are identified better than the estimates for the scenarios with shallow GW tables at -150 and -100 cm. The α, n, and θ_{res} values for GW -200 cm in the 1st layer correspond well with the target values, while only the α values for GW -150 and -100 cm have a good matching with the target values. For the 2nd layer, the solutions of the hydraulic parameters (α, n, θ_{res}, and θ_{sat}) for all the ground water table depths of -200, -150, and -100 cm are better matched compared to the results of the 1st layer indicating that it may be affected by the root activities more than other layers as shown in Table 2.5. Overall, the parameter estimations at the GW -200 cm are matched better with the target values than those at the GW -150 and -100 cm. Especially, as the ground water table is lowered, the n, θ_{sat}, and K_{sat} values at the GW -150 cm in the 1st layer were identified better than those for the GW -100 cm indicating that the parameter estimations at the upper layers are influenced by the upward flows from shallow ground water table [see *Ines and Mohanty*, 2008a].

In the 3rd layer, close to ground water boundary, only α and K_{sat} values are identifiable with the target values. The errors of estimation in the 3rd layer are considerably worse than those in the 1st and 2nd layers. The inverse solutions for Case 2

(in the presence of ground water tables) have more uncertainties than for Case 1 (well drained). In general, the uncertainty range ($\pm$ 95PCI) of soil hydraulic parameters with GW at -200 cm is smaller than those for GW at -150 and -100 cm. It confirms that soil hydraulic estimates in the layered soil column are governed not only by soil layering but also by the bottom boundary conditions, especially in the presence of shallow ground water table.

Table 2.6: Solutions of the layer-specific soil moisture assimilation scheme of CB 5 (1st: clay loam, 2st: sandy loam, 3st: silt loam) for Case 2 – layered soil column with ground water tables (-200, -150, and -100 cm)

Soil Layers	Parameter	Target Values*	GW -200 cm		GW -150 cm		GW -100 cm	
			Mean	PCI	Mean	PCI	Mean	PCI
1st	α	0.030	0.032	0.029-0.034	0.031	0.027-0.034	0.032	0.030-0.033
	n	1.370	1.372	1.235-1.508	1.389	1.237-1.542	1.459	1.189-1.728
	θ_{res}	0.129	0.127	0.098-0.156	0.136	0.089-0.184	0.136	0.076-0.196
	θ_{sat}	0.470	0.502	0.444-0.560	0.481	0.465-0.497	0.498	0.415-0.582
	K_{sat}	1.840	5.815	-1.383-13.013	2.390	1.448-3.333	3.637	0.107-7.166
2nd	α	0.021	0.020	0.019-0.021	0.020	0.016-0.023	0.022	0.016-0.029
	n	1.610	1.587	1.563-1.611	1.600	1.577-1.622	1.600	1.588-1.612
	θ_{res}	0.067	0.063	0.057-0.070	0.065	0.058-0.072	0.097	0.042-0.152
	θ_{sat}	0.370	0.375	0.365-0.385	0.374	0.362-0.385	0.371	0.367-0.375
	K_{sat}	41.600	30.290	24.131-36.449	27.303	25.044-29.562	34.917	-1.796-71.629
3rd	α	0.012	0.013	0.006-0.019	0.009	0.003-0.016	0.008	0.003-0.013
	n	1.390	1.497	1.366-1.628	1.434	1.109-1.760	1.515	1.419-1.610
	θ_{res}	0.061	0.141	0.104-0.178	0.122	0.069-0.175	0.125	0.052-0.198
	θ_{sat}	0.430	0.423	0.356-0.490	0.411	0.361-0.462	0.422	0.405-0.440
	K_{sat}	30.500	31.778	14.576-48.980	20.487	10.982-29.992	51.981	37.939-66.023

*UNSODA database [*Leij et al.*, 1999]

2.4.3 Case 3: Field Validation Experiment

Figures 2.6 and 2.7 show the daily rainfall and simulated/observed soil moisture in the 1st (LW 02: 0-30 cm and LW 07: 0-20 cm), 2nd (LW 02: 30-60 cm and LW 07: 20-40 cm), and 3rd (LW 02: 60-90 cm and LW 07: 40-60 cm) layers at the field sites during the simulation period based on the inverse modeling. In general, $\theta(h)$ at the LW 02 and 07 sites derived by the layer-specific soil moisture assimilation scheme matched well with the observations, although uncertainties exist in the estimated $\theta(h)$ functions for the 1st, 2nd, and 3rd layers. When we compared the K_{sat} values of UNSODA database and laboratory-based experiments in Table 2.1 and 2.3, the laboratory-based K_{sat} values were extremely higher than those of UNSODA database due to measurement errors. Thus, we excluded the K(h) functions for further analysis due to its non-sensitivity. The simulated soil moisture (1st: R^2=0.998 and MAE=0.011, 2nd: R^2=0.997 and MAE=0.001, 3rd: R^2=0.992 and MAE=0.002 for LW 02, and 1st: R^2=0.991 and MAE=0.005, 2nd: R^2=0.992 and MAE=0.003, 3rd: R^2=0.993 and MAE=0.003 for LW 07) estimates in the layered soil column identified well with the estimates derived by the soil hydraulic parameters taken near the LW 02 and 07 sites. The soil moisture estimates in the 1st layer at the field sites have more uncertainties than those in 2nd and 3rd layers. These results are in agreement with the results of CB5 in Case 1, which indicate that the estimated soil moisture with the grass cover in the 2nd and 3rd layers are closer to the target values than those in the 1st layer, as shown in Figure 2.5. The derived soil hydraulic properties compared well with UNSODA, based on dominant textural class.

Figure 2.8 shows the measured (TDR-based) and simulated soil moisture dynamics in the soil layers at the LW 11 site. Overall, the simulated results (R^2: 0.891, MAE: 0.018 for the 1st layer; R^2: 0.967, MAE: 0.006 for the 2nd layer; R^2: 0.894, MAE: 0.034 for the 3rd layer) for the calibration period matched well with the measurements. The result (R^2: 0.965, MAE: 0.051 for the 1st layer; R^2: 0.891, MAE: 0.035 for the 2nd layer; R^2: 0.949, MAE: 0.037 for the 3rd layer) for the validation period also shows the good matching in the soil layers at the field site. The derived soil hydraulic parameters by inverse modeling in the 1st layer (0-20cm) compared well with the independently measured soil hydraulic parameters from laboratory. Table 2.7 presents the uncertainty analysis of estimated soil moisture dynamics for the soil layers using various objective functions (additive absolute value, multiplicative absolute value, and additive square delta forms) with three different methods (MAE, MBE, and RMSE) at the LW 11 site. It is clear that the additive absolute value form of the objective function used in this study produced better results than by the multiplicative and square delta forms for the calibration and validation. Also, The MAE and RMSE performed similarly during the calibration and validation whereas the MBE was less sensitive than others.

Although this method has a limitation (available measurements in the soil layers), it gives us insights of the implication/impact of soil heterogeneity and layering in quantifying soil hydraulic parameters in the layered soil column. With more *in-situ* soil moisture networks in place globally (e.g., Oklahoma Mesonet, Soil Climate Analysis Network: SCAN, USDA-Agricultural Research Service: USDA-ARS network, National Ecological Observatory Network: NEON, International Soil Moisture Network: ISMN,

etc.) and data available in the recent years at multiple soil depths from benchmark experiments, this layer-specific assimilation method can prove to be quite useful for predicting the soil moisture dynamics in the soil layers.

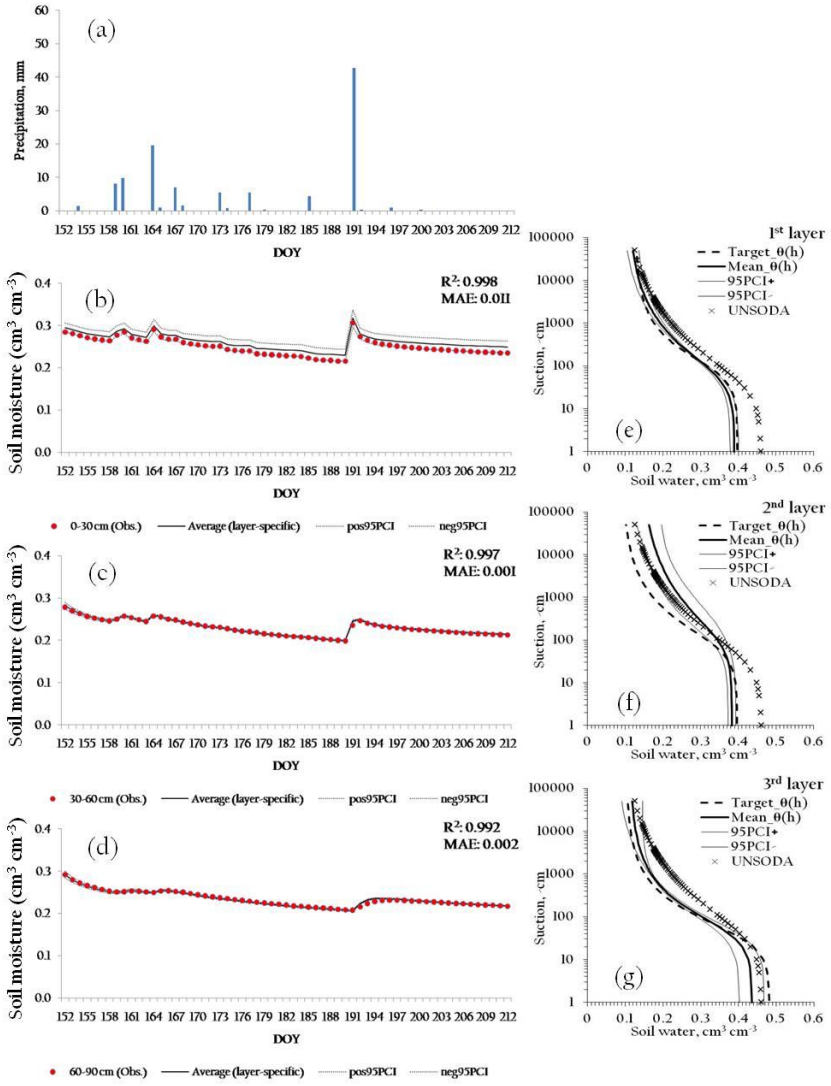

Figure 2.6: (a) Daily precipitation (mm), (b-d) observed/simulated root zone soil moisture dynamics, and (e-g) θ(h) functions of target, derived solutions, and UNSODA database (dominated by loam soil) at the 1st (0-30 cm), 2nd (30-60 cm), and 3rd (60-90 cm) in the LW 02 site (ARS 136) using the layer-specific soil moisture assimilation scheme

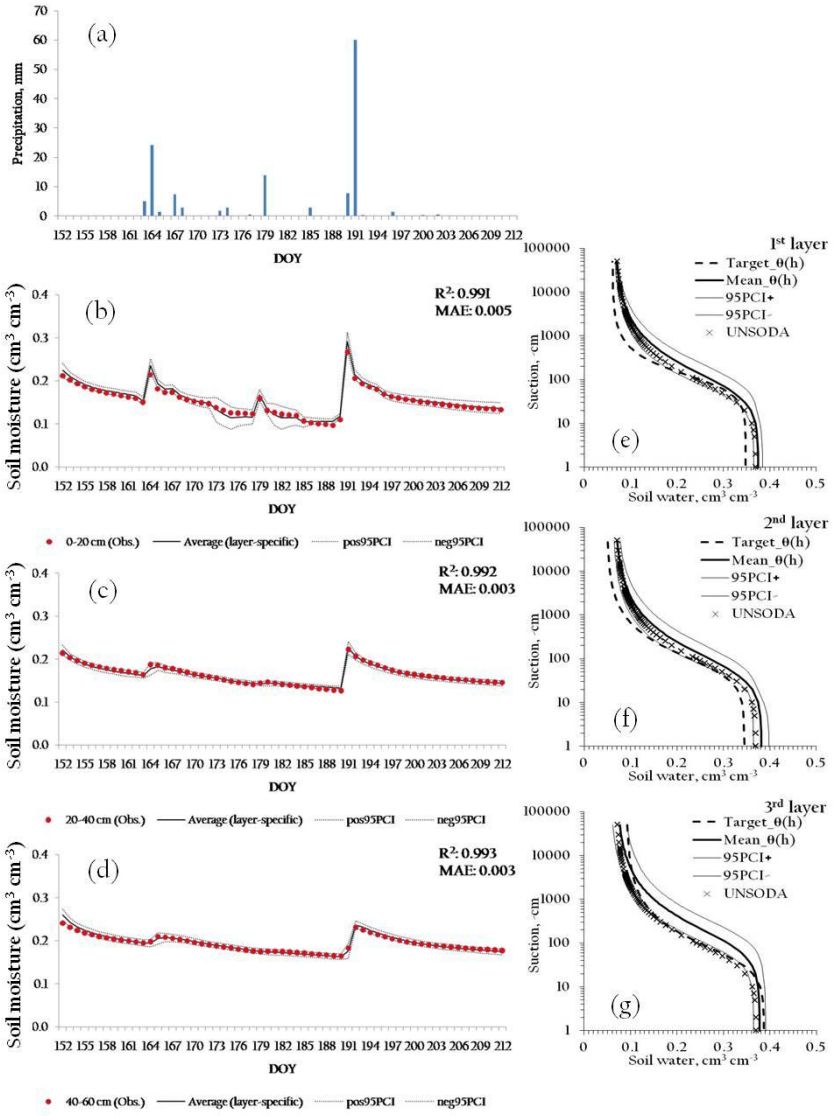

Figure 2.7: (a) Daily precipitation (mm), (b-d) observed/simulated root zone soil moisture dynamics, and (e-g) θ(h) functions of target, derived solutions, and UNSODA database (dominated by sandy loam soil) at the 1st (0-30 cm), 2nd (30-60 cm), and 3rd (60-90 cm) in the LW 07 site (ARS 151) using the layer-specific soil moisture assimilation scheme

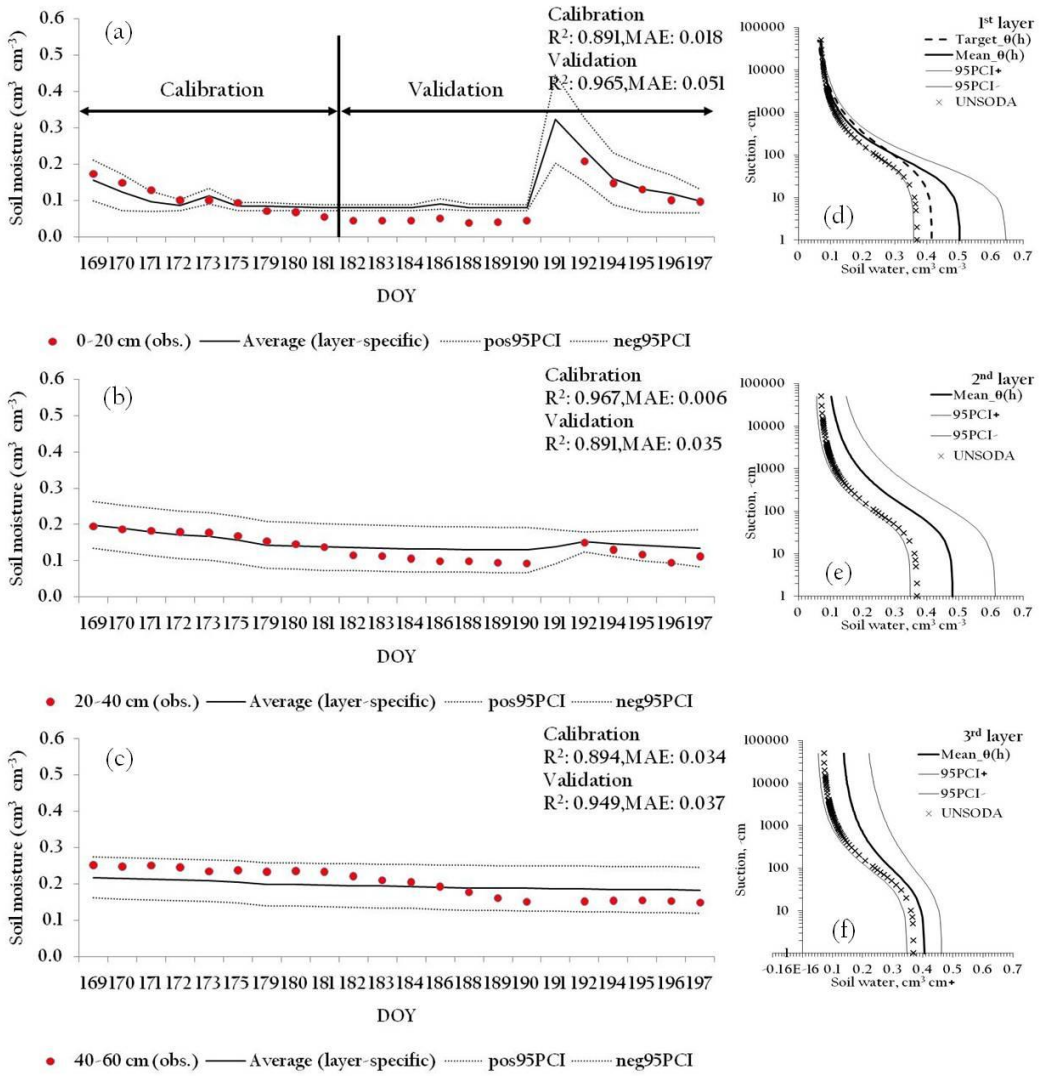

Figure 2.8: (a-c) Observed (*in-situ*) and simulated root zone soil moisture dynamics and (d-f) θ(h) functions of target (1st layer was only measured), derived solutions, and UNSODA database (dominated by sandy loam soil) at the 1st (0-20 cm), 2nd (20-40 cm), and 3rd (40-60 cm) in the LW 11 site (ARS 136) using the layer-specific soil moisture assimilation scheme during the calibration (DOY 169-181) and validation (DOY 182-197) periods

Table 2.7: Uncertainty analysis using various objective functions (additive absolute value, multiplicative absolute value, and additive square delta forms) with three different methods (Mean Absolute Error-MAE, Mean Bias Error-MBE, and Root Mean Square Error-RMSE) based on the layer-specific soil moisture assimilation scheme during the calibration (DOY 169-181) and validation (DOY 182-197) periods

	Soil Layers	Additive Absolute form				Multiplicative form				Additive Square delta form			
		R^2	MAE	MBE	RMSE	R^2	MAE	MBE	RMSE	R^2	MAE	MBE	RMSE
Calibration	1st*	0.891	0.018	0.004	0.006	0.972	0.017	-0.017	0.007	0.979	0.027	-0.027	0.010
	2nd**	0.967	0.006	0.004	0.002	0.971	0.058	-0.058	0.019	0.963	0.013	0.008	0.005
	3rd***	0.894	0.034	0.034	0.011	0.886	0.067	-0.067	0.022	0.891	0.021	0.021	0.007
Validation	1st*	0.965	0.051	-0.051	0.026	0.994	0.073	-0.073	0.032	0.996	0.079	-0.079	0.034
	2nd**	0.891	0.035	-0.035	0.013	0.023	0.094	-0.094	0.028	0.009	0.047	-0.046	0.016
	3rd***	0.949	0.037	-0.029	0.016	0.940	0.133	-0.133	0.040	0.943	0.044	-0.041	0.018

*0-20 cm
**20-40 cm
***40-60 cm

2.5 Conclusions

In this study, a layer-specific soil moisture assimilation procedure based on simulation-optimization (SWAP-GA) scheme was developed to quantify effective soil hydraulic parameters in the layered soil profile. Various numerical experiments with the conditions of free drainage, presence of ground water tables at several different depths, different vegetation covers, and field experiments were conducted. The impacts of soil layers, heterogeneity of different soil textures, and different land covers in a vertically layered soil column were evaluated in Case 1 using the layer-specific soil moisture assimilation scheme. Case 2 was conducted to evaluate the impacts of various ground water table depths with a grass cover. The field experiments of Case 3 were conducted for assessing the applicability of this approach at the field-scales (LW 02, 07, and 11 sites).

The results of Case 1 show that the soil layers and order/sequence of vertical heterogeneity of soil textures affect the uncertainties of parameter estimations due to complex signature of soil water in the layered soil profile. Also, the estimated parameters in the 1^{st} and 3^{rd} layers with the grass and wheat covers have relatively more errors than that of the bare soil. It may indicate that the hydrological model has the own weakness for simulating plant root activities in the root zone. In Case 2, we found that as the ground water table becomes deeper, the estimates of soil hydraulic parameters improved as well as the results with the free drainage condition. These results suggest that the bottom boundary condition has a large influence on the hypothesis of layer-specific data assimilation studies. In the field experiments of Case 3, the soil moisture

dynamics and θ(h) functions were only estimated at the LW 02 and 07 sites using the layer-specific soil moisture assimilation scheme, because of the limited observations. The simulated near-surface and sub-surface soil moisture estimates at the field sites identified well with the field observations (derived by the soil hydraulic parameters obtained from the soil core samples collected at the field sites), even though the soil moisture estimates near the land surface have slightly higher uncertainties than those in the deeper soil layers. The simulated soil moisture dynamics in the soil layers were also matched well with the *in-situ* measurements for the LW 11 site. It suggests that the layer-specific assimilation scheme based on inverse modeling could be used to model soil moisture dynamics in the layered soil profile even with the limited soil moisture measurements in the real world conditions. We envisaged that the new soil moisture assimilation procedure would be useful for vadose zone and land surface modeling in well-instrumented hydrologic system. In future, with the advent of advanced soil moisture remote sensing capabilities with deeper penetrating depths, this layer-specific assimilation platform can be useful for estimating large-scale effective soil hydraulic properties under heterogeneous/layered soil condition, as the near-surface assimilation proved to be useful in homogeneous soil conditions in the past studies (*Ines and Mohanty,* 2008a,b, 2009).

CHAPTER III

QUANTIFYING EFFECTIVE SOIL HYDRAULIC PROPERTIES USING PIXEL-BASED SOIL MOISTURE AND EVAPOTRANSPIRATION PRODUCTS AT MULTIPLE SCALES

3.1 Synopsis

With the development of many earth-observing remote sensing platforms, pixel-based products are becoming critical inputs in many hydrological and meteorological models. Remotely sensed soil moisture (SM) and ground-based evapotranspiration (ET) have been used in the past to estimate pixel-scale soil hydraulic parameters [*Ines and Mohanty,* 2008a,b]. However, effect of disparate spatial support scales for SM and ET estimation were not accounted in those efforts. In this study, we developed a new inverse modeling algorithm by combining soil moisture based Noisy Monte Carlo Genetic Algorithm (NMCGA) [*Ines and Mohanty* 2009] and Surface Energy Balance Algorithm for Land (SEBAL) [*Daroonwan et al.,* 2008] based ET products (using remote sensing data) for quantifying pixel-scale effective soil hydraulic properties. This coupled NMCGA and SEBAL algorithm was evaluated for estimating the effective soil hydraulic parameters of the root zone (0-200 cm) under synthetic scenarios (e.g., different soil textures, weather conditions, vegetation covers, and ground water table depths) and the results were tested using field validation experiments from point- to satellite-scale. For comparison purposes, inverse modeling results were analyzed under three scenarios; 1) evapotranspiration (ET) only in the optimization algorithm, 2) soil moisture (SM) only in the optimization algorithm, and 3) soil moisture and evapotranspiration (SM+ET)

jointly in the optimization criteria. When we considered both SM and ET components jointly in the optimization algorithm, it improved the estimations of effective soil hydraulic properties and soil moisture fluxes and reduced their uncertainties better than those of using SM-only (NMCGA) or ET-only (SEBAL). Similar results were shown in the presence of ground water table depth (-100 cm from the soil surface) indicating that the ET component plays a key role in the soil profile dominated by upward flow from ground water. For the field validation experiments, the SM and ET estimates derived by the optimized parameters under the (SM+ET) joint criterion matched the observations at various spatial scales better than those of the SM-only criterion. Also, the estimated SM and ET at the airborne sensing-scale provide more reasonable statistics in both the spatial and temporal scales than those of the point- and satellite-scales. These results demonstrate the robustness of our approach providing estimates of effective soil hydraulic properties using pixel-based SM and ET products at multiple scales.

3.2 Introduction

Soil hydraulic properties at the land surface and in the unsaturated zone are key variables for many hydro-climatic processes [*Hansen et al.,* 1999; *Mohanty et al.,* 2002; *Mohanty and Zhu,* 2007], because they are necessary to characterize effective hydraulic behavior of the soil system [*Wood,* 1994; *Vrugt et al.,* 2003; *Jana and Mohanty,* 2012a,b,c]. In general, laboratory methods are used to determine the soil hydraulic functions by direct measurement using soil core samples extracted from field. The major concern of this traditional method is the question of whether the soil parameters derived by a cm-scale soil core sample with pre-defined boundary conditions can represent field-

and regional-scale features [*Kool and Parker*, 1988; *Van Dam*, 2000, *Das et al.,* 2008a, 2011]. With increasing availability of remote sensing (RS) platforms characterized by large spatial and temporal coverage, an inverse modeling is better suited for estimating spatially distributed effective soil hydraulic properties [*Yeh,* 1989; *Ines and Mohanty,* 2008a,b, 2009] and provides an attractive alternative to direct measurement.

From the hydrologic process point of view, soil moisture (SM) dynamics and evapotranspiration (ET) dynamics play significant and complementary roles for water balance in the vadose zone [*Xevi et al.,* 1996]. As both the SM and ET estimates are spatially distributed, using remote sensing data (instead of point measurements) in the inverse estimation provide a more uniform representation of effective soil hydraulic properties for a regional extent. In this context remotely sensed soil moisture products have been used by *Ines and Mohanty* [2008a,b]. However, to our knowledge, no remote sensing based ET product has been used to date for inverse estimation of effective soil hydraulic properties at the field-/footprint-scale. Over the past decade, several pixel-based RS schemes (SEBAL [*Bastiaanssen et al.,* 1998, 2005], METRIC [*Allen et al.,* 2007], simplified-SEBI [*Roerink et al.,* 2000], and SHEBA [*Su,* 2002]) have been developed for estimating spatially distributed ET.

Estimation of soil hydraulic parameters using different hydrological models for the inversion could be quite complex indicating that a unique optimum soil hydraulic parameter set is impractical. Alternatively, *Aronica et al.,* [2002] suggested Generalized Likelihood Uncertainty Estimation (GLUE) [*Beven and Binley,* 1992*; Beven and Freer,* 2001] using Monte Carlo simulation techniques [*Kuczera and Parent,* 1998; *Bates and*

Campbell, 2001] to improve the search algorithm in a Bayesian framework, which provide dependable simulations for a range of model inputs. From the effective parameter estimation perspective, a number of methods have been developed such as Genetic Algorithm (GA) [*Wang,* 1991; *Ines and Droogers,* 2002; *Ines and Mohanty,* 2008a,b, 2009], Shuffled Complex Evolution-University of Arizona (SCE-UA) [*Duan et al.*, 1992], and Artificial Neural Networks (ANNs) [*Pachepsky et al,* 1996; *Schaap and Bouten,* 1996; *Schaap and Leij,* 1998; *Schaap, et al.,* 1998; *Jana et al.,*2007]. *Ines and Mohanty* [2008a,b, 2009] developed the near-surface soil moisture assimilation scheme and Noisy Monte Carlo Genetic Algorithm (NMCGA) based on the inverse modeling for quantifying the effective soil hydraulic properties using the remote sensing (RS) soil moisture (SM) products. In their study, effective soil hydraulic properties and their uncertainties were estimated simultaneously using the near-surface soil moisture derived from a RS footprint. Both point-scale SM and ET values were also considered for the parameter estimations under the synthetic condition [*Ines and Mohanty*, 2008a]. But there was no consideration of remotely sensed ET products while searching for the effective soil hydraulic parameters with the pixel-based SM. The land surface heterogeneity composed of various soil textures and vegetation covers indicate that point-scale ET may influence uncertainties in extracting soil hydraulic parameters from RS soil moisture products.

In this study, our primary goal is to develop an improved inverse modeling algorithm for pixel-based effective soil hydraulic property estimation at multiple scales using remotely sensed SM and ET products. Main objectives are three-fold: 1) to

develop the coupled data assimilation algorithm based on NMCGA and SEBAL, which integrates pixel-based SM and ET components in order to quantify the pixel-scale effective soil hydraulic properties, 2) to analyze the impact of SM-only, ET-only, and SM+ET-jointly in the optimization criteria for the soil hydraulic parameter estimations under various wetness range, and 3) to evaluate the robustness of this approach at different spatial scales ranging from point, airborne sensing, to satellite.

3.3 Materials and Methods

3.3.1 Coupled Data Assimilation Algorithm

Figure 3.1 shows the framework of coupled NMCGA and SEBAL based near-surface SM and ET assimilation scheme. We can define this approach as the coupled NMCGA and SEBAL algorithm to determine the effective soil water content ($\theta(h)$) and hydraulic conductivity ($K(h)$) functions in the soil profile using the remotely sensed near-surface (0-1 or 0-5 cm) SM and ET estimates based on a simulation-optimization approach. The data assimilation scheme repeatedly runs the hydrologic model until the derived solutions (soil hydraulic properties) have converged [*Abbaspour et al.,* 1997; *Kool and Parker,* 1988; *Ines and Droogers,* 2002].

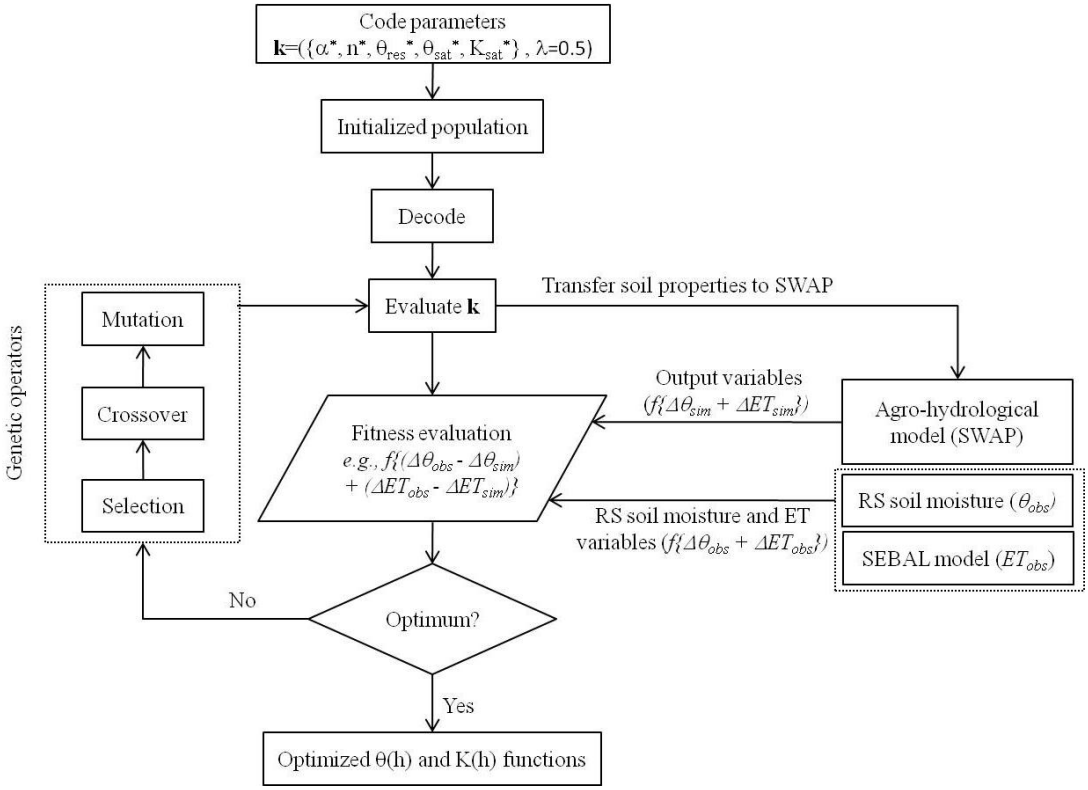

Figure 3.1: Framework of the coupled NMCGA and SEBAL algorithm used in this study

3.3.2 Hydrological Model

The Soil Water Atmosphere Plant (SWAP) model is a one-dimensional (1-D) physically-based model for simulating water flow across soil, water, atmosphere, and plant system [*Kroes et al.,* 1999*; van Dam et al.,* 1997]. In the SWAP model, water flow is modeled using Richards' equation (Eq. (3.1)). The soil hydraulic functions in the soil column can be described by analytical expressions of *van Genuchten* [1980] and *Mualem* [1976] using the relationship between the soil water content (θ), pressure head (h), and unsaturated hydraulic conductivity (K).

$$\frac{\partial \theta}{\partial t} = C(h)\frac{\partial h}{\partial t} = \frac{\partial [K(h)(\frac{\partial h}{\partial z} + 1)]}{\partial z} - S(h) \qquad (3.1)$$

where θ is the volumetric water content ($cm^3\ cm^{-3}$), K is the hydraulic conductivity (cm d^{-1}), h is the soil water pressure head (-cm), z is the vertical soil depth (cm) taken positively upward, t is the time (d), C is the differential soil water capacity (cm^{-1}), and S(h) is the actual soil water extraction rate by plants ($cm^3\ cm^{-3}\ d^{-1}$) defined as Eq. (3.2).

$$S(h) = \alpha_w(h)\frac{T_{pot}}{Z_r} \qquad (3.2)$$

where T_{pot} is the potential transpiration (cm d^{-1}), Z_r is the rooting depth (cm), and α_w is a reduction factor as function of h and accounts for water deficit and oxygen stress [*Feddes et al.*, 1978]. The Richards' equation (3.1) using the finite difference scheme as described by *Belmans et al.* [1983] allows the use of soil hydraulic databases and all kinds of management scenarios.

$$S_e = \frac{\theta(h) - \theta_{res}}{\theta_{sat} - \theta_{res}} = [\frac{1}{1 + |\alpha h|^n}]^m \qquad (3.3)$$

$$K(h) = K_{sat} S_e^{\lambda} [1 - (1 - S_e^{1/m})^m]^2 \qquad (3.4)$$

where S_e is the relative saturation (-), θ_{res} is the residual water content ($cm^3\ cm^{-3}$) in the dry range, and θ_{sat} is the saturated water content ($cm^3\ cm^{-3}$), α, n, m, and λ are shape parameters, K_{sat} is the saturated hydraulic conductivity (cm d^{-1}), and $m=1 - \frac{1}{n}$, respectively.

Various conditions for the top (atmospheric) and bottom boundary conditions (free drainage condition and ground water table depths) can be simulated by the SWAP

model [*van Dam et al.,* 1997]. The SWAP model contains simple and detailed (WOFOST) crop growth routines. A simple model simulates the impacts of climate, soil textures, plant types, and water managements. In this study, the simple crop model combined with the water management modules (e.g., irrigation and drainage) [*van Dam et al.,* 1997, *van Dam,* 2000] is used. Using the Penman-Monteith equation, the SWAP model calculates not only the potential and actual soil evaporation (E_{pot} and E_{act}), but also plant transpiration (T_{pot} and T_{act}) partitioned by the leaf area index (LAI) or soil cover fraction (SC) of the land unit. This model performs well under various climatic and environmental conditions [*Wesseling and Kroes,* 1998; *Sarwar et al.,* 2000; *Droogers et al.,* 2000; *Singh et al.,* 2006a,b].

3.3.3 SEBAL Model

Evapotranspiration (ET) across the land-atmosphere boundary is generated by the water loss from open water, soil, and plant surface. ET is governed by the energy and heat exchanges at the land surface as illustrated in Figure 3.2. The computation of ET by the Surface Energy Balance Algorithm for Land (SEBAL) model includes three steps: 1) preliminary processing of remote sensing data (producing normalized difference vegetation index, NDVI, albedo, and temperature images), 2) selection of hot and cold pixel (hot pixel assumes all the energy used to heat the surface and cold pixel assumes all the energy used to evaporate water), and 3) estimation of pixel-based ET [*Daroonwan et al.,* 2008]. The key input datasets for the SEBAL model consist of spectral radiances in the visible, near-infrared, and thermal infrared wavelengths (e.g., Moderate Resolution Imaging Spectroradiometer (MODIS) surface reflectance products,

solar zenith angle, quality control band, land surface temperature, and band emissivities 31 and 32), weather data (wind speed), and regional information (e.g., height of vegetation, altitude of target area, specific locations of hot and cold pixels), and day of year (DOY). ET is related to the surface-energy balance (Eq. (3.5)) which is defined as

$$R_n = G_0 + H + LE \quad (3.5)$$

where R_n ($W \cdot m^{-2}$) is the net radiation absorbed at the land surface, G_0 ($W \cdot m^{-2}$) is the soil heat flux, H ($W \cdot m^{-2}$) is the sensible heat flux, and LE ($W \cdot m^{-2}$) is the latent heat flux associated with ET.

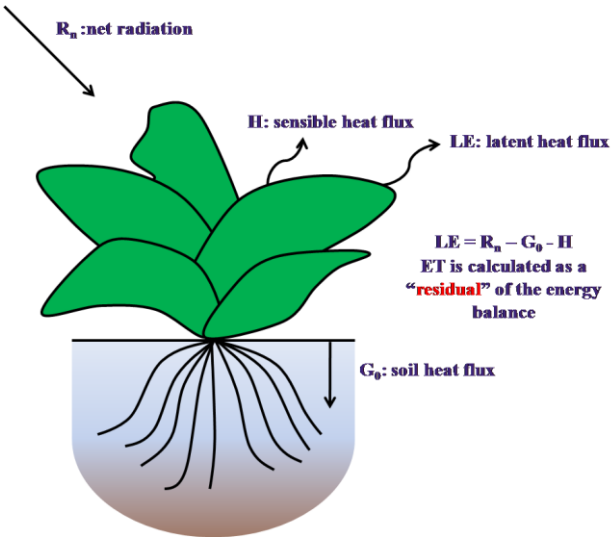

Figure 3.2: Schematic diagram of the surface energy balance for computing evapotranspiration (ET)

3.3.4 Genetic Algorithm Implementation

Genetic Algorithms (GAs) developed by Holland and his colleagues are based on the theory of "survival of the fittest" by Darwin [*Holland*, 1975; *Goldberg*, 1989]. GAs

are useful for searching optimal solutions from unknown variable spaces. Basically, GAs create the chromosome which contains the variable information using a "string" structure of binary digits (i.e. 0 and 1). These binary strings recursively compete to survive in the mating pool to produce the next generation through GA operations of selection, crossover, and mutation.

1) Selection: In the selection process, the strongest chromosome competes against others and survives to be selected. Then the rest of them die (discard).
2) Crossover: The surviving binary chromosomes then exchange their genetic information through mating during crossover to produce their offspring.
3) Mutation: The freshly selected chromosomes are mutated to generate new genetic materials for the next generation. Finally, the mutated chromosomes are restored instead of the certain genetic characteristics lost by degeneration.

Usually, GAs have numerous uncertainties ('noise') for solving real-world problems. Noisy genetic algorithms (NGAs) operated in a noisy environment are suggested for fitness evaluation of the chromosome variables subjected to a stochastic field [*Miller and Goldberg*, 1996; *Wu et al*., 2006]. *Ines and Mohanty* (2008b) integrated a NGA with a re-sampling (ensemble: e) algorithm for the Monte Carlo (MC) simulation [*Efron,* 1982; *Miller and Goldberg,* 1996; *Miller,* 1997; *Ines and Mohanty,* 2008b], called the noisy Monte Carlo genetic algorithm (NMCGA), to determine the fittest chromosomes (effective soil hydraulic parameter sets: $\mathbf{P} = \{\alpha, n, \theta_{res}, \theta_{sat}, K_{sat}\}$). The input parameter sets for the hydrological model can be shown as $\mathbf{k} = \{\mathbf{P}, \lambda\}$. The shape parameter of lamda (λ=0.5) is fixed in modeling. We transferred the parameter set (**P**) as

parameter set ($\mathbf{P}^*=\{\alpha^*, n^*, \theta_{res}^*, \theta_{sat}^*, K_{sat}^*\}$) to consider the uncertainties (e.g., representing heterogeneities of land surface in a RS product) of individual soil parameters in modeling. Thus, the set ($\mathbf{k}=\{\mathbf{P}^*, \lambda\}$) is used in this approach. In the NMCGA, GA estimates the combinations of parameter statistics ($\mathbf{P}^*$) in Eq. (3.6). Then, the MC algorithm derives realizations (r) of parameter combinations based on statistics. The noisy fitness of parameter combinations for all the realizations derived by the MC simulation in one resampling event is estimated along the given generations.

In this study, we integrated the SEBAL model with the NMCGA for the parameter estimations. The coupled NMCGA and SEBAL algorithm quantifies the effective hydraulic parameters by minimizing the difference between the observed (pixel-based) and simulated estimates of both soil moisture and ET time series (t). The representations of parameter statistics (Eq. (3.6)) are shown in Table 3.1. The objective functions (OF_{SM} and OF_{ET}) for the SM and ET components are shown in Eq. (3.7) and (3.8) below.

$$\mathbf{P}^*=\{\mu_{(\alpha)}, \sigma^2_{(\alpha)}, \mu_{(n)}, \sigma^2_{(n)}, \mu_{(\theta res)}, \sigma^2_{(\theta res)}, \mu_{(\theta sat)}, \sigma^2_{(\theta sat)}, \mu_{(Ksat)}, \sigma^2_{(Ksat)}\} \quad (3.6)$$

$$OF_{SM}(\mathbf{k})_e=\mathrm{Min}\left(\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{N_{resample}}\left(\sum_{r=1}^{N_{resample}}SimSM(k_r)_{t,e}\right)-ObsSM_t\right|\right)\ \forall e \quad (3.7)$$

$$OF_{ET}(\mathbf{k})_e=\mathrm{Min}\left(\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{N_{resample}}\left(\sum_{r=1}^{N_{resample}}SimET(k_r)_{t,e}\right)-ObsET_t\right|\right)\ \forall e \quad (3.8)$$

Table 3.1: Representation of the effecitve soil hydraulic properties for the coupled NMCGA[a] and SEBAL algorithm

Parameters (P^*)	Initial Chromosomes	Case1 to 3		Number of Bits (L)	2^L
		Minimum values	Maximum values		
$\mu_{(\alpha)}$ (cm^{-1})	00101	0.006	0.033	5	32
$\sigma_{(\alpha)}$ (cm^{-1})	00101	0.000	0.033	5	32
$\mu_{(n)}$ (-)	110010	1.200	1.610	6	64
$\sigma_{(n)}$ (-)	110010	0.000	1.000	6	64
$\mu_{(\theta res)}$ ($cm^3\ cm^{-3}$)	0001111	0.000	0.800	7	128
$\sigma_{(\theta res)}$ ($cm^3\ cm^{-3}$)	0001111	0.000	0.020	7	128
$\mu_{(\theta sat)}$ ($cm^3\ cm^{-3}$)	00001	0.370	0.550	5	32
$\sigma_{(\theta sat)}$ ($cm^3\ cm^{-3}$)	00001	0.000	0.200	5	32
$\mu_{(Ksat)}$ ($cm\ d^{-1}$)	0101000101	1.840	55.700	10	1024
$\sigma_{(Ksat)}$ ($cm\ d^{-1}$)	0101000101	0.000	10.000	10	1024

[a]Global search space = 32×32×64×64×128×128×32×32×1024×1024 = 7.3787E+19

Note that $\mathbf{P}^*$ is the statistics ($\mu_{(\ )}$: means, $\sigma_{(\ )}$: standard deviations) of effective soil hydraulic parameters, $\mathbf{k}_r$ is the combinations of effective soil hydraulic parameters ($\mathbf{k}$) with realizations (r) generated by the MC resampling, where e is the resampling numbers or ensemble, $N_{resample}$ is the number of realizations (r) derived from each resampling event, SimSM($\mathbf{k}_r$) is the simulated soil moisture with $N_{resample}$(r), SimET($\mathbf{k}_r$) is the simulated ET with $N_{resample}$(r), ObsSM is the observed (pixel-based) soil moisture, and ObsET is the observed (pixel-based) ET, t is the time index, respectively. All variables (SM-$cm^3\ cm^{-3}$ and ET-mm d^{-1}) were normalized for the objective functions. The optimization is constrained as

$$\text{Constrained SM}(\mathbf{k})_{t,e}=\begin{cases}+\text{PCI}_{SM}(\text{SimSM}(k_r)_{t,e}) > \text{ObsSM}_t \\ -\text{PCI}_{SM}(\text{SimSM}(k_r)_{t,e}) < \text{ObsSM}_t\end{cases}, \forall t\ (\text{where } t \le T), \forall e \quad (3.9)$$

$$\text{Constrained ET}(\mathbf{k})_{t,e}=\begin{cases}+\text{PCI}_{ET}(\text{SimET}(k_r)_{t,e}) > \text{ObsET}_t \\ -\text{PCI}_{ET}(\text{SimET}(k_r)_{t,e}) < \text{ObsET}_t\end{cases}, \forall t \text{ (where } t \le T), \forall e \quad (3.10)$$

where

$$\pm \text{PCI}_{SM}(\text{SimSM}(k_r))_{t,e} = \frac{1}{N_{resample}} \sum_{r=1}^{N_{resample}} \text{SimSM}(k_r)_{t,e} \pm (\text{PCIfactor}_{SM}) \times$$

$$\sqrt{\frac{\sum_{r=1}^{N_{resample}} (\text{simSM}(k_r)_{t,e} - (\frac{1}{N_{resample}} \sum_{r=1}^{N_{resample}} \text{SimSM}(k_r)_{t,e}))^2}{(N_{resample}-1)}} \forall t \text{ (where } t \le T), \forall e \quad (3.11)$$

$$\pm \text{PCI}_{SM}(\text{SimET}(k_r))_{t,e} = \frac{1}{N_{resample}} \sum_{r=1}^{N_{resample}} \text{SimET}(k_r)_{t,e} \pm (\text{PCIfactor}_{ET}) \times$$

$$\sqrt{\frac{\sum_{r=1}^{N_{resample}} (\text{simET}(k_r)_{t,e} - (\frac{1}{N_{resample}} \sum_{r=1}^{N_{resample}} \text{SimET}(k_r)_{t,e}))^2}{(N_{resample}-1)}} \forall t \text{ (where } t \le T), \forall e \quad (3.12)$$

where, Constrained SM(**k**) is the logical constraint of SM, Constrained ET(**k**) is the logical constraint of ET, ±PCIfactor$_{SM}$ are the ±95 percent (%) confidence intervals of SM, and ±PCI factor$_{ET}$ are the ±95 percent (%) confidence intervals of ET.

The MC spectrum of simulated soil moisture and ET estimates is constrained by the observed soil moisture and ET with ± PCIfactor$_{SM}$ in Eq. (3.9) and ± PCIfactor$_{ET}$ in Eq. (3.10). For the fitness evaluation (Z(**k**)), the modified penalty method suggested by *Chan-Hilton and Culver* [2000] is used with the weighing factor (f):

$$Z(\mathbf{k})_e = \{f \times (OF_{SM}(\mathbf{k})_e) \times (1+\text{Penalty}_{SM}(\mathbf{k})_e) + (1-f) \times (OF_{ET}(\mathbf{k})_e) \times (1+\text{Penalty}_{ET}(\mathbf{k})_e)\}$$

$$\begin{cases} f = 1, & SM-\text{only criterion} \\ f = 0, & ET-\text{only criterion} \quad \forall e \\ 0 < f < 1, & SM+ET-\text{joint criterion} \end{cases} \quad (3.13)$$

$$\text{fitness}\,(\mathbf{P}^{*})_{e} = \frac{1}{Z(k)_{e}} \forall e \tag{3.14}$$

The weighting factor (f) is used to assign different weights to SM and ET in the objective function. $\mathbf{k}_r$ is highly subjected to a stochastic field indicating that the fitness ($\mathbf{P}^{*}$) is not always similar for each MC re-sampling event. The noisy fitness is minimized by estimating the so-called sampling fitness (Sfitness($\mathbf{P}^{*}$)) suggested by *Ines and Mohanty* [2009] by averaging the fitness ($\mathbf{P}^{*}$) of each ensemble (e) from the MC re-sampling.

$$\text{Sfitness}(\mathbf{P}^{*}) = \frac{1}{E}\sum_{e=1}^{E}\text{fitness}(P^{*}) \tag{3.15}$$

where E is the ensemble domain for the MC resampling. $\text{Penalty}_{SM}(\mathbf{k})_e$ and $\text{Penalty}_{ET}(\mathbf{k})_e$ in the Eq. (3.13) are determined by Constrained $SM(\mathbf{k})_{te}$ of Eq. (3.9) and Constrained $ET(\mathbf{k})_{te}$ of Eq. (3.10),

$$\text{Penal}_{SM}(\mathbf{k})_{e,j} = \sum_{t=1}^{T}\xi_{t,j}\,(\text{ObsSM}_t - \text{PCI}_{SMj}(\text{SimSM}(\mathbf{k}_r))_{t,e})^2 \,\forall e \tag{3.16}$$

$$\text{Penal}_{ET}(\mathbf{k})_{e,j} = \sum_{t=1}^{T}\xi_{t,j}\,(\text{ObsET}_t - \text{PCI}_{ETj}(\text{SimET}(\mathbf{k}_r))_{t,e})^2 \,\forall e \tag{3.17}$$

where $j = \begin{cases} 1, & +\text{PCI}_{SM}, +\text{PCI}_{ET} \\ 2, & -\text{PCI}_{SM}, +\text{PCI}_{ET} \end{cases}$, the penalty coefficient $\xi_{t,j} = \begin{cases} 10, \text{if FALSE}, \\ 0, \text{otherwise} \end{cases} \forall j, \forall t$ (3.18)

$$\text{Penalty}_{SM}(\mathbf{k})_e = \sum_{j=1}^{2}\text{Penal}_{SM}\,(\mathbf{k})_{e,j}\,\forall e \tag{3.19}$$

$$\text{Penalty}_{ET}(\mathbf{k})_e = \sum_{j=1}^{2}\text{Penal}_{ET}\,(\mathbf{k})_{e,j}\,\forall e \tag{3.20}$$

3.3.5 Numerical Experiments

We conducted this study under the assumption that the effective soil hydraulic parameters in the unsaturated zone can be quantified using the pixel-based remotely sensed soil moisture (e.g., Polarimetric Scanning Radiometer: PSR, Electronically Scanned Thinned Array Radiometer: ESTAR, Advanced Microwave Scanning Radiometer - Earth Observing System: AMSR-E, etc.) and ET (e.g., MODIS, Landsat). To evaluate the parameter estimation algorithm, two synthetic and field validation experiments were conducted: 1) Case 1: homogeneous soil column with free drainage, 2) Case 2: homogeneous soil column with a ground water table depth (GW -100 cm from the soil surface), and 3) Case 3: field validation experiments under various hydroclimatic conditions in Iowa, Illinois, and Texas. For Case 1 and 2, the near-surface soil moisture (0-1cm) and ET time series (as target) values were generated by the SWAP model using the available soil hydraulic parameters from the UNSODA database [*Leij et al.*, 1999]. The numerical experiments have no complexities unlike the RS pixel or *in-situ* soil moisture, which have uncertainties due to various vegetation covers, soil textures, land management practices, climatic conditions, etc. Thus, the numerical experiments are suitable for evaluating the efficiency of the coupled NMCGA and SEBAL algorithm. In Case 3, several field sites at Brown (BRW) in Illinois, Walnut Creek (WC 11 to 14) in Iowa, and Lubbock in Texas were selected for validation studies as shown in Fig. 3.3. The homogeneous soil columns with free drainage condition representing arid/semi-arid regions and ground water table depth of -200, -150, and -100 cm (as shown in Fig. 3.4) representing humid and semi-humid regions were used in this study.

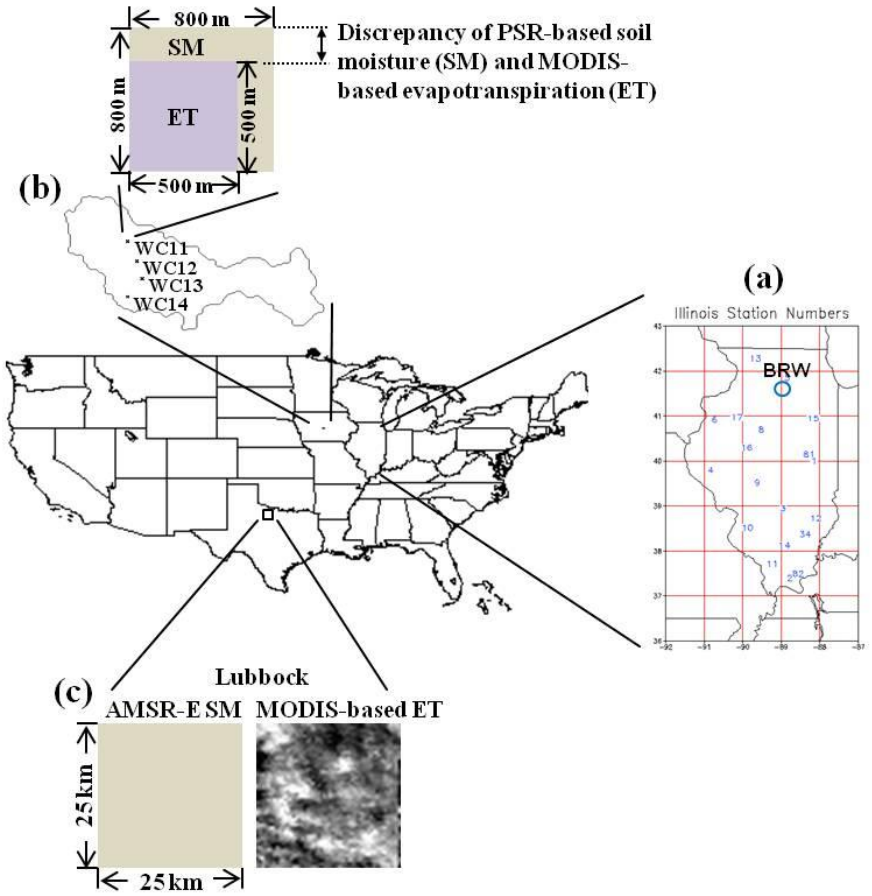

Figure 3.3: Study area; (a) the Brown (BRW) site in Illinois, (b) the Walnut Creek (WC 11-14) sites in Iowa, (c) the Lubbock site in Texas

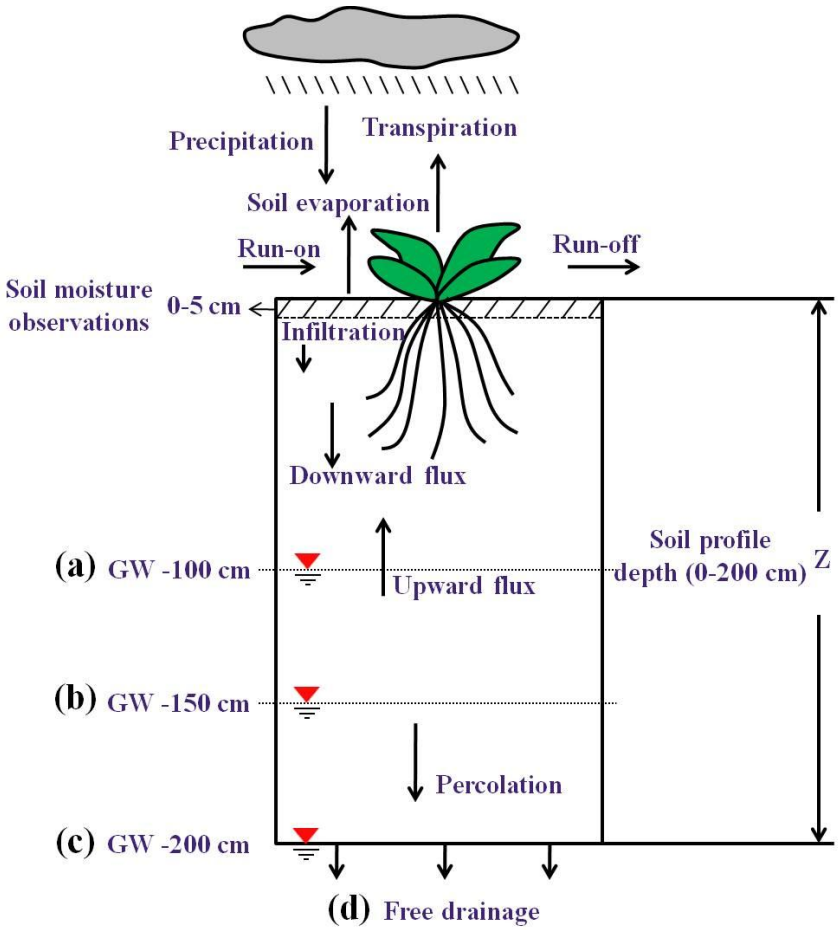

Figure 3.4: Homogeneous soil column for numerical experiments under the free drainage and shallow ground water (GW) tables; (a) GW -100 cm, (b) GW -150 cm, (c) GW -200 cm, (d) free drainage

3.3.5.1 Case 1: Homogeneous Soil Column with Free Drainage

We generated the soil moisture and ET using three different soil textures including sandy loam, silt loam, and clay loam soils with various climate conditions and maize cover. The SWAP model simulations used a standard crop growing season (May 1 - October 31, 2005) in Lubbock, Texas, under the rain-fed condition. For the Lubbock site, weather data are available at Texas ET network (http://texaset.tamu.edu/index.php).

For these (Case 1) numerical studies, it is assumed that the bottom boundary of the soil column is well drained. To evaluate the impact of ET component for quantifying the effective soil hydraulic parameters, three different scenarios were used: 1) the evapotranspiration (ET only: f=0) criterion, 2) the soil moisture (SM only: f=1) criterion, and 3) the soil moisture + evapotranspiration (SM+ET jointly: 0<f<1) criterion. We tested various weighing factors ("f" ranging from 0.1 to 0.9) assigning different weights to SM and ET and selected the weighting factor of 0.8, which has a better match (both simulated SM and ET) with observations.

3.3.5.2 Case 2: Homogeneous Soil Column with a Ground Water Table Depth of -100 cm

According to the findings of *Ines and Mohanty* [2008a,b, 2009], the estimation of effective soil hydraulic properties was affected by the presence of ground water table. The ET component contributing to the water balance in the unsaturated zone was examined in estimating the soil parameters for sandy loam, silt loam, and clay loam soils with the presence of ground water table at -100 cm.

3.3.5.3 Case 3: Field Validation Experiments

Field testing of our proposed inverse modeling based data assimilation algorithm was carried out for evaluating the effective soil hydraulic parameter estimation at the point-, airborne-, and satellite-scales. The Brown site in Illinois (April 1 – October 31, 2002), Walnut Creek (WC) field 11 to 14 in Iowa (May 1 – October 31, 2002), and Lubbock, Texas (March 1 – July 31, 2002) sites were selected for field validation. Vegetation cover of the field sites consisted of corn (WC 11 and WC 12), soybean (WC

13 and WC 14), and grass (Brown and Lubbock), respectively. *In-situ* soil moisture data (for 13 days) for the Brown site at the Global Moisture Soil Moisture Data Bank (GMSDB: http://climate.envsci.rutgers.edu/soil_moisture/) was measured by the neutron probe technique. Airborne PSR (800 m × 800 m footprints) soil moisture datasets (for 10 days) [*Bindlish et al.*, 2006] during the Soil Moisture Experiment 2002 (SMEX02) were used for the WC sites (http://nsidc.org/). For the larger scale, we used the AMSR-E (~25 km X 25 km footprints, [*Njoku*, 2008]) soil moisture product (soil depth of 0-1 cm) for 7 days at the Lubbock site.

The observed ET (4-6 days during the simulation period because of the limited available MODIS datasets due to the weather conditions and scanning intervals) for the Brown, WC 11-14, and Lubbock sites were estimated by the SEBAL model. Daily weather data (e.g., precipitation, wind speed, maximum and minimum temperature, relative humidity, and solar radiation) were collected from the Soil Climate Analysis Network (SCAN: http://www.wcc.nrcs.usda.gov/scan/) in Iowa and the Illinois Climate Network (ICN: http://www.isws.illinois.edu/warm/datatype.asp) in Illinois. For the Lubbock site, we used the Tropical Rainfall Measuring Mission (TRMM: http://trmm.gsfc.nasa.gov/)-based precipitation which has the same resolution matching with the AMSR-E product.

The limitation of the field validation study is the resolution discrepancy between PSR-based soil moisture (800 m X 800 m) and MODIS-based ET (500 m X 500 m) products. Also, the initial and bottom boundary conditions were unknown at the field sites. Thus, we tested our approach with various combinations of initial and bottom

boundary conditions such as the ground water table depths of -200 cm, -150 cm, and -100 cm from the soil surface and selected the best conditions (GW -150 cm for the WC 11 to 14 sites and GW -200 cm for the Brown and Lubbock sites), which have the highest fitness for the individual sites. We assumed that the initial conditions (h(z,t=0) = -150 cm for WC 11 to 14 sites and h(z,t=0) = -200 cm for Brown and Lubbock sites) were in equilibrium with the ground water tables.

Pearson's correlation (R^2) and root mean square error (RMSE) of the observed and simulated soil moisture and evapotranspiration (ET uses the same equations in Eq. (3.21) and (3.22)) were used to evaluate our numerical and field validation experiments of the inverse modeling based data assimilation (using integrated NMCGA and SEBAL) algorithm:

$$R^2 = \frac{\sum_{t=1}^{n}(\theta_{sim,t} - \bar{\theta}_{sim})(\theta_{obs,t} - \bar{\theta}_{obs})}{\sqrt{\sum_{t=1}^{n}(\theta_{sim,t} - \bar{\theta}_{sim})^2 \sum_{t=1}^{n}(\theta_{obs,t} - \bar{\theta}_{obs})^2}} \tag{3.21}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(\theta_{obs,t} - \theta_{sim,t})^2}{n}} \tag{3.22}$$

where $\theta_{sim,t}$ is the simulated soil moisture with the time index (t), $\bar{\theta}_{sim}$ is the average soil moisture of $\theta_{sim,t}$, $\theta_{obs,t}$ is the observed soil moisture with the time index (t), $\bar{\theta}_{obs}$ is the average soil moisture of $\theta_{obs,t}$, and *t* is the time index, respectively.

3.4 Results and Discussion

3.4.1 Numerical Case Studies

3.4.1.1 Case 1: Homogeneous Soil Column with Free Drainage

Table 3.2 shows the summary of solutions (P^*: $\mu_{(\alpha)}$, $\mu_{(n)}$, $\mu_{(\theta res)}$, $\mu_{(\theta sat)}$, $\mu_{(Ksat)}$) for the numerical experiments derived by the proposed data assimilation algorithm under the ET-only, SM-only, and SM+ET-joint criteria with different soil textures and maize crop cover for Case 1 scenario. Estimated soil hydraulic parameters are shown in terms of their arithmetic means (3 ensembles × 30 realizations) and ±95PCI.

Table 3.2: Solutions for sandy loam, silt loam, and clay loam soils under the ET-only, SM-only, and SM+ET-joint criteria for Case 1

Soil Types	Parameters	Target values[a]	ET criteria		SM criteria		SM+ET criteria	
			Average	±95PCI	Average	±95PCI	Average	±95PCI
Sandy loam	α	0.021	0.023	0.011-0.036	0.022	0.014-0.030	0.024	0.012-0.036
	n	1.610	1.589	1.566-1.613	1.586	1.549-1.623	1.579	1.530-1.627
	θ_{res}	0.067	0.077	0.074-0.080	0.067	0.060-0.074	0.065	0.063-0.068
	θ_{sat}	0.370	0.387	0.387-0.387	0.370	0.370-0.370	0.377	0.368-0.387
	K_{sat}	41.600	53.321	50.006-56.637	50.875	45.129-56.620	47.432	41.040-53.824
Silt loam	α	0.012	0.014	0.004-0.024	0.013	0.002-0.024	0.013	0.007-0.019
	n	1.390	1.457	1.233-1.681	1.498	1.322-1.674	1.429	1.190-1.669
	θ_{res}	0.061	0.067	0.059-0.076	0.064	0.060-0.069	0.076	0.060-0.091
	θ_{sat}	0.430	0.481	0.396-0.566	0.457	0.355-0.560	0.452	0.417-0.487
	K_{sat}	30.500	45.970	39.322-52.619	34.505	16.419-52.591	33.567	26.510-40.624
Clay loam	α	0.030	0.028	0.021-0.035	0.030	0.030-0.030	0.032	0.032-0.032
	n	1.370	1.409	1.159-1.659	1.304	1.304-1.304	1.413	1.183-1.643
	θ_{res}	0.129	0.119	0.109-0.129	0.125	0.097-0.153	0.135	0.109-0.160
	θ_{sat}	0.470	0.473	0.370-0.576	0.452	0.354-0.549	0.494	0.413-0.575
	K_{sat}	1.840	4.854	0.883-8.824	2.380	1.493-3.267	4.168	0.968-7.368

Max generation: 30, Number of chromosomes in a generation: 10, Number of ensembles: 3, Number of resample: 30, [a]UNSODA database [*Leij et al*., 1999]

In general, the shape (α) parameter under most conditions (described by the three criteria) for sandy loam soil closely identified with the target value, while the estimated (n) parameters are more variable. The scaling (θ_{res} and θ_{sat}) parameters under the SM-only and SM+ET-joint criteria are better than those of the ET-only criterion suggesting that they have more uncertainties in the later case. The K_{sat} values under all the criteria are considerably overestimated compared to the target values. But it is apparent that the K_{sat} value estimated under the SM+ET-joint criterion is better than those of the ET-only and SM-only criteria. Figure 3.5 shows the $\theta(h)$ and K(h) functions derived by the estimated soil hydraulic parameters (Table 3.2) for sandy loam soil. The $\theta(h)$ functions under all the criteria appeared to be identifiable with the observations, but the K(h) functions have more (mean) bias errors. Uncertainties of the K(h) values under the ET-only and SM-only criteria are considerably higher compared to the target values. The SM+ET-joint criterion also has variations in the K(h) function. However, the estimated values are closer to the true observations compared to those of the SM-only and ET-only criteria as the target value is only adequately defined in a range of ±95PCI under the SM+ET-joint criterion. This may imply that by adding an ET component in the near-surface soil moisture assimilation algorithm improves the estimates of K(h) functions.

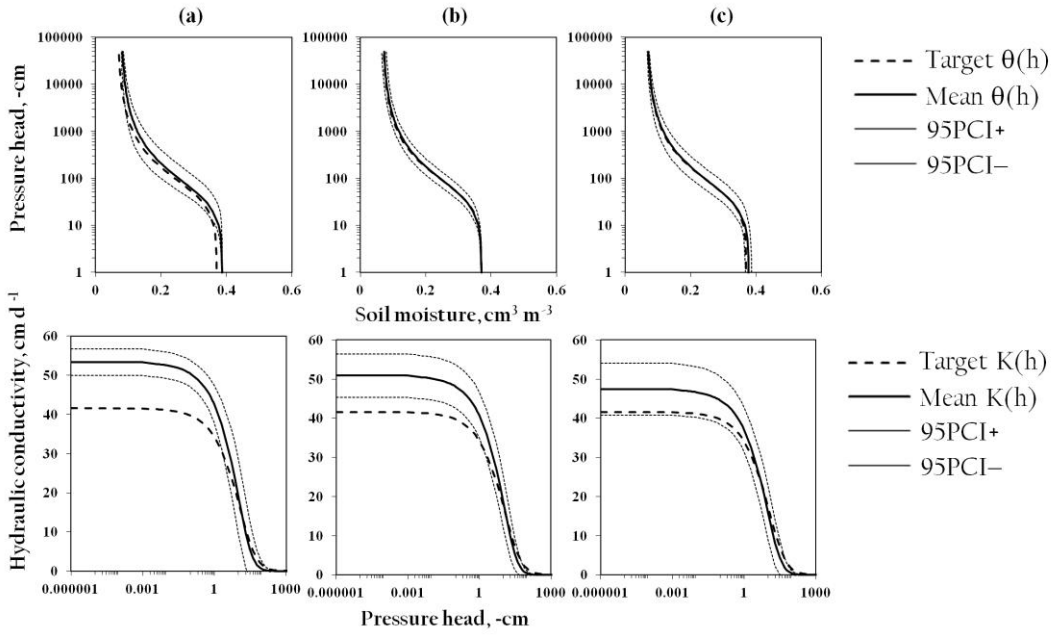

Figure 3.5: θ(h) and K(h) functions from MC simulations for sandy loam soil under rain-fed condition; (a) ET-only criterion, (b) SM-only criterion, (c) SM+ET-joint criterion

In silt loam soil, the α and θ_{res} values under the ET-only and SM-only criteria matched well with the target values. In contrast, for the SM+ET-joint criteria, only α is predictable with high accuracy. The K_{sat} value fared better in performance under the SM+ET-joint criterion compared to the SM-only criterion, while the ET-only criterion showed large uncertainties. The θ(h) functions under the SM-only and SM+ET-joint criteria are found to be superior to those obtained for the ET-only criteria as shown in Figure 3.6. But uncertainty bound (±95PCI) for the θ(h) functions under the SM+ET-joint criterion are smaller than those of the ET-only and SM-only criteria. The most profound impact of ET component is visible more clearly in the K(h) functions, because

variations of the simulated K(h) function under the SM+ET-joint criterion are considerably decreased compared to those of the SM-only and ET-only criteria.

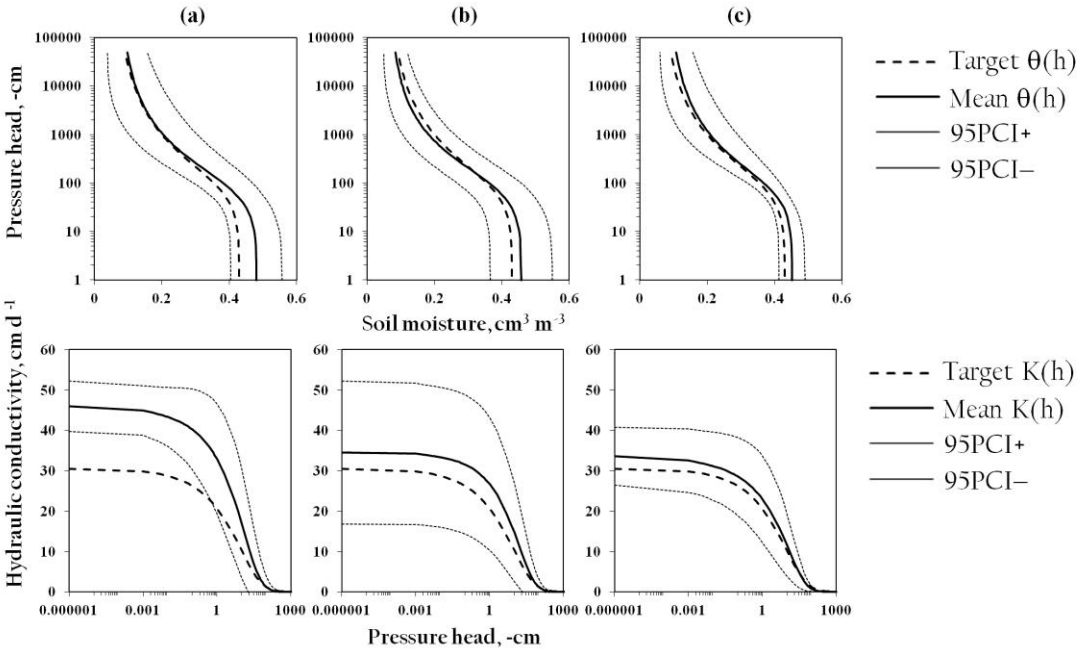

Figure 3.6: θ(h) and K(h) functions from MC simulations for silt loam soil under rain-fed condition; (a) ET-only criterion, (b) SM-only criterion, (c) SM+ET-only criterion

For clay loam soil (see Figure 3.7), the α and θ_{sat} values under the ET-only criterion are easily identifiable with the target values, but the n, θ_{res}, and K_{sat} estimates are relatively less clear (Table 3.2). The α, θ_{res}, and K_{sat} values under the SM-only criterion are matched well. Only the α value under the SM+ET-joint criterion matched well with the observation unlike the solutions for sandy loam and silt loam soils. Figure 3.7 shows the θ(h) and K(h) functions of clay loam soil for all the criteria considered.

The θ(h) and K(h) functions estimated under the SM-only criterion are slightly better than those for the ET-only and SM+ET-joint criteria. Also, uncertainties of ±95PCI under the SM-only criterion are narrow compared to those for the ET-only and SM+ET-joint criteria.

Figure 3.8 shows the comparisons of observed and simulated ET for various criteria with clay loam soil under the synthetic conditions. Note that ET comparisons for sandy loam and silt loam soils are not presented for the sake of brevity. For DOY 129, the extremely high (synthetic/observed) ET value (0.55 mm d^{-1}) was generated in this numerical study due to the inherent weakness of the adopted hydrological model. This may imply that an error in ET influences the uncertainties in the parameter estimation and modeling performance for clay loam soil in the unsaturated zone. Although the simulated K(h) functions of the ET-only and SM+ET-joint criteria have bias (caused by an unusual synthetic/observed ET value), the simulated θ(h) functions still correspond well to the target values. The results of clay loam soil indicate that the hydraulic conductivity function (K(h)) is more sensitive to ET component than the soil water retention function (θ(h)). Overall, most of the θ(h) and K(h) functions under the SM+ET-joint criterion are better identifiable than those for the ET-only and SM-only criteria. Furthermore, the effect of ET component is relatively less sensitive to the soil water retention (θ(h)) than the hydraulic conductivity (K(h)).

Figure 3.9 shows the soil moisture in the deeper soil depth (180-200 cm) for three soil textures. In sandy loam soil, simulated soil moisture under the SM-only (RMSE: 0.001) and SM+ET-joint (RMSE: 0.004) criteria matched well with the target

values. But the soil moisture (RMSE: 0.010) for the ET-only criterion was overestimated compared to the true value. We can confirm similar patterns for silt loam and clay loam soils as well. The soil moisture estimates under the SM+ET-joint criterion for silt loam (RMSE: 0.008) and clay loam (RMSE: 0.002) soils are closer to the target value than those for the ET-only (RMSE: 0.015 for silt loam, RMSE: 0.008 for clay loam) and SM-only (RMSE: 0.011 for silt loam, RMSE: 0.004 for clay loam) criteria, respectively. These results demonstrate that an ET component is attributable to improving the predictive skill of soil moisture flux at the deeper soil depths. These findings are quite significant for quantifying the effective soil hydraulic parameters of the vadose zone with the coupled NMCGA and SEBAL algorithm using only near-surface (0-1 or 0-5cm) soil moisture estimates from airborne/satellite RS platform. Uncertainties may result in identifying the soil hydraulic parameters as the soil depth increases due to textural layering, rooting depth and density, pore-size distribution, etc. Summarily, results of Case 1 indicate that the coupled NMCGA and SEBAL algorithm including an ET component (with SM), that provides more information in the unsaturated zone hydrologic behavior, improves the estimates of effective soil hydraulic parameters.

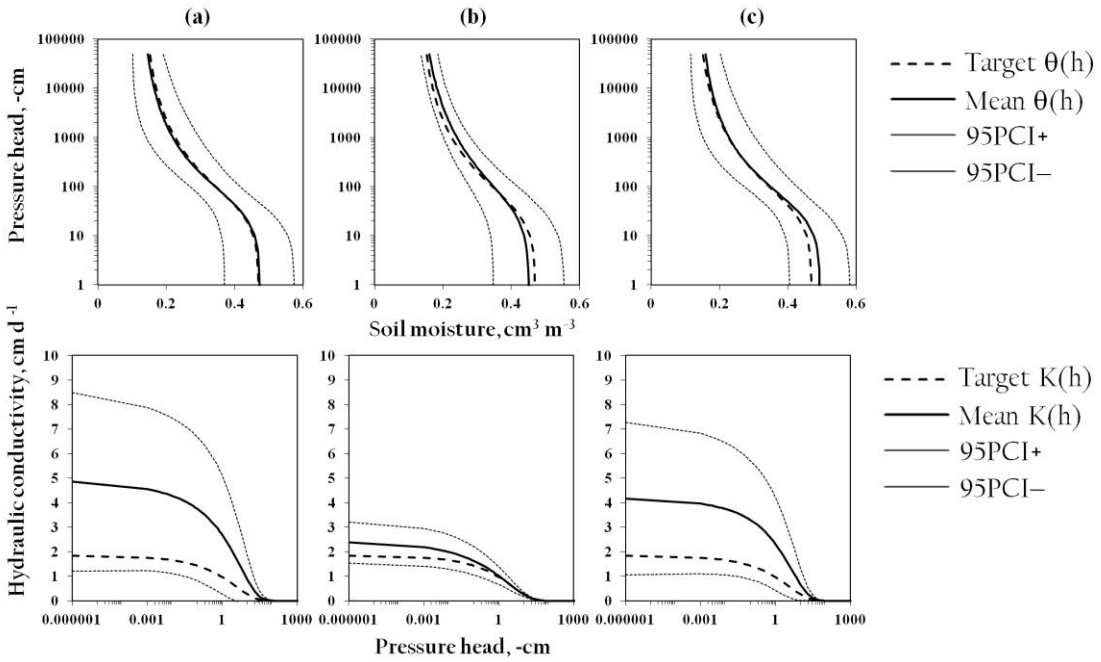

Figure 3.7: θ(h) and K(h) functions from MC simulations for clay loam soil under rain-fed condition; (a) ET-only criterion, (b) SM-only criterion, (c) SM+ET-joint criterion

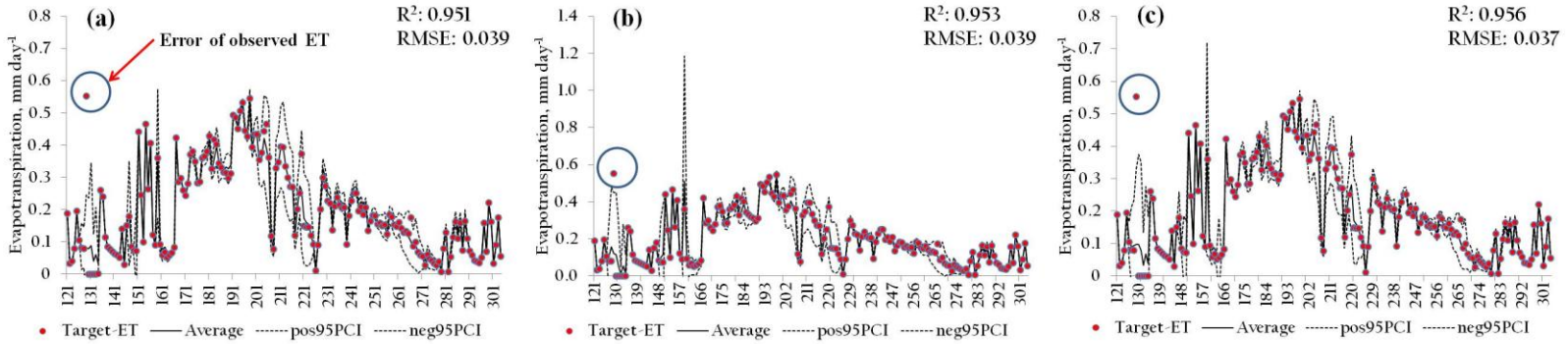

Figure 3.8: Comparisons of evapotranspiration for clay loam soil under rain-fed criterion; (a) ET-only criterion, (b) SM-only criterion, (c) SM+ET-joint criterion.

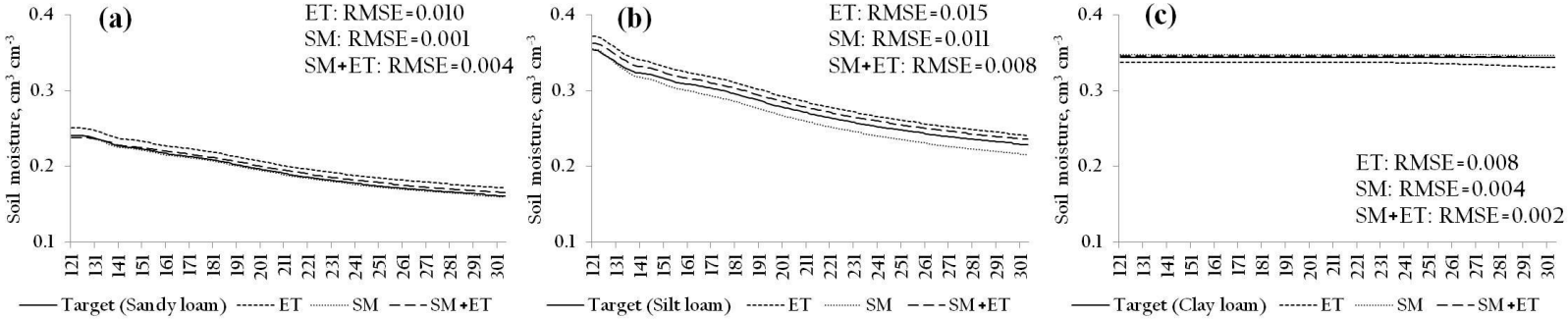

Figure 3.9: Comparisons of the deep (180-200 cm) soil moisture dynamics for sandy loam, silt loam, and clay loam soil; (a) ET-only criterion, (b) SM-only criterion, (c) SM+ET-joint criterion

3.4.1.2 Case 2: Homogeneous Soil Column with Ground Water Table

We simulated Case 2 scenario to evaluate the impact of an ET component in the presence of a ground water table (-100 cm from the soil surface) in the homogeneous soil column under the SM-only and SM+ET-joint criteria (the ET criteria is excluded). Table 3.3 shows the derived effective soil hydraulic parameters with the ground water table depth of -100 cm.

In sandy loam soil, the soil hydraulic parameters under the SM+ET-joint criterion are a better match than those for the SM-only criterion with the true values except of θ_{res} and θ_{sat}. Although the standard deviation (SD) of K_{sat} under the SM+ET-joint criterion is slightly higher than for the SM-only criterion, the average value of K_{sat} is estimated more successfully. The results for silt loam soil are also similar compared to the findings of sandy loam soil. The α, n, θ_{res}, and θ_{sat} values under the SM+ET-joint criterion are identifiable well with the target values, while under the SM-only criterion only the α value matched well. The K_{sat} value under the SM-only criterion has large uncertainties compared to that of the SM+ET-joint criterion. Usually, silt loam soil with shallow water tables is less sensitive than those for sandy loam and clay loam soils. In this study, when an ET component is included in quantifying the effective soil hydraulic properties, the solutions for the silt loam soil in the presence of a ground water table are much improved. In clay loam soil, the parameters under the SM+ET-joint criterion are perfectly identified with the target value, while the n value has small variations. However, the α and θ_{res} values under the SM-only criterion are only well matched. In general, SM+ET-joint criterion performed better for predicting the soil hydraulic

parameters and their uncertainties better than SM-only criterion under the ground water table condition.

Mostly, the solutions with the shallow water table depth of -100 cm are not as well identified as those under the free drainage condition indicating that the parameter estimations have more uncertainties in the soil profile dominated by upward flows from the ground water table rather than the free drainage condition. Based on the results of Case 2, we suggest that this approach can reflect the hydrological condition in the unsaturated zone affected by both the top (atmospheric) and bottom boundary conditions.

Table 3.3: Solutions with a shallow water table depth of -100 cm for sandy loam, silt loam, and clay loam soils under the SM and SM+ET criteria for Case 2

Soil Types	Parameters	Target values	GW -100			
			SM		SM+ET	
			Average	SD	Average	SD
Sandy loam	α	0.021	0.029	0.000	0.016	0.005
	n	1.610	1.399	0.118	1.510	0.055
	θ_{res}	0.067	0.074	0.009	0.080	0.013
	θ_{sat}	0.370	0.370	0.000	0.471	0.037
	K_{sat}	41.600	50.157	2.274	38.422	3.042
Silt loam	α	0.012	0.013	0.006	0.013	0.002
	n	1.390	1.591	0.012	1.397	0.115
	θ_{res}	0.061	0.082	0.011	0.067	0.003
	θ_{sat}	0.430	0.456	0.048	0.445	0.055
	K_{sat}	30.500	45.642	5.276	36.049	5.050
Clay loam	α	0.030	0.033	0.000	0.031	0.002
	n	1.370	1.386	0.145	1.401	0.135
	θ_{res}	0.129	0.127	0.015	0.130	0.004
	θ_{sat}	0.470	0.486	0.014	0.472	0.054
	K_{sat}	1.840	4.522	0.880	2.050	0.055

3.4.1.3 Case 3: Field Validation Experiments

In Case 3 we tested the applicability of the coupled NMCGA and SEBAL algorithm under the SM-only and SM+ET-joint criteria at various spatial scales in the field conditions. Table 3.4 shows the statistics of simulated soil moisture and ET with respect to the measurements. Mostly, the statistics (R^2 and RMSE) of soil moisture and ET estimates under the SM+ET-joint criterion show good performance and generally better than those under the SM-only criterion. Figs. 3.10 and 3.11 show the observed and simulated results (SM and ET) at different spatial scales (point-scale for the Brown, Illinois site, airborne-scale for the WC 11, Iowa site, and satellite-scale for the Lubbock, Texas site) under the SM-only and SM+ET-joint criteria. As shown in the synthetic Case 1 and 2 (in Table 3.2 and 3.3), not only the simulated SM and ET estimates but also their uncertainty ranges under the SM+ET-joint criterion improved compared to those of the SM-only criterion in these field situations at different spatial scales.

Table 3.4: Correlation (R^2) and RMSE of the pixel-based (*in-situ*)/simulated soil moisture and ET under the SM and SM+ET criteria at the Brown, WC (11-14) and Lubbock sites for Case 3

	SM				SM+ET			
	Soil moisture		ET		Soil moisture		ET	
Sites	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
[a]Brown	0.927	0.085	0.874	1.663	0.936	0.101	0.874	0.990
[b]WC 11	0.777	0.050	0.955	0.554	0.781	0.047	0.952	0.500
[b]WC 12	0.799	0.042	0.948	0.605	0.792	0.049	0.957	0.432
[b]WC 13	0.777	0.054	0.954	1.022	0.782	0.054	0.949	0.978
[b]WC 14	0.784	0.045	0.943	1.208	0.785	0.044	0.944	1.200
[c]Lubbock	0.537	0.053	0.600	1.316	0.569	0.060	0.665	1.296

[a]*In-situ* soil moisture and pixel-based ET datasets are used for the Brown site in Illinois
[b]Pixel-based (PSR) soil moisture and ET datasets are used for the Walnut Creek (WC 11 to 14) sites in Iowa
[c] Pixel-based (AMSR-E) soil moisture and ET datasets are used for the Lubbock site in Texas

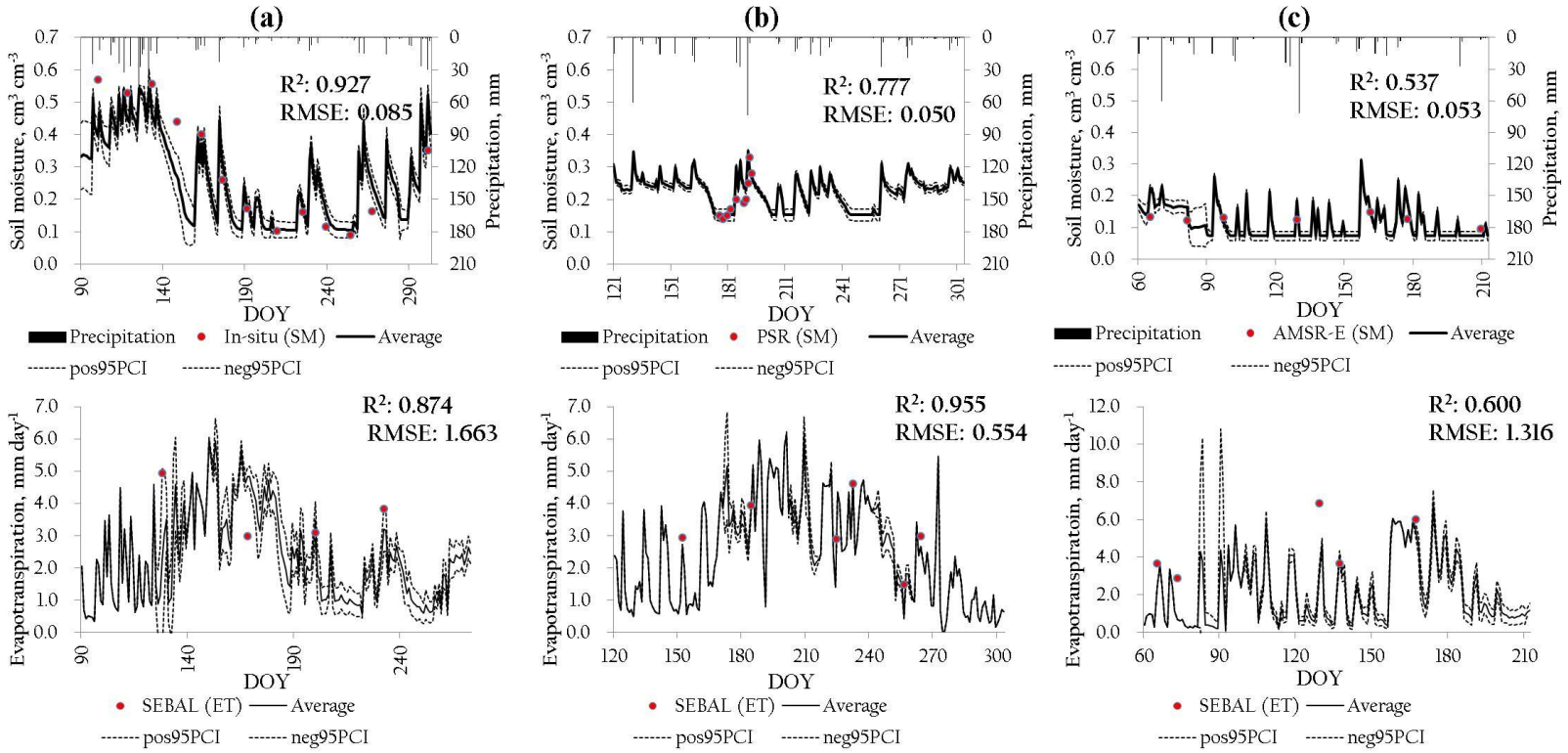

Figure 3.10: Comparisons of the observed and simulated soil moisture (SM) and evapotranspiration (ET) at the different scales using the coupled NMCGA and SEBAL algorithm under the SM criterion; (a) point-scale (*in-situ*) SM and pixel-based ET, (b) airborne-scale (PSR) pixel-based SM and ET, (c) satellite-scale (AMSR-E) pixel-based SM and ET

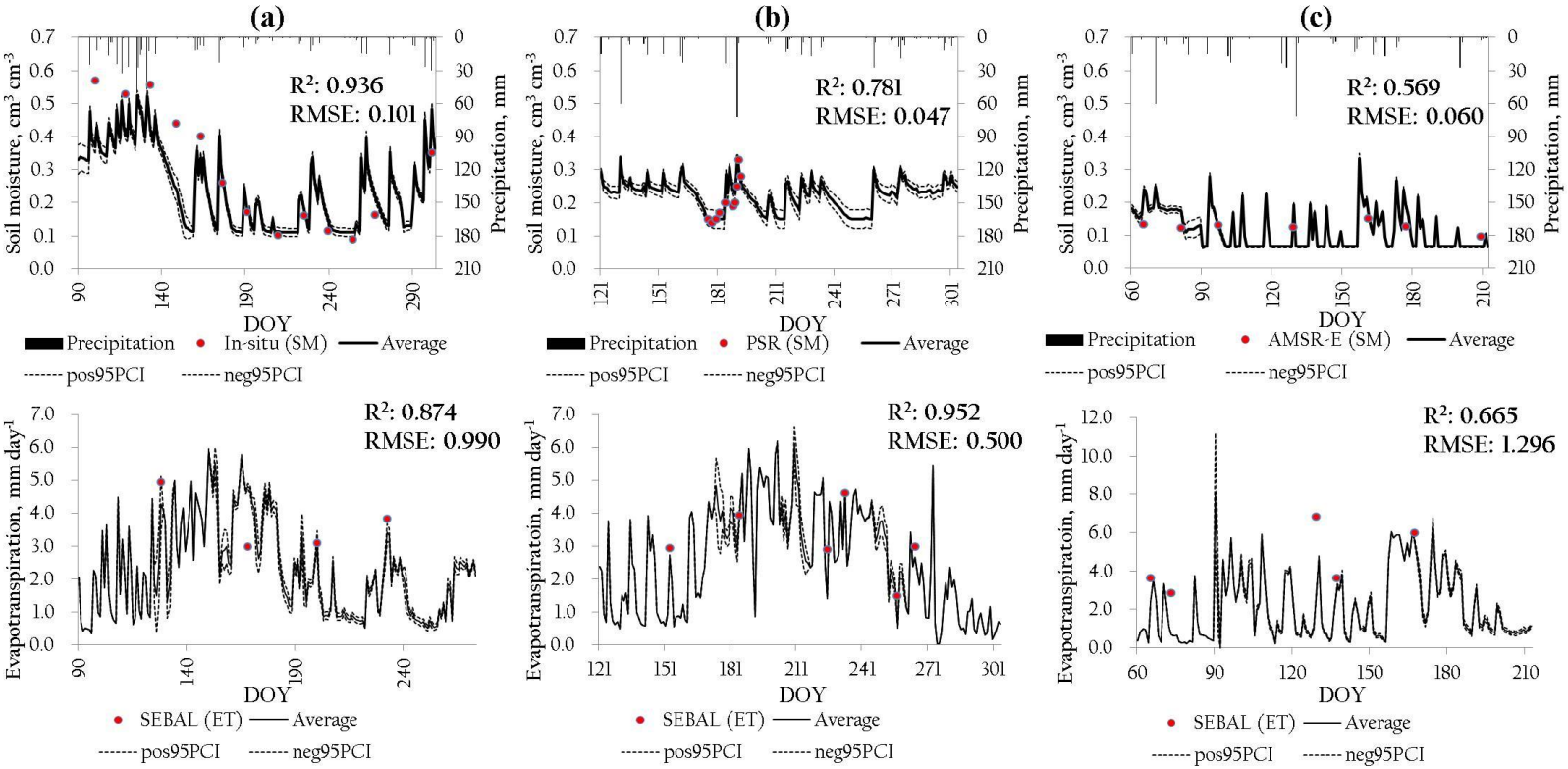

Figure 3.11: Comparisons of the observed and simulated soil moisture (SM) and evapotranspiration (ET) at the different scales using the coupled NMCGA and SEBAL algorithm under the SM+ET-joint criterion; (a) point-scale (*in-situ*) SM and pixel-based ET, (b) airborne-scale (PSR) pixel-based SM and ET, (c) satellite-scale (AMSR-E) pixel-based SM and ET

At the point-scale, simulated soil moisture dynamics have large uncertainties due to various factors (e.g., climatic data from sparse weather stations, limited information on rooting depth and density, lack of plant growth functions, inherent weakness of hydrologic model structure) with respect to measurements (including measurement errors) during the early simulation period (DOY: 90-150). Note that soil moisture lacked response to the rainfall event on DOY 175-177. At the footprint-scale, model estimated soil moisture matched relatively well with the PSR-based soil moisture measurements. Compared to point-scale, remote sensing footprint-scale result has smaller uncertainties. Footprint-scale discrepancy could be attributed to the data qualities of the specific pixels, weather conditions and variations within the pixel, and discrepancy between PSR- and MODIS-scales for input data. Especially, the observed (PSR) and simulated soil moisture patterns have shown good match with the rainfall pattern indicating that the pixel-based soil moisture at the field scale reflects well the spatial correspondence rather than the point-scale. The observed/simulated soil moisture dynamics at the satellite-scale are relatively lower than those of the point- and airborne-scales. Although various uncertainties from the heterogeneities of areal soil textures and vegetation covers across the land surface are included in a remote sensing pixel, the results estimated by this approach matched well with the measurements.

Above results show the impact of different spatial scales in parameter estimation. As the scale increases from the point- to satellite-scale, soil moisture quantities trend to be decreased compared to rainfall amounts, because of the bias of different scales. However, the uncertainty (RMSE) at the point-scale was higher than at the other scales

in the limited conditions undertaken for this study. It demonstrates that the airborne-scale provides us relatively reasonable statistics (both R^2 and RMSE) than those of the point- and satellite-scales. These results beg a question, because the results with the *in-situ* measurements at the point-scale have more variations in the modeling performance than those at the satellite-scale. *In-situ* datasets usually have a high accuracy indicating that the large errors of *in-situ* soil moisture in this study might be an extreme (unusual) case. However, it shows one of the potential uncertainties, which can be incurred at fields. Thus, pixel-based datasets can provide more stable soil moisture information with predictable uncertainties (e.g., weather conditions, quality of a pixel, etc).

The ET values have similar patterns as the soil moisture. The estimated ET values matched the pixel-based ET, but the observation in DOY 169 is considerably lower than the simulated ET value, which means that the MODIS images as input data for the SEBAL algorithm may have noise. The ET estimates at the field-scale have a good match with the MODIS-based ET. But the large-scale ET results have more variations, since an AMSR-E product consists of areal heterogeneities of soil textures and vegetation covers whereas the 1-D physical model (SWAP) is limited to model a homogeneous land surface. Based on these results, it is evident that the (pixel-based) ET component improves the parameter estimations and contributes towards the reduction of uncertainty ranges in simulating soil moisture and ET under the inverse modeling.

These results further support the robustness of the coupled NMCGA and SEBAL algorithm for quantifying effective soil hydraulic parameters at different scales in a real world scenario.

3.5 Conclusions

This study was conducted to evaluate the applicability of a new coupled NMCGA and SEBAL based data assimilation algorithm for quantifying effective soil hydraulic properties using the RS pixel-based near-surface soil moisture and ET products at the point- to satellite-scales. Numerical experiments for the synthetic conditions and field validations are undertaken with different soil textures, climate scenarios, presence of ground water table, and vegetation covers under the ET-only, SM-only, and SM+ET-joint criteria. Generally, the $\theta(h)$ and $K(h)$ functions (derived by the searched soil hydraulic parameters) under the SM+ET-joint criterion are identified better than those of the ET-only and SM-only criteria for Case 1, although the $K(h)$ functions still have small variations. As the simulated soil moisture dynamics under the SM+ET-joint criterion in the deeper soil depth (180-200 cm) for three soil textures (sandy loam, silt loam, and clay loam soils) have a better match with the target values compared to those for the ET-only and SM-only criteria, it confirms that this approach improves not only the parameter estimations, but also the soil moisture flux in the deep soil depth (180-200 cm). In Case 2, the soil hydraulic properties in the presence of a ground water table (-100 cm) under the SM+ET-joint criterion are superior to those for the SM-only criterion. It is evident that this inverse modeling based data assimilation approach including an ET component contributes towards the reduction of uncertainties generated by the upward flow of ground water table.

In a real world situation, the results of soil moisture and ET estimates under the SM+ET-joint criterion still matched the measurements from point- to satellite-scales

more than those of the SM-only criterion as shown in the synthetic experiments, although the correlations (WC 12 site for the SM, WC 11 and 13 sites for the ET) and RMSE (Brown and Lubbock sites for the SM) under the SM-only criterion were slightly better. At the point-scale, the estimated SM has no matching with the measurements during the initial simulation period indicating that the discrepancy of spatial scales (measuring location and weather station) and measuring errors may cause large uncertainties in estimating the parameters. The satellite-scale results were influenced by the areal heterogeneity of land surface (e.g., soil texture, vegetation covers, etc.), but the SM estimates were comparable with the AMRE-E products. In the airborne-scale, the estimated soil moisture was more reasonable for the spatial and temporal scales, although the discrepancy between PSR- and MODIS-based resolutions was included. The ET estimates at the satellite-scale have more variations than those of the point- and airborne-scales, because of the limitation of model structure, noise of pixels, etc.

Thus, these results of synthetic and field validation experiments demonstrate that although the RS pixel-based product has variations in the inverse modeling, the coupled NMCGA and SEBAL algorithm can be useful for estimating the hydraulic parameters at the multiple scales across the land surface and contributes towards the reduction of uncertainties.

CHAPTER IV

DEVELOPMENT OF A DETERMINISTIC DOWNSCALING ALGORITHM FOR REMOTE SENSING SOIL MOISTURE FOOTPRINT USING SOIL AND VEGETATION CLASSIFICATIONS

4.1 Synopsis

Soil moisture at the local scale is required to account for small-scale spatial heterogeneity of land surface, because many hydrological processes manifest at scales ranging from cm to km. Although remote sensing (RS) platforms provide large-scale soil moisture dynamics, scale discrepancy between observation scale (e.g., ~ 40 km) and modeling scale (e.g., ~ 1 km) leads to uncertainties in the performance of land surface hydrologic models. To overcome this drawback, we developed a new deterministic downscaling algorithm (DDA) for estimating fine-scale soil moisture with large pixel-based remotely sensed (RS) soil moisture (SM) and evapotranspiration (ET) products using a genetic algorithm. This approach was evaluated under various synthetic and field experiment (Little Washita (LW) 13 and 21, Oklahoma) conditions including homogeneous and heterogeneous land surface conditions composed of different soil textures and vegetation. Our algorithm is based on determining effective soil hydraulic properties for different sub-pixels within a RS pixel and estimating the long-term soil moisture dynamics of individual sub-pixels using the hydrological model with the extracted soil hydraulic parameters.

The soil moisture dynamics of sub-pixels from synthetic experiments matched well with the observations under heterogeneous land surface condition, although

uncertainties (MBE: -0.049~0.049) exist. Field experiments typically have more variations due to weather conditions, measurement errors, unknown bottom boundary conditions, and scale discrepancy between remote sensing pixel and model grid resolution. However, the soil moisture estimates of individual sub-pixels (from ESTAR footprints) downscaled by this approach matched well (R^2: 0.724~0.914, MBE: -0.203~-0.169 for the LW 13; R^2: 0.343~0.865, MBE: -0.165~-0.122 for the LW 21) with the *in-situ* soil moisture measurements. The good correspondence of observed θ(h) functions (from the soil core samples) and GA searched soil parameters at the LW 13 and 21 sites demonstrated the robustness of this algorithm. Although this algorithm is tested under limited conditions at field-scale, this approach improves the availability of remotely sensed soil moisture product at finer-resolution for various land surface and hydrological model application.

4.2 Introduction

Land surface soil moisture is a pivotal factor for hydrology, agronomy, and meteorology. In general, soil moisture data is limited to a few spatio-temporal scales. Point-scale soil moisture datasets are time consuming, expensive, and do not provide a uniform representation at larger scales. Remote sensing (RS) techniques can provide an attractive alternative to direct measurement. *Ottlé and Vidal-Madjar* [1994] derived land surface soil moisture using thermal infrared remote sensing. Directly active [*Ulaby et al*., 1996] and passive microwave [*Njoku and Entekkabi*, 1996] remote sensing approach were developed to estimate surface soil moisture dynamics. However, the use of RS pixel-based data is limited due to the scale discrepancy between observed RS resolution

and required modeling resolution [*Engman*, 1991; *Entekhabi et al*., 1999]. In this regard, downscaling schemes are necessary to improve the availability of sub-pixel soil moisture products from RS footprints/pixels for agriculture and water resources management at the field scale.

A few studies have explored downscaling or disaggregation methods for extracting subgrid soil moisture estimates within a RS pixel. *Crow et al.,* [2000] downscaled spaceborne soil moisture products to obtain surface soil dielectric values approximating to volumetric soil moisture content using a soil dielectric inversion model. *Merlin et al.,* [2005] developed a downscaling method using fine-scale optical data during the Soil Moisture and Ocean Salinity (SMOS) mission [*Kerr et al.,* 2001] to improve the availability of SMOS near-surface soil moisture at the sub-grid scale. Also, an interpolation approach of passive microwave data based on fine-scale active microwave data was developed by *Kim and Barros* [2002] and further refined by *Das et al.* [*2008b*] particularly for the Soil moisture Active and Passive (SMAP) mission. More recently, *Ines et al.,* [2012] developed a stochastic disaggregation method for soil moisture using a simulation-assimilation scheme. This approach extracts soil type identification (representing soil hydraulic properties) and sub-area fractions of corresponding soil-vegetation combinations within a RS soil moisture product. However, the stochastic disaggregation method estimates only the soil characteristics (soil ID values) and sub-area fractions (%) by the soil-vegetation combinations within a pixel in a probabilistic sense without their specific locations practically recognized. In other words, few studies have addressed the issue of downscaling remotely sensed soil

moisture with footprints ranging from several hundred meters to several kilometers (e.g., airborne Electronically Scanned Thinned Array Radiometer, ESTAR, Polarimetric Scanning Radiometer, PSR, space-borne SMOS, and Soil Moisture Active and Passive, SMAP sensors), and resolving them at fine scale (e.g., for individual fields with homogeneous soil and vegetation).

The main goal of this study is to develop and test a new downscaling algorithm with pixel-based soil moisture (SM) and evapotranspiration (ET). The primary objectives are two-fold: 1) to develop a deterministic downscaling algorithm (DDA) using a genetic algorithm (GA) scheme for producing sub-pixel level soil moisture products from large spatial scale data at various resolutions and 2) to assess the robustness of this approach for remotely sensed data under real and synthetic experiments in the time domain.

4.3 Materials and Methods

4.3.1 Deterministic Downscaling Algorithm (DDA)

Heterogeneity in land surface comprises different soil textures and vegetation covers (assuming the case of flat topography). It is assumed that various soil-vegetation combinations have their unique characteristics of soil moisture and evapotranspiration dynamics as illustrated in Fig. 4.1(a).

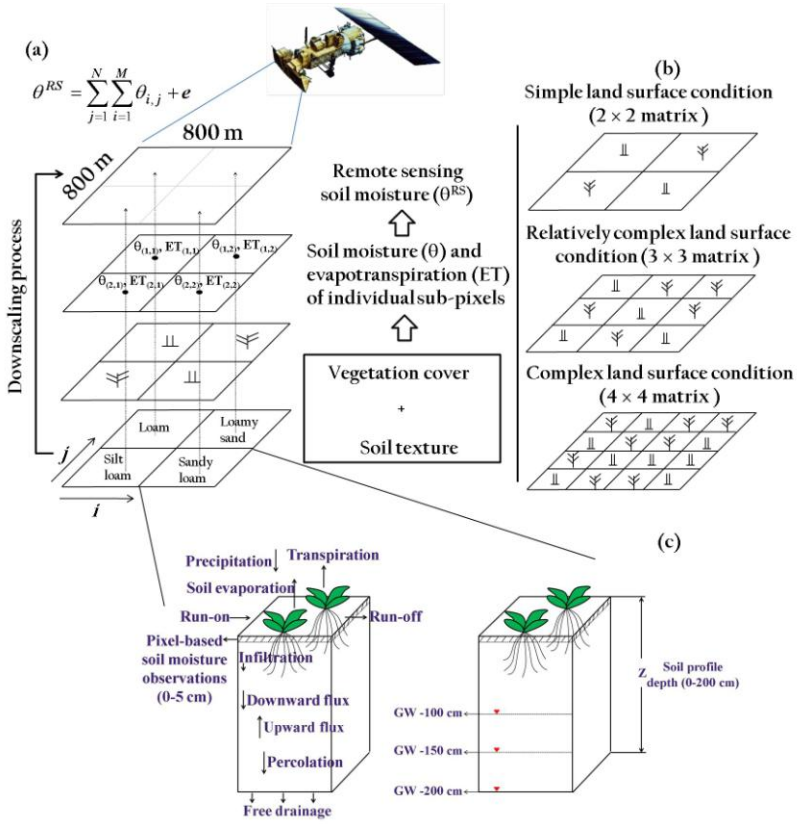

Figure 4.1: (a) Description of the deterministic downscaling algorithm, (b) simple (2 × 2 matrix), relatively complex (3×3 matrix), and complex (4×4 matrix) land surface conditions, (c) homogeneous soil columns with the free drainage condition and various shallow ground water table depths (GW -200, -150, -100 cm)

As we estimate the distributed ET values of various soil-vegetation combinations within a pixel, the SM estimates and their locations corresponding to the ET values can be also obtained. High resolution RS images provide finer-scale ET products across land surface. Thus, in this study we developed a deterministic downscaling algorithm (DDA) for extracting fine-scale soil moisture (for sub-pixels) within a RS footprint using pixel-based SM and ET. Basically, the spectral mixtures within a RS pixel-based product have

linear relationships. The response of each sub-pixel in any spectral wavelength can be considered as a linear combination of the responses of each component which is assumed to be in the mixture. Thus, each image (sub-pixel) contains land surface information with respect to the fraction and spectral response of each component within the ground resolution unit. Hence, individual sub-pixel spectral images ($a_{i=1,\ldots M, j=1,\ldots,N}$), which have the soil components ($\mathbf{P}=\{ s_{i,j} \}$) in Eq. (4.1), can be practically designed as a linear mixture [*Ferreira et al*., 2007; *Ines et al.*, 2012] (Eq. (4.2-4)) as below,

$$\mathbf{P} = \{ s_{i,j} = s_{1,1}, \ldots, s_{M,N} \} \tag{4.1}$$

$$\theta^{sum,t}(\mathbf{P}) = \frac{1}{N} \frac{1}{M} \{ \sum_{j=1}^{N} \sum_{i=1}^{M} \theta^{sub,i,j,t} \} \quad \forall t \tag{4.2}$$

$$ET^{sum,t}(\mathbf{P}) = \frac{1}{N} \frac{1}{M} \{ \sum_{j=1}^{N} \sum_{i=1}^{M} \mathrm{ET}^{\mathrm{sub,i,j,t}} \} \quad \forall t \tag{4.3}$$

$$ET^{RSsum,t}(\mathbf{P}) = \frac{1}{N} \frac{1}{M} \{ \sum_{j=1}^{N} \sum_{i=1}^{M} ET^{RSsub,i,j,t} \} \quad \forall t \tag{4.4}$$

$$Min S_{i,j} \leq s_{i,j} \leq Max S_{i,j} \; (1 \leq s_{i,j} \leq 29) \tag{4.5}$$

where **P**: the variables whose $s_{i,j}$ component is the soil hydraulic properties of all the soil contained within the *ij* pixel, $\theta^{sub,i,j,t}$: the simulated soil moisture of individual sub-pixel in the time index (*t*), *i*: the row number of sub-pixels with the domain (*M*), *j*: the column number of sub-pixels with the domain (*N*), $\theta^{sum,t}$: the sum of simulated soil moisture ($\theta^{sub,i,j,t}$) of individual sub-pixels in the time index (*t*), $ET^{sub,i,j,t}$: the simulated evapotranspiration of individual sub-pixel with the time index (*t*), $ET^{sum,t}$: the sum of simulated evapotranspiration ($ET^{sub,i,j,t}$) of individual sub-pixels with the time index (t),

$ET^{RSsub,i,j,t}$: the RS evapotranspiration product of individual sub-pixel with the time index (t), $ET^{RSsum,t}$: the sum of RS evapotranspiration products ($ET^{RSsub,i,j,t}$) of individual sub-pixels with the time index (*t*), and *t*: the time index, respectively. The $s_{i,j}$ component was constrained in the Eq. (4.5).

We used the un-mixing model (Eq. (4.6)) designed to solve P ($s_{i,j}$ component) using a simulation-optimization scheme based on inverse modeling [*Ines and Droogers*, 2002; *Ines and Mohanty*, 2008a,b,2009]. The un-mixing model indicates that a RS soil moisture product ($\theta^{RS,t}$) can be estimated by the sum ($\theta^{sum,t}$) of simulated soil moisture ($\theta^{sub,i,j,t}$) of individual sub-pixels and adding an error term (e_t). To minimize the error (e_t) between the observed/simulated SM and ET by tuning a hydrological model, we used a genetic algorithm. The objective (Z(**P**)) and fitness (*Fitness*(**P**)) functions are shown in Eq. (4.7-8) as below,

$$\theta^{RS,t}(\mathbf{P}) = \frac{1}{N}\frac{1}{M}\{\sum_{j=1}^{N}\sum_{i=1}^{M}\theta^{sub,i,j,t}\} + e_t \quad \forall t \tag{4.6}$$

$$Z(\mathbf{P}) = e_t = \text{Min}\{\frac{1}{T}\sum_{t=1}^{T}\{f\times|\theta^{sum,t} - \theta^{RS,t}| + (1-f)\times|ET^{sum,t} - ET^{RSsum,t}|\}\} \tag{4.7}$$

$$Fitness(\mathbf{P}) = Max[Z(\mathbf{P})]^{-1} \tag{4.8}$$

where $\theta^{RS,t}$: the remotely sensed soil moisture product in the time index (*t*), *Z*: the objective (minimizing) function (Note: All variables were weighted and normalized for the objective function), *f*: the weighting factor ($0 < f < 1.0$) indicating that the weighting factor of 0.1 means more weight is given to ET and the weighting factor of 0.9 means more weight to SM, and *Fitness*(**P**): the maximization function if Z(**P**) is minimized.

Genetic algorithms (GAs) are search algorithms to solve the optimized solutions for complex problems based on the survival of competing mechanism [*Holland*, 1975; *Goldberg*, 1989]. GAs are influenced by not only initial random generator seeds (e.g., idum: -3000, -2000, -1000, etc.), but also by the number of parameters (**P**) to be searched from unknown spaces. In this study, we selected the Ensemble Multiple Operators Genetic Algorithm (EMOGA, [*Shin and Mohanty*, 2012]). The Multiple Operators Genetic Algorithm (MOGA) has the unique ability to reproduce the fittest chromosomes (**P**) in the individual population as much as the number of chromosomes (parameters: $M \times N$) to be searched (see Fig. 4.2).

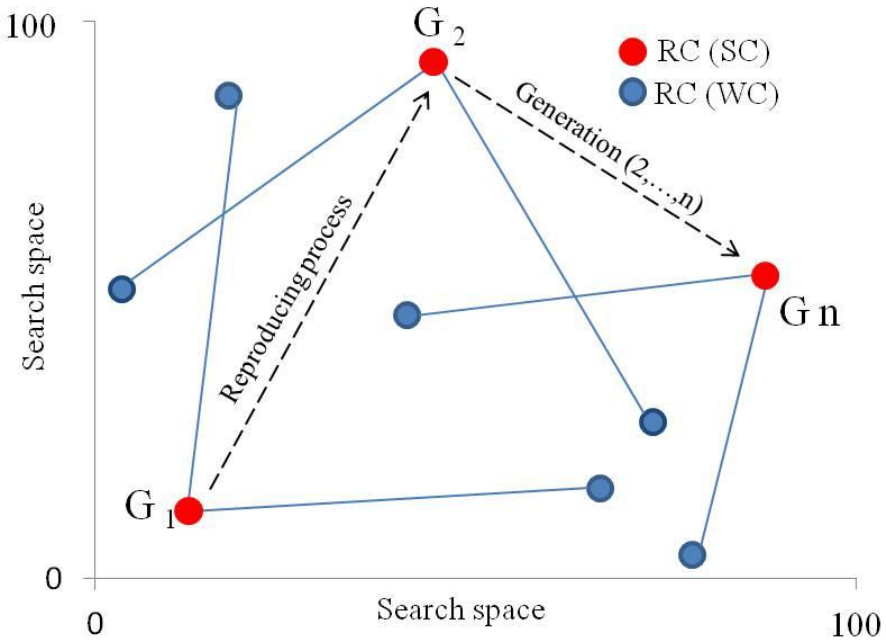

Figure 4.2: Schematic of ensemble multiple operators genetic algorithm (EMOGA)
RC: reproduced chromosomes; SC: strongest chromosome; WC: weak chromosomes

The reproduced chromosomes ($RC^{r=1,...M\times N}$) have new genetic information through the GA operators (selection, crossover, and mutation) and explore more search spaces. Then, the MOGA restarts when the chromosomes are converged to one region, which means that the better chromosomes are not searched than the previous strongest one for sequential 50 generations (note that the number of generations is subjective) at the converged region before all generations are completed. With the restarting technique, the MOGA provides new genetic materials through the creep and jump mutation operators [*Ines and Honda*, 2005]. The MOGA always remember the previous (g-1) elite chromosomes and reproduce in the next generation [*Ines and Mohanty*, 2008a]. We integrated a random re-sampling (ensemble) algorithm [IBM Programmers' Guide; Efron, 1982] into the MOGA for searching more unknown spaces, called EMOGA.

This approach with the EMOGA uses a physically-based (1-D) soil water atmosphere plant (SWAP) model to simulate soil water flow between the soil, water, atmosphere, and the plant system [*Kroes et al.*, 1999*; van Dam et al.*, 1997]. The SWAP model calculates the soil water flow based on Richards' equation in Eq. (4.9). We can describe the soil hydraulic functions by analytical expressions (*van Genuchten* [1980] and *Mualem* [1976]) using the relationship between the soil water content (θ), pressure head (h), and unsaturated hydraulic conductivity (K),

$$\frac{\partial \theta}{\partial t} = C(h)\frac{\partial h}{\partial t} = \frac{\partial [K(h)(\frac{\partial h}{\partial z} + 1)]}{\partial z} - S(h) \qquad (4.9)$$

where θ: the soil moisture content ($cm^3\ cm^{-3}$), K: the hydraulic conductivity ($cm\ d^{-1}$), h: the pressure head (-cm), z: the soil depth (cm) taken positively upward, t: the time

domain (d), *C*: the differential water capacity (cm^{-1}), and *S(h)*: the actual soil moisture extraction rate by plants ($cm^3\ cm^{-3}\ d^{-1}$) defined as Eq. (4.10).

$$S(h) = \alpha_w(h)\frac{T_{pot}}{Z_r} \tag{4.10}$$

where T_{pot}: the potential transpiration ($cm\ d^{-1}$), Z_r: the rooting depth (cm), and α_w: the reduction factor as function of h and accounts for water deficit and oxygen stress [*Feddes et al.*, 1978]. The Richards' equation (4.11) with the finite difference approach [*Belmans et al.*, 1983] allows the use of soil hydraulic database and various management scenarios.

$$S_e = \frac{\theta(h) - \theta_{res}}{\theta_{sat} - \theta_{res}} = \left[\frac{1}{1 + |\alpha h|^n}\right]^m \tag{4.11}$$

$$K(h) = K_{sat} S_e^{\lambda} \left[1 - \left(1 - S_e^{1/m}\right)^m\right]^2 \tag{4.12}$$

where S_e: the relative saturation (-), θ_{res}: the residual water contents ($cm^3\ cm^{-3}$), and θ_{sat}: the saturated water contents ($cm^3\ cm^{-3}$), α (cm^{-1}), n (-), m (-), and λ(-): the shape parameters of the retention and the conductivity functions, K_{sat}: the saturated hydraulic conductivity ($cm\ d^{-1}$), and m=1– 1/n, respectively.

The SWAP model considers various top and bottom boundary conditions such as weather conditions, water table depths, flux, aquifer, and surface drain, etc. [*van Dam et al.*, 1997]. The SWAP model has three crop routines: i) a simple model to simulate the impacts of weather, soil feature, and plant type, ii) a detailed model (WOFOST), and iii) the same model attuned to simulate crop growth. Also, the water management modules (irrigation and drainage) were combined with this model [*van Dam et al.*, 1997, *van Dam*, 2000]. The SWAP model estimates the potential evapotranspiration (ET_{pot}) using

the Penman-Monteith equation and partitions potential transpiration (T_{pot}) and soil evaporation (E_{pot}) by the leaf area index or the soil cover fraction. Then, ET_{pot} into actual ET (ET_{act}) is reduced through adjusting the E_{pot} and T_{pot} into actual values based on empirical relationships as the soil becomes dry. This model conducts well with various meteorological and environmental criteria [*Wesseling and Kroes, 1998; Sarwar et al., 2000; Droogers et al., 2000, Singh et al., 2006a*].

4.3.2 Physical Soil Texture Database

When the land surface within a RS product is comprised of several soil textures (e.g., sandy loam, silt loam, and clay loam, etc.), we traditionally need to search the effective soil hydraulic parameters (α, n, θ_{res}, θ_{sat}, K_{sat}) for each soil unit, which means that the parameter estimation would take a large combinational problem. We suggest a physical soil texture database including various soil information from the UNSODA [*Leij et al.,* 1999], Staring soil database [*Wösten et al.,* 1994] and Rosetta [*Schaap et al.,* 1999] in Table 4.1. This database contains the soil hydraulic properties (29 soil textures) of Mualem-van Genechthen. Using the physical soil texture database, the soil information ($s_{i,j}$) corresponding to the soil ID values can be provided for individual soil unit for the model performance.

Table 4.1: Physical soil texture database for the deterministic downscaling algorithm

Soil ID	Soil textures	Shape parameters		Scaling parameters		Hydraulic conductivity	
		α	n	θ_{res}	θ_{sat}	K_{sat}	λ
1[a]	Sandy Loam	0.021	1.61	0.067	0.37	41.6	0.5
2[b]	Sandy Loam	0.075	1.89	0.065	0.41	106.1	0.5
3[c]	Sandy Loam	0.027	1.45	0.039	0.39	38.3	-0.861
4[a]	Loam	0.025	1.31	0.083	0.46	38.3	0.5
5[b]	Loam	0.036	1.56	0.078	0.43	25.0	0.5
6[c]	Loam	0.011	1.47	0.061	0.40	12.1	-0.371
7[a]	Silt	0.006	1.53	0.123	0.48	55.7	0.5
8[b]	Silt	0.016	1.37	0.034	0.46	60.0	0.5
9[c]	Silt	0.007	1.68	0.050	0.49	43.8	0.624
10[a]	Silt Loam	0.012	1.39	0.061	0.43	30.5	0.5
11[b]	Silt Loam	0.020	1.41	0.067	0.45	10.8	0.5
12[c]	Silt Loam	0.005	1.66	0.065	0.44	18.2	0.365
13[a]	Sandy Clay Loam	0.033	1.49	0.086	0.40	9.7	0.5
14[b]	Sandy Clay Loam	0.059	1.48	0.100	0.39	31.4	0.5
15[c]	Sandy Clay Loam	0.021	1.33	0.063	0.38	13.2	-1.280
16[a]	Clay Loam	0.030	1.37	0.129	0.47	1.8	0.5
17[b]	Clay Loam	0.019	1.31	0.095	0.41	6.2	0.5
18[c]	Clay Loam	0.016	1.42	0.079	0.44	8.2	-0.763
19[a]	Silty Clay Loam	0.027	1.41	0.098	0.55	7.4	0.5
20[b]	Silty Clay Loam	0.010	1.23	0.089	0.43	1.7	0.5
21[c]	Silty Clay Loam	0.008	1.52	0.090	0.48	11.1	-0.156
22[b]	Sandy Clay	0.027	1.23	0.100	0.38	2.9	0.5
23[c]	Sandy Clay	0.033	1.21	0.117	0.39	11.4	-3.665
24[a]	Silty Clay	0.023	1.39	0.163	0.47	8.4	0.5
25[b]	Silty Clay	0.005	1.09	0.070	0.36	0.5	0.5
26[c]	Silty Clay	0.016	1.32	0.111	0.48	9.6	-1.287
27[a]	Clay	0.021	1.20	0.102	0.51	26.0	0.5
28[b]	Clay	0.008	1.09	0.068	0.38	4.8	0.5
29[c]	Clay	0.015	1.25	0.098	0.46	14.8	-1.561

[a]UNSODA[*Leij et al.,* 1999]

[b]SoilSurvey[*Wösten et al.,* 1994]

[c]Rosetta[*Schaap et al.,* 1999]

4.3.3 Simplified-Surface Energy Balance Index (S-SEBI) Model

Evapotranspiration is the process by which liquid water evaporates from open water, soil, and plant surfaces to the atmosphere across the land-atmosphere boundary. Remote sensing measurements of the surface energy balance provide a means to evaluate spatially and temporally distributed vegetation conditions at large scales [*Moran et al.,* 1995*; Moulin et al.,* 1998]. In this study, we selected a Simplified-Surface Energy Balance Index (S-SEBI) model for calculating pixel-based evapotranspiration (ET) estimates using RS products [*Roerink et al.,* 2000]. The land surface energy balance can be solved on a pixel by pixel basis using RS (e.g., LANDSAT5-TM, Moderate Resolution Imaging Spectroradiometer-MODIS, Advanced Very High Resolution Radiometer-AVHRR, etc.) datasets. The land surface energy balance is given by,

$$R_n = G_0 + H + \lambda E \tag{4.13}$$

where, R_n: the net radiation [W/m^2], G_0: the soil heat flux [W/m^2], H: the sensible heat flux [W/m^2], and λE: the latent heat flux [W/m^2].

The S-SEBI model uses scanned spectral radiance (visible, near-infrared, and thermal infrared range) for estimating surface reflectance, surface temperature, and vegetation index under clear weather conditions. We used the LANDSAT5-TM images (30 m$\times$30 m) to compute fine-scale pixel-based ET using the S-SEBI model in Table 4.2.

Table 4.2: Description of the LANDSAT5-TM

Image Character	Value
Sensor	LANDSAT5-TM
Path/Row	28/36
Acquisition date (time)	20 April 1997 (16:35:29) 07 June 1997 (16:37:10) 09 July 1997 (16:38:16)
Reference system	UTM-24N
Resolution	Band 1 to 7 (30m×30m)

4.3.4 Data Assimilation Framework

We conducted numerical data assimilation experiments for assessing the deterministic downscaling algorithm based on inverse modeling. The experiments were composed of synthetic and field validation experiments under rain-fed conditions including: i) synthetic experiments with various land surface conditions, ii) impacts of different vegetation covers and ground water tables under the synthetic conditions, and iii) field validation experiments, respectively.

We designed three (simple - 2×2 matrix, relatively complex - 3×3 matrix, and complex - 4 × 4 matrix) land surface conditions representing homogeneous and heterogeneous land surface with free drainage (indicating semi/arid regions) for testing this approach under the synthetic condition at the Lubbock site (March 1 to July 31, 2002) in Texas (Fig. 4.1b). We collected the weather datasets at the Irrigation Technology Center (http://texaset.tamu.edu/index.php).

An evapotranspiration component is the major factor for downscaling RS soil moisture products. To confirm the impact of ET component for the downscaling process, we analyzed the characteristics of (synthetic) soil moisture (0-1 cm) and ET dynamics by the combinations of different soil textures (see Table 4.1: e.g., Soil ID: 1-sandy loam, 4-loam, 7-silt, and 13- sandy clay loam) and various vegetation covers (wheat, soybean, grass, and maize). The various weighting factors (0.1 to 0.9) for the objective function (Eq. (4.7)) were also tested under the relatively complex land surface condition.

Additional experiments were conducted to evaluate the impacts of various vegetation covers (wheat, soybean, grass, and maize) and varying ground water table depths of -200, -150, and -100 cm (Fig. 4.1c) under the relatively complex land surface condition. The bottom flux (positive upward) is set to test that the interdependency assumption used in the inverse modeling is still executable when the soil water flow in the unsaturated zone is dominated by the bottom boundary condition, e.g., by major upward flows from the ground water table (*van Dam*, 2000).

For the field validation experiments, we selected the Little Washita watershed (LW 13 and 21 footprints) in Oklahoma during the simulation period (March 1 to July 31, 1997) as shown in Fig. 4.3. The *in-situ*/pixel-based soil moisture (0-5 cm) were measured during the Southern Great Plains experiment (SGP97) from June 18, 1997 to July 18, 1997 [*Mohanty et al.*, 2002]. The daily volumetric soil moisture datasets (24 days for the LW 13 and 17 days for the LW 21 sites) were measured at the 49 (7×7) sampling points, except for the heavy rainfall events. We used the airborne

Electronically Scanned Thin Array Radiometer (ESTAR) pixel-based soil moisture products for 17 days at the LW sites [*Jackson et al.*, 1999].

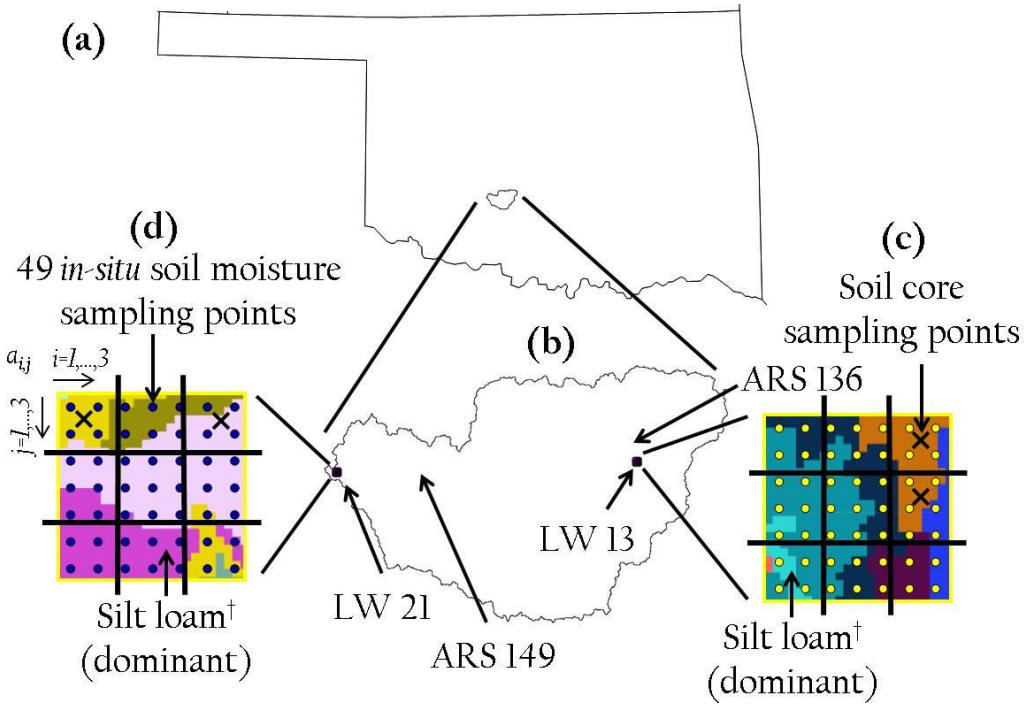

Figure 4.3: (a) Oklahoma, (b) Little Washita (LW) watershed, (c) LW 13 site, (d) LW 21 site including the *in-situ* soil moisture sampling points and weather stations for the field experiments

†Silt loam soil with different slope ranges.
Cross indicates the soil core sampling points.

This approach downscaled the ESTAR soil moisture products under relatively complex land surface condition and validated subgrid (or sub-pixel) soil moisture values with the *in-situ* soil moisture measurements. The soil core samples in the soil depth of 3-9 cm were collected to obtain the effective soil hydraulic properties [*Mohanty et al.*, 2002] at the field sites. The sub-pixels $a_{1,3}$ and $a_{2,3}$ at the LW 13 site (Fig. 4.3) have three

and two soil core samples (i.e., soil hydraulic properties), respectively. Soil core sample at the sub-pixels $a_{1,1}$ and $a_{1,3}$ were taken at the LW 21 site. Besides soil moisture dynamics ($\theta(t)$) at 49 *in-situ* sampling locations for LW13 or LW21 (Fig 4.3), we further validated the downscaling approach through the comparison of soil water retention functions ($\theta(h)$). They were derived by the searched soil ID values (in terms of their soil hydraulic parameters in Table 4.1) and field-observed soil hydraulic properties in Table 4.3. We excluded the hydraulic conductivity ($K(h)$) functions for validation, because of the extreme variations in K_{sat} (5.063-129.427 mm day^{-1}).

Table 4.3: Soil hydraulic properties derived by the soil core samples obtained at the LW 13 and 21 sites

Sites	Soil depth	Sub-pixel (a_{ij})	α^a	n^a	$\theta_{res}{}^a$	$\theta_{sat}{}^a$	$K_{sat}{}^a$
LW 13	3-9 cm	3 (a_{13})	0.009	1.430	0.115	0.354	6.679
			0.015	1.204	0.002	0.322	5.063
			0.026	1.229	0.109	0.366	5.702
		6 (a_{23})	0.012	1.262	0.106	0.435	129.427
			0.009	1.387	0.068	0.387	22.550
LW 21	3-9 cm	1 (a_{11})	0.006	1.581	0.117	0.429	31.795
		3 (a_{13})	0.009	1.734	0.115	0.432	17.885

[a]Field observations [*Mohanty et al*., 2002]

The pixel-based ET (30m×30m) estimates using the LANDSAT5-TM datasets were computed by the S-SEBI model [*Roerink et al.,* 2000] for a few days (i.e., 3 days for LW13 and 2 days for LW 21 during the simulation period due to the limited

available LANDSAT5-TM datasets and weather condition. The LW 13 and 21 sites have the silt loam (predominant), sandy loam, and loam soils (from SSURGO: http://www.nrcs.usda.gov/). During the simulation period in the LW 21 site, two-thirds were covered by the winter wheat, and the other third was covered by the short native grass. The wheat cover area is flat (slope less than 1%) and the grass cover area has a rolling slope (3-12%). The LW 13 site also has rolling topography with grass cover and a small pond in sub-pixel $a_{2,1}$ as shown in Fig. 4.4.

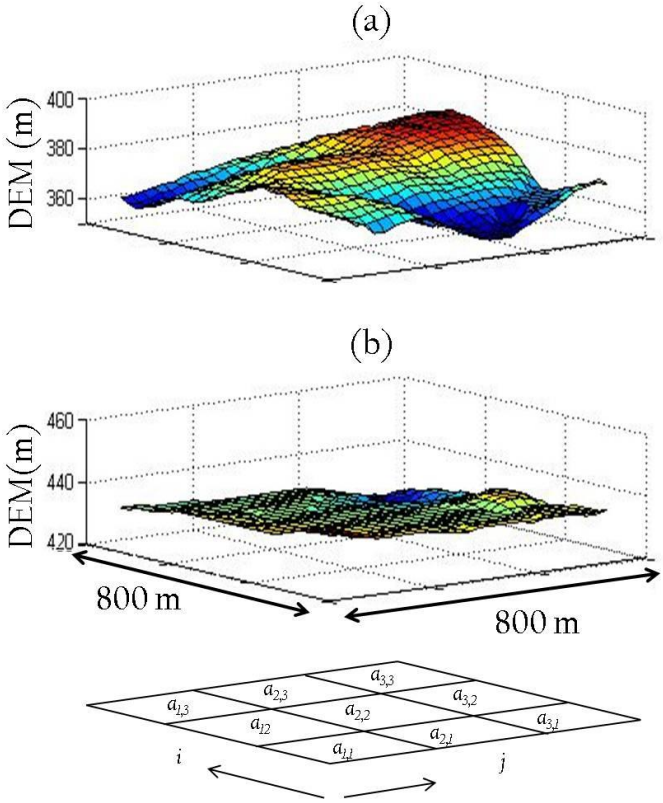

Figure 4.4: Digital elevation method (DEM) for the study sites at the Little Washita (LW) watershed in Oklahoma; (a) LW 13 site, (b) LW 21 site

We set the different crop growing periods (March 1 to June 27 for the wheat and grass cover at the LW 21 site and March 1 to July 31 for the LW 13 site), because the wheat was harvested on June 27 at the LW 21 site. Daily climatic datasets (e.g., precipitation, wind speed, maximum and minimum temperature, and solar radiation) for the model input were collected from the USDA-Agricultural Research Service micronet weather station (ARS 133 for the LW 13 and ARS 149 for the LW 21, http://ars.mesonet.org/) in Oklahoma.

In this study, we assumed that the remote sensing pixel (field) comprised of parallel stream tubes or soil columns (sub-pixels) and the vertical soil columns are discretized by 33 texturally-homogeneous computational layers. The soil profile was discretized at the intervals of 1 cm for the top 10 cm (1-10th layer) from the soil surface. For the soil depths of 10-60 cm (11-20th layer) and 60-200 cm (20-32nd layer), vertical grid intervals of 5 cm and 10 cm were used (except 20 cm in the 33rd layer). The initial conditions ($h(z,t=0)$ = -100 cm) were specified for the soil columns in the synthetic cases. We tested this approach under the free drainage and various ground water table depths (-200, -150, and -100 cm), because of the unknown bottom boundary conditions at the field sites. We assumed that the initial conditions were in equilibrium with the bottom boundary conditions. The model parameters used for genetic algorithm and SWAP model for the synthetic and field validation experiments are shown in Table 4.4.

Table 4.4: Modeling conditions of genetic algorithm and SWAP model

	Synthetic experiments	Field experiments
GA parameters:		
No. population	30	30
No. seed	-3000, -2000, -1000	-3000, -2000, -1000
No. generation	5000	5000
No. search restart	4 (only complex condition)	-
No. ensemble	10	10
Modeling Conditions:		
Top boundary	Time dependent flux/head	Time dependent flux/head
Bottom boundary	Free drainage, ground water table depth (-100, -150, -200 cm)	Free drainage, ground water table depth (-100, -150, -200 cm)
Initial conditions	$h(z,t=0)$=-100 cm	Equilibrium with bottom boundary conditions

For the uncertainty analysis, we used the simulation-optimization scheme (SWAP-EMOGA) with re-sampling (ensemble e) and multi-population by various random generator seeds (-3000, -2000, and -1000) and selected the maximum fitness for the individual simulations, respectively. Using the selected solutions (e×three different random seeds), we estimated the Pearson's correlation (R^2) and mean bias error (MBE). The 95 percent confidence interval (95PCI) of the selected solutions was calculated as,

$$Range_{t,i,j} = 95PCI_{t,i,j+} - 95PCI_{t,i,j-} \quad (4.14)$$

where $95PCI_{t,i,j+}$ is the upper boundary, $95PCI_{t,i,j-}$ is the lower boundary, and t is the time (running) index.

4.4 Results and Discussion

4.4.1 Synthetic Experiments - Various Land Surface Conditions

The various combinations of soil textures (soil ID: 1, 4, 7, 13) and vegetation covers (wheat, soybean, grass, and maize) have their unique characteristics of soil water retention ($\theta(h)$, $cm^3\ cm^{-3}$), hydraulic conductivity ($K(h)$, $cm\ day^{-1}$), soil moisture dynamics $\theta(t)$ ($cm^3\ cm^{-3}$), and evapotranspiration ET (t) ($mm\ day^{-1}$) in Fig. 4.5. The $\theta(h)$ function of silt (ID: 7) soil is higher than those of loam (ID: 4), sandy clay loam (ID: 13), and sandy loam (ID: 1) soils (in Fig. 4.5a,b). Fig. 4.5c,d show the impacts of various vegetation covers on estimating the SM and ET dynamics with a loam (ID: 4) soil. Overall, the grass cover with the average soil moisture (0.287 $cm^3\ cm^{-3}$) contains more moisture in the soil matrix than others (soybean-0.233, maize-0.221, and wheat-0.183 $cm^3\ cm^{-3}$) during the simulation period. However, average ET estimates (wheat-2.504, soybean-2.278, maize-1.866, and grass-1.159 $mm\ day^{-1}$) are contrary to the SM trends, except for the soybean cover, which has both relatively high SM and ET. It shows that the vegetation covers influence not only the ET estimates but also soil moisture.

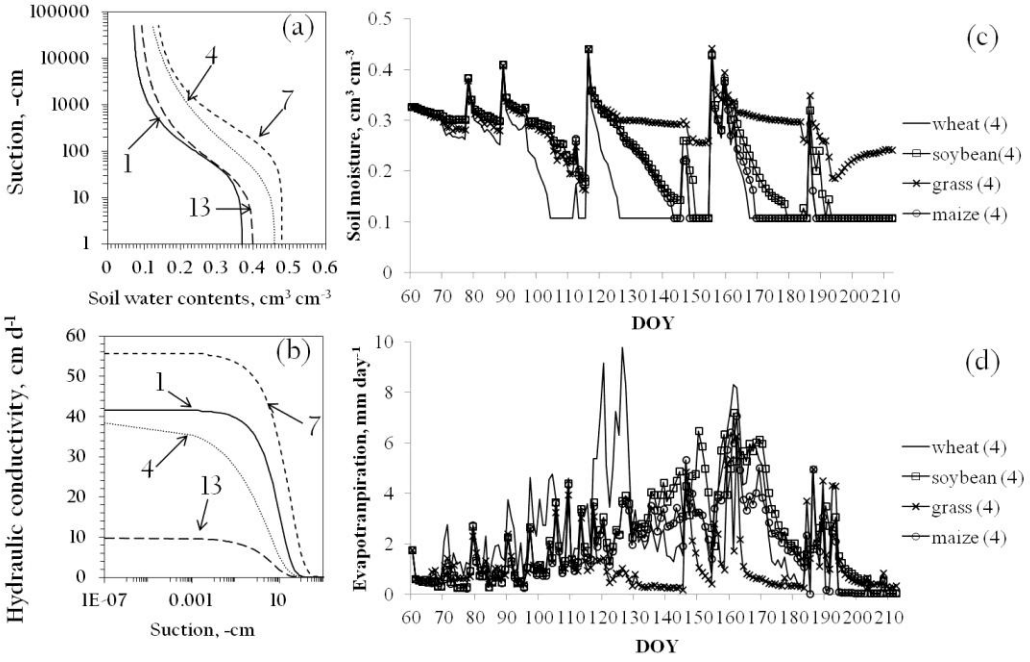

Figure 4.5: Various characteristics of the combinations of soil textures (soil ID: 1-sandy loam, 4-loam, 7-silt, and 13-sandy clay loam) and vegetation covers (wheat, soybean, grass, and maize); (a) soil water contents (θ(h)), (b) hydraulic conductivities (K(h)), (c) soil moisture dynamics (cm^3 cm^{-3}), (d) evapotranspiration (mm day^{-1})

We tested the range of weighting factors f (0.1~0.9) for the DDA with the relatively complex land surface condition. When the weighting factors (f) were in the ranges of 0.1 to 0.5, the derived solutions were identified well with the synthetic observations in Table 4.5. The derived soil ID values ($s_{i=1,\ldots,3,j=1,\ldots,3}$) of individual sub-pixels with the weighting factor (f=0.1) have the highest fitness than others (f=0.2~0.5). But when more weights (f=0.6~0.9) were given to the soil moisture, maximum fitness decreased. Furthermore, the soil ID values for sub-pixels with weighting factors (f=0.6 and 0.8) did not match well with the observations compared to the results for weighting

factors (f=0.1 to 0.5). Most of all, the DDA showed good matching with the synthetic observations (soil ID), but the specific locations of derived soil ID values were not matched in the range of f=0.6~0.9. It demonstrated that the ET component assigns the downscaled soil moisture estimates (by the searched soil ID values) of individual sub-pixels to the specific locations within a RS product.

Table 4.6 shows the derived solutions (soil ID) under various land surface conditions. The solutions for each sub-pixel are fairly well identified (minimum uncertainties) with the synthetic observations under the simple land surface condition. The results with the relatively complex land surface condition also matched well with the observations, even though they have small errors for sub-pixels ($a_{1,1}$, $a_{1,2}$, $a_{1,3}$, $a_{2,1}$, $a_{2,2}$, $a_{3,3}$). However, the uncertainty ranges of solutions increased considerably when the land surface condition becomes complex. The DDA approach found the exact soil ID values ($a_{1,1}$: 4, $a_{1,2}$: 7, $a_{1,3}$: 7, $a_{1,4}$: 13, $a_{2,1}$: 27, $a_{2,2}$: 4, $a_{2,3}$: 1, $a_{2,4}$: 20, $a_{3,1}$: 20, $a_{3,2}$: 7, $a_{3,3}$: 4, $a_{3,4}$: 7, $a_{4,1}$: 13, $a_{4,2}$: 27, $a_{4,3}$: 27, $a_{4,4}$: 1 from the physical soil texture database in Table 4.1) of sub-pixels, but the solution uncertainties for the complex condition were considerably higher than those of the simple and relatively complex conditions.

Table 4.5: Solutions (soil ID) of sub-pixels derived by the deterministic downscaling algorithm using the genetic algorithm (EMOGA) based on the various weighting factors (f=0.1 to 0.9) under the relatively complex land surface condition for Case 1

Sub-pixel ($a_{i,j}$)	Observations (Soil ID)*	Relatively complex land surface condition								
		Weighting factor								
		f=0.1	*f*=0.2	*f*=0.3	*f*=0.4	*f*=0.5	*f*=0.6	*f*=0.7	*f*=0.8	*f*=0.9
Maximum Fitness		**1.11E+07**	9.44E+06	8.39E+06	7.33E+06	6.27E+06	5.84E+05	1.82E+02	7.86E+05	2.21E+02
$a_{1,1}$	4	**4**	**4**	**4**	**4**	**4**	7	7	1	5
$a_{1,2}$	7	**7**	**7**	**7**	**7**	**7**	1	22	**7**	18
$a_{1,3}$	13	**13**	**13**	**13**	**13**	**13**	**13**	**13**	**13**	1
$a_{2,1}$	1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
$a_{2,2}$	4	**4**	**4**	**4**	**4**	**4**	13	18	1	7
$a_{2,3}$	7	**7**	**7**	**7**	**7**	**7**	4	**7**	**7**	5
$a_{3,1}$	13	**13**	**13**	**13**	**13**	**13**	7	11	4	1
$a_{3,2}$	1	**1**	**1**	**1**	**1**	**1**	4	**1**	4	20
$a_{3,3}$	1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	13	7

*Soil identification (soil ID) values from the simplified soil hydraulic database
Vegetation cover: wheat crop
Bold: the exact solution searched

Table 4.6: Solutions (soil ID) derived by the deterministic downscaling algorithm using the genetic algorithm (EMOGA) under the simple, relatively complex, and complex land surface conditions for Case 1

Sub-pixel ($a_{i,j}$)	Simplex land surface condition			Relatively complex condition			Complex land surface condition		
	Observations (Soil ID)*	Vegetation cover**	Solution Soil ID	Observations (Soil ID)*	Vegetation cover**	Solution Soil ID	Observations (Soil ID)*	Vegetation cover**	Solution Soil ID
$a_{1,1}$	4	1-4	**4**	4	1	**4,7**	4	1	**4**,27
$a_{1,2}$	7	1-4	**7**	7	1	**4,7**	7	2	**7**,15,27
$a_{1,3}$	-	-	-	13	1	1,**13**	7	1	6,**7**,26,29
$a_{1,4}$	-	-	-	-	-	-	13	1	5, **13**
$a_{2,1}$	13	1-4	**13**	1	1	**1**,4	27	2	4,10,**27**
$a_{2,2}$	1	1-4	**1**	4	1	**4**,13	4	2	**4**,7,10,27
$a_{2,3}$	-	-	-	7	1	**7**	1	1	**1**
$a_{2,4}$	-	-	-	-	-	-	20	1	**20**
$a_{3,1}$	-	-	-	13	1	**13**	20	2	4,**20**,27
$a_{3,2}$	-	-	-	1	1	**1**	7	2	**7**,27
$a_{3,3}$	-	-	-	1	1	**1,**13	4	1	**4**,27
$a_{3,4}$	-	-	-	-	-	-	7	1	6,**7**,29
$a_{4,1}$	-	-	-	-	-	-	13	1	5,**13**
$a_{4,2}$	-	-	-	-	-	-	27	2	10,11,20,**27**
$a_{4,3}$	-	-	-	-	-	-	27	2	4,7,8,10,11,18,**27**,28
$a_{4,4}$	-	-	-	-	-	-	1	1	**1**

*Soil identification (soil ID) values from the simplified soil hydraulic database; **Vegetation covers: 1-wheat, 2-soybean, 3-grass, 4-maize
Bold: the exact solution searched

Figures 4.6 and 4.7 show the soil moisture dynamics (derived by the searched solutions in Table 4.4) estimated by this approach and ET for the sub-pixels under the complex land surface condition (Figures for the simple and relatively complex conditions are not shown). The soil moisture dynamics of sub-pixels ($a_{1,1}$, $a_{2,2}$, and $a_{4,3}$) have a bias compared to the observations as the mean bias error (MBE) were -0.031, -0.049, and 0.049, respectively. But the other sub-pixels matched well to the observations with -0.021 ~ -0.013 for the MBE. The simulated ET values tend to be similar with the results of soil moisture as the sub-pixels ($a_{1,1}$, $a_{2,2}$, and $a_{4,3}$) have more uncertainties (MBE: -0.009, -0.014, -0.011) than those (MBE: -0.006 ~ 0.000) of the other sub-pixels. From the results for the sub-pixels of $a_{1,2}$ (silt and soybean), $a_{1,3}$ (silt and wheat), and $a_{2,2}$ (loam and soybean), we confirmed that different vegetation covers of sub-pixels (soybean for $a_{1,2}$ and wheat for $a_{1,3}$) with the same soil texture (ID: 7) influence not only the soil moisture, but also ET estimates. The soil moisture dynamics of sub-pixels ($a_{1,2}$ and $a_{2,2}$) with different soil textures (ID: 4, 7), which have the same vegetation cover (soybean), were affected by different soil textures (silt vs. loam), but the vegetation type influences the ET values slightly more than soil textures. On the basis of above findings, we suggest that the ET component is the key factor for this downscaling approach.

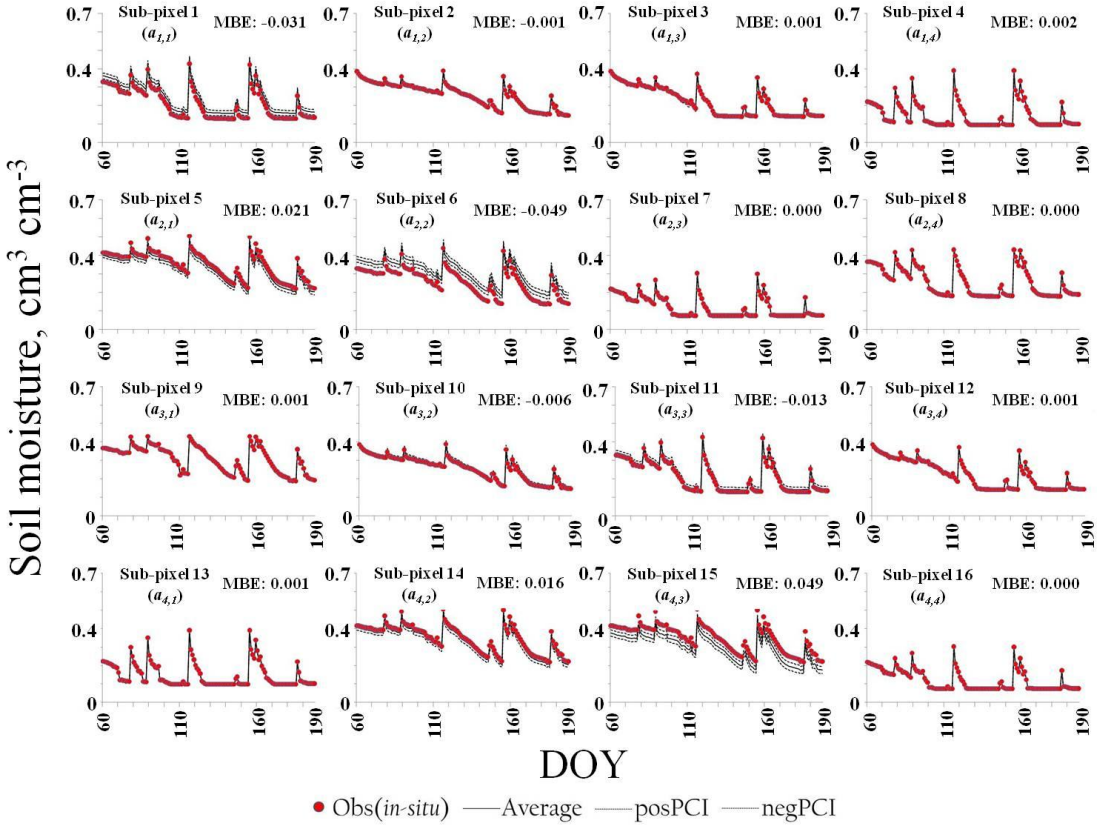

Figure 4.6: Soil moisture dynamics (0-1 cm) of sub-pixels downscaled by the deterministic downscaling algorithm using the genetic algorithm (EMOGA) under the complex land surface condition for Case 1

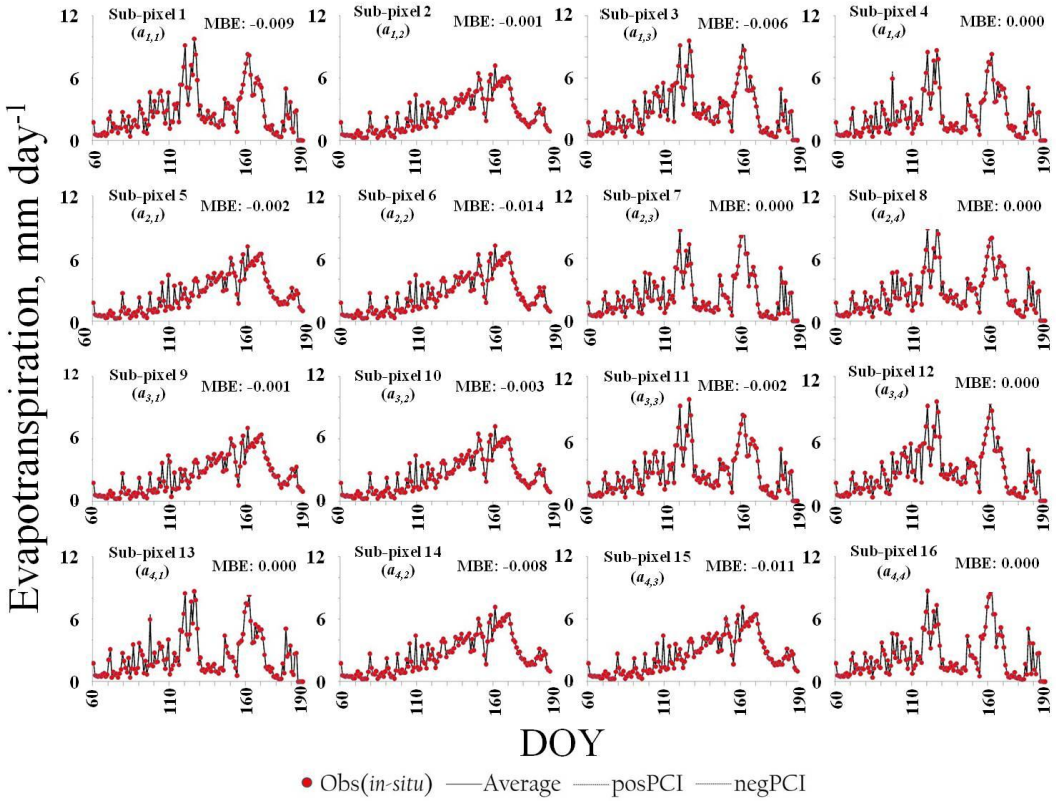

Figure 4.7: Evapotranspiration of sub-pixels by the deterministic downscaling algorithm using the genetic algorithm (EMOGA) under the complex land surface condition for Case 1

4.4.2 Impacts of Various Vegetation Covers and Ground Water Tables

Table 4.7 shows the results of various vegetation covers (wheat, soybean, grass, and maize) under synthetic condition. Overall, the DDA approach searched well the soil ID values of sub-pixels with various land covers under relatively complex condition, although the solutions with the wheat and soybean covers have more uncertainties than those of the grass and maize crops. The derived soil ID values for the maize cover perfectly matched with the synthetic observations. The results with different water table depths (-200, -150, and -100 cm) show somewhat the similar trend with those of the free drainage condition (Table 4.8). The solutions with the GW -100 cm have more variations than those of GW -150 cm indicating that the estimations of soil parameters were disturbed due to the upward flow of shallow ground water table. The soil ID values with the GW -200 cm relatively have less uncertainties in modeling, but the results (ID: 1-sandy loam) of sub-pixels ($a_{1,1}$ and $a_{2,2}$) were not matched with the observation (ID: 4-Loam). The soil ID values (ID: 1, 4) have the similar soil hydraulic properties (see Table 4.1) indicating that this approach can be affected by the non-sensitivity of the similar soil water content ($\theta(h)$) and hydraulic conductivity ($K(h)$) curves. Overall, the impacts of ground water tables were less sensitive than the vegetation covers, which mean that the land surface conditions (e.g., soil textures, land covers, atmospheric condition, etc.) influence this approach more than the ground water tables.

Table 4.7: Solutions (soil ID) derived by the deterministic downscaling algorithm using the genetic algorithm (EMOGA) with various vegetation covers (wheat, soybean, grass, and maize) under the relatively complex land surface condition for Case 2

Sub-pixel ($a_{i,j}$)	Observations* (Soil ID)	Relatively complex land surface condition			
		wheat	soybean	grass	maize
$a_{1,1}$	4	**4**	**4**	**4**,10,18	**4**
$a_{1,2}$	7	**7**,10	**7**	**7**	**7**
$a_{1,3}$	13	1,5,**13**	7,**13**	**13**	**13**
$a_{2,1}$	1	**1**,4	**1**,13	**1**	**1**
$a_{2,2}$	4	**4**,13	**4**,27	**4**	**4**
$a_{2,3}$	7	**7**,10	4,**7**,13	**7**	**7**
$a_{3,1}$	13	**13**	1,7,**13**	**13**	**13**
$a_{3,2}$	1	**1**	**1**	**1**	**1**
$a_{3,3}$	1	**1**	**1**	**1**	**1**

*Soil identification (soil ID) values from the simplified soil hydraulic database
Bold: the exact solution searched

Table 4.8: Solutions (soil ID) derived by the deterministic downscaling algorithm using the genetic algorithm (EMOGA) for different ground water table depths (-200, -150, and -100 cm) with the wheat crop under the relatively complex land surface condition for Case 2

Sub-pixel ($a_{i,j}$)	Observations* (Soil ID)	GW -100 cm	GW -150 cm	GW -200 cm
$a_{1,1}$	4	**4**,27	**4**	1
$a_{1,2}$	7	**7**,12	**7**,9,12	**7**,12
$a_{1,3}$	13	**13**	**13**	**13**
$a_{2,1}$	1	**1**	**1**	**1**
$a_{2,2}$	4	**4**,27	**4**	1
$a_{2,3}$	7	**7**,9,12	**7**,9,12	**7**,12
$a_{3,1}$	13	5,**13**	**13**	1,**13**
$a_{3,2}$	1	**1**	**1**	**1**,13
$a_{3,3}$	1	**1**	**1**	**1**

*Soil identification (soil ID) values from the simplified soil hydraulic database
Bold: the exact solution searched

4.4.3 Field Validation Experiments

Figure 4.8 shows the statistics (maximum, minimum, median, and 1st/3rd quartiles of solutions) of model performance with free drainage condition and various ground water table depths for the field validation experiments.

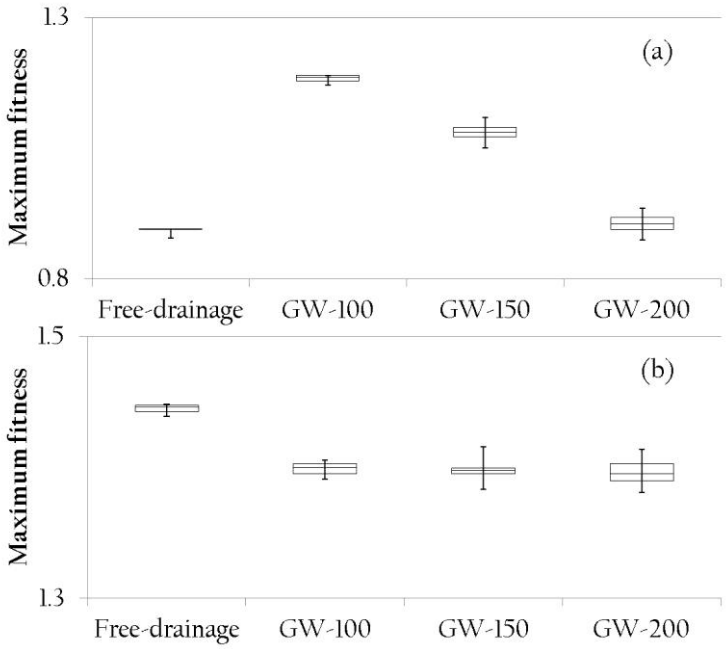

Figure 4.8: Maximum fitness of the field experiments (10 ensembles × three different random number generator seeds) with the various bottom boundary conditions (free drainage condition and GW-100, -150, -200 cm); (a) LW 13 site, (b) LW 21 site

We tested the various bottom boundary conditions for the field sites under the free drainage and ground water tables (-200, -150, and -100 cm) in modeling and the derived solutions for the ground water depth (GW) of -100 cm and free drainage condition at the LW 13 and 21 sites identified better with the measurements than those

of the other conditions, respectively. Thus, we selected the bottom boundary conditions of GW -100 cm (LW 13) and free drainage condition (LW 21) for the field sites. The soil ID values derived by this approach were shown in Table 4.9. The solutions (soil ID values) of sub-pixels vary across the range of sandy loam (ID: 2, 3), loam (ID: 6), silt (ID: 8), silt loam (ID: 10), and silty clay loam (ID: 21) at the LW 13 site. However, silt and silt loam soils (ID: 8, 9) are dominant (30-70 %) in the solutions of sub-pixels compared to the actual soil textures (silt loam-predominant, sandy loam, and loam soils). Similar trends are also shown in the results of the LW 21 site. The range of derived solutions including sandy loam (ID: 1, 2, 3), loam (ID: 5), silt (ID: 7, 9), silt loam (ID: 12), and sandy clay loam (ID: 13, 14) is more variable than those of the LW 13 site. As mentioned above, the soil water retention ($\theta(h)$) and hydraulic conductivity ($K(h)$) curves from the physical soil texture database across a range of pressure heads have similarity, indicating that the search of soil textures (soil ID values) by this approach can be limited due to the non-sensitivity of $\theta(h)$ and $K(h)$ functions (as in Fig. 4.9). However, this approach still searched silt and silt loam soils (ID: 7, 9, 12) as the dominant soils (36.7-73.3 %) for the individual sub-pixels, except for the sub-pixel of $a_{2,3}$. When we compared the solutions at the LW 13 and 21 sites, a sandy clay loam soil (ID: 14) was relatively more prominent in the derived solutions ($a_{1,1}$ and $a_{2,1}$) at the LW 21 site. This is because soil moisture at the LW 21 site were measured only for 17 days during the dry season indicating that the actual soil conditions could favor more a sandy or sandy clay loam soil. On the contrary, the LW 13 site has a small pond (sub-pixel: $a_{2,1}$) (Fig. 4.4), which means that the actual field site has more moisture explaining better

solutions for the shallow GW of -100 cm than the others. Thus, it is not unusual that this approach searched different soil ID values, even though the field sites have similar soil textures.

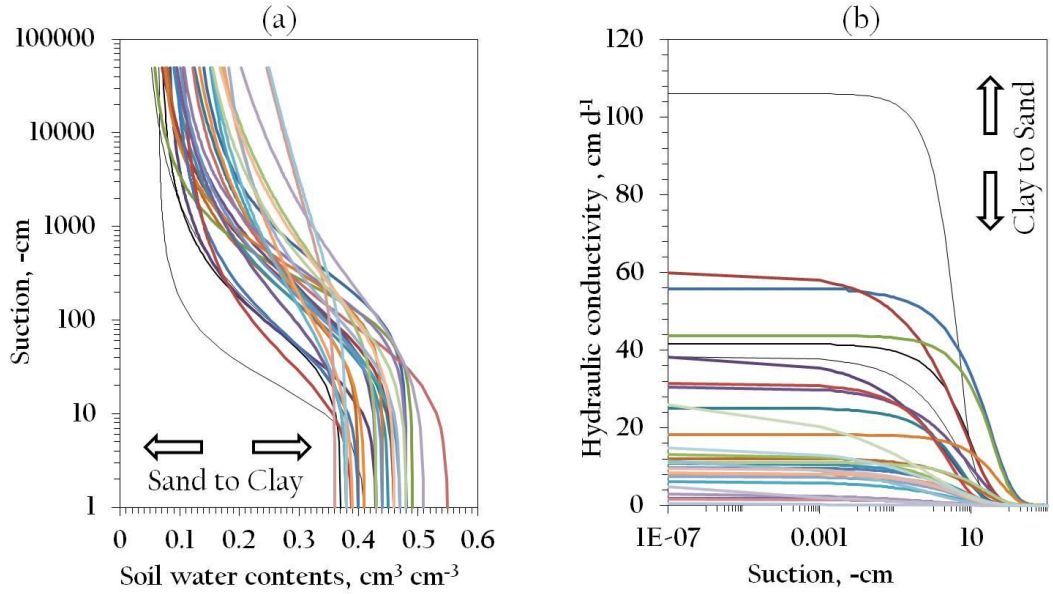

Figure 4.9: (a) Soil water contents ($\theta(h)$) and (b) hydraulic conductivities ($K(h)$) of physical soil texture database

UNSODA[*Leij et al.*, 1999]; SoilSurvey[*Wösten et al.*, 1994]; Rosetta[*Schaap et al.*, 1999]

Table 4.9: Solutions (soil ID) derived by the deterministic downscaling algorithm using the genetic algorithm (EMOGA) at the LW 13 and 21 sites for Case 3

Study sites	Sub-pixel ($a_{i,j}$)																	
	$a_{1,1}$		$a_{1,2}$		$a_{1,3}$		$a_{2,1}$		$a_{2,2}$		$a_{2,3}$		$a_{3,1}$		$a_{3,2}$		$a_{3,3}$	
	Soil ID	%	Soil ID	%	Soil ID	%	Soil ID	%	Soil ID	%	Soil ID	%	Soil ID	%	Soil ID	%	Soil ID	%
LW 13	3	10.0	2	10.0	3	6.7	3	16.7	2	3.3	2	6.7	2	3.3	2	3.3	2	3.3
	6	23.3	**3**	**30.0**	6	23.3	6	36.7	3	6.7	3	10.0	3	23.3	3	6.7	3	3.3
	8	**63.3**	**6**	**30.0**	**8**	**66.7**	**8**	**46.6**	6	33.3	6	40.0	6	16.7	6	30.0	6	26.7
	21	3.4	**8**	**30.0**	10	3.3			**8**	**50.0**	**8**	**43.3**	**8**	**56.7**	**8**	**56.7**	**8**	**56.7**
									21	6.7					21	3.3	10	6.7
																	21	3.3
	Sum	100		100		100		100		100		100		100		100		100
LW 21	**9**	**43.3**	**9**	**30.0**	**12**	**26.7**	**14**	**50.0**	5	13.3	**2**	**33.3**	**9**	**73.3**	**9**	**53.3**	1	13.3
	5	6.7	1	10.0	2	16.7	9	43.3	**9**	**30.0**	12	20.0	14	23.3	1	23.3	2	10.0
	14	**43.3**	2	23.3	**9**	**26.7**	3	3.3	1	20.0	1	23.3	1	3.4	5	6.7	5	10.0
	13	3.3	12	13.3	1	16.7	1	3.4	12	23.3	9	13.3			12	13.3	**9**	**43.3**
	1	3.4	5	20.0	5	6.7			2	13.4	7	3.3			2	3.4	12	23.4
			7	3.4	3	3.3					5	3.3						
					7	3.4					14	3.4						
	Sum	100		100		100		100		100		100		100		100		100

See Table 4.1 for soil identification (soil ID) values from the simplified soil hydraulic database
Bold: the exact solution searched

Figure 4.10a shows the measured (*in-situ*) and downscaled soil moisture dynamics of sub-pixels ($a_{i:1,..,3,j:1,..,3}$) at the LW 21 site. Mostly, the approach downscaled reasonably well the soil moisture estimates (R^2: 0.343 ~ 0.865 and MBE: -0.165 ~ -0.122) for the individual sub-pixels with the measurements, although the soil moisture values in the sub-pixels ($a_{1,3}$, $a_{2,2}$, $a_{3,1}$) have uncertainties (below 0.1 cm^3 cm^{-3}). The lowest soil moisture simulated by the hydrological model is about 0.1 (cm^3 cm^{-3}), but the *in-situ* soil moisture measurements are even lower than the simulated estimates (<1.0 cm^3 cm^{-3}), which means that the hydrological model is less sensitive during the dry condition. We also compared the $\theta(h)$ functions derived by the dominant soil ID values (ID: 9, 14 for the sub-pixel 1, and ID: 9, 12 for the sub-pixel 3) for the individual sub-pixels with the observed $\theta(h)$ functions measured with the soil core samples collected at the LW 21 sites (shown in Fig. 4.10(b,c)). The derived $\theta(h)$ functions, which have silt (ID: 9) and silt loam (ID: 12) soils, in the sub-pixel $a_{1,3}$ matched well with the observed soil water retention curve, but the $\theta(h)$ functions of sub-pixel $a_{1,1}$ have a bias compared to the observation, especially for sandy clay loam (ID: 14) soil.

Pixel-based simulated ET estimates for the LW 21 site are illustrated in the Fig. 4.11. Simulated ET estimates (MBE: -4.378 ~ -3.630, R^2 is excluded, because of few datasets) with wheat crop (at the sub-pixels $a_{1,2}$, $a_{1,3}$, $a_{2,2}$, $a_{2,3}$, $a_{3,2}$, $a_{3,3}$) are higher than those of the grass cover (sub-pixels: $a_{1,1}$, $a_{2,1}$, $a_{3,1}$). On DOY 190 ET is relatively underestimated compared to the observations, because wheat crop was harvested at LW21 site on June 27 (DOY 178) and converted to bare ground. Overall, although DDA estimated soil moisture and ET at sub-pixel level have errors due to inherent weaknesses

of the adopted hydrological model (SWAP) and limited field measurements, the approach generally shows good performance for downscaling the remote sensing (ESTAR) soil moisture product.

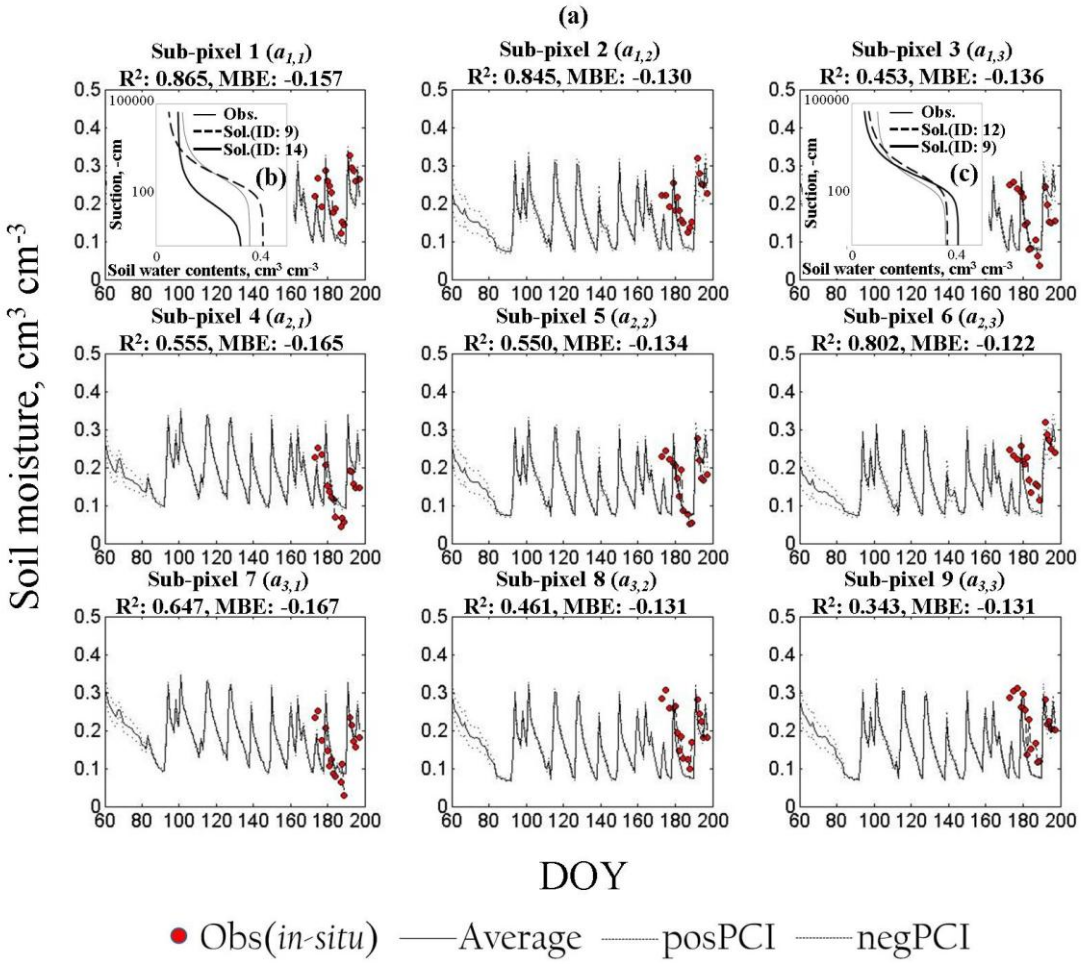

Figure 4.10: (a) Soil moisture dynamics (0-5 cm) of sub-pixels downscaled by the deterministic downscaling algorithm using the genetic algorithm (EMOGA) at the LW 21 site in Oklahoma, (b) $\theta(h)$ functions of the observation and solutions (soil ID: 9, 14), (c) $\theta(h)$ functions of the observation and solutions (soil ID: 9, 12)

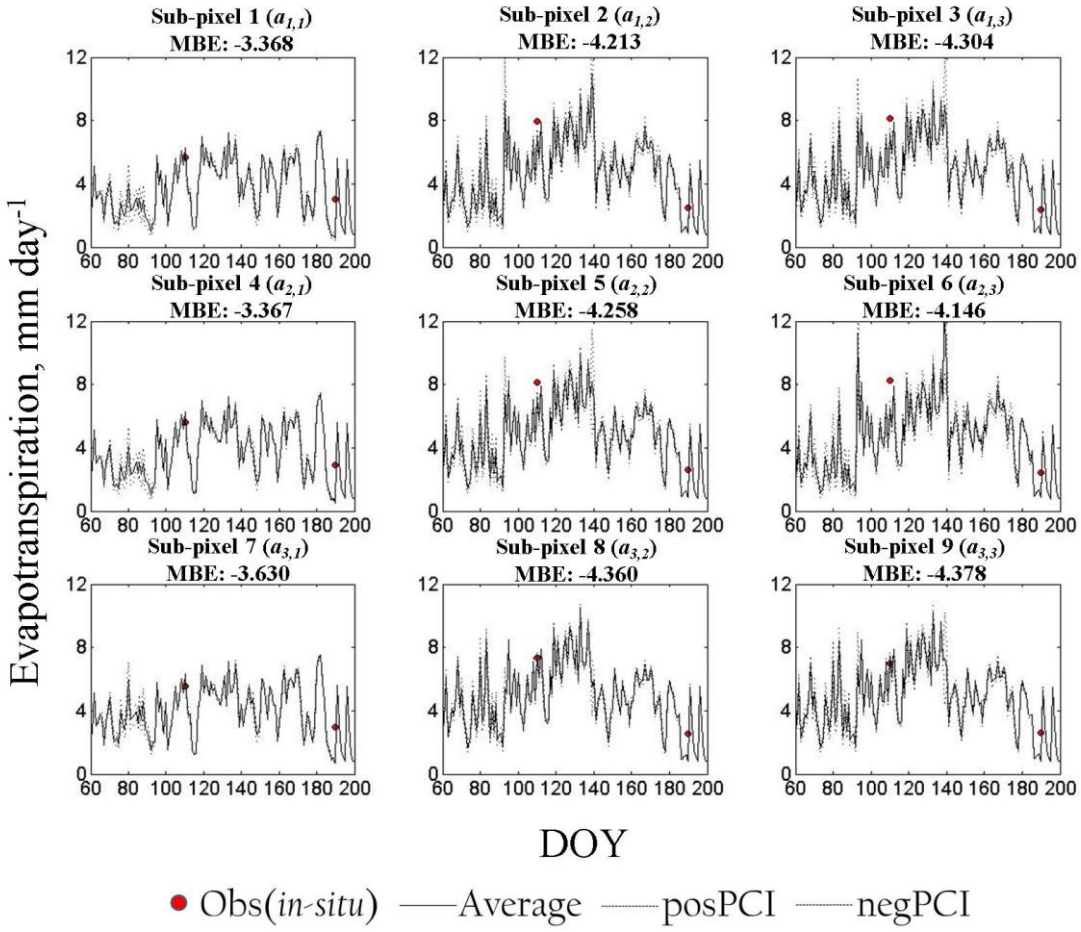

Figure 4.11: Evapotranspiration of sub-pixels by the deterministic downscaling algorithm using the genetic algorithm (EMOGA) at the LW 21 site in Oklahoma

The downscaled soil moisture dynamics (MBE: -0.203 ~ -0.169) at the LW 13 site showed more variations than those (MBE: -0.165 ~ -0.122) of the LW 21 site, although the correlations (R^2: 0.724 ~ 0.914) for LW13 are better than those (R^2: 0.343 ~ 0.865) for the LW 21 site in Fig. 4.12(a). The simulated soil moisture for the sub-pixels ($a_{1,1}$, $a_{1,2}$, $a_{1,3}$, $a_{2,1}$, $a_{2,3}$, $a_{3,1}$) are underestimated compared to the *in-situ* measurements ($a_{1,1}$, $a_{1,2}$, $a_{1,3}$, $a_{2,1}$, $a_{2,3}$, $a_{3,1}$) at LW13. We suggest that as the land surface at the LW 13 site has a significant slope compared to the flat terrain of LW 21 site (<2% slope) indicating that topography causes more uncertainties in downscaling of the RS soil moisture products than for a flat site. The derived $\theta(h)$ functions with the solutions (ID: 8-dominant, silt soil) of the sub-pixels (3 and 6) are compared with the observations in Fig. 4.12(b,c). The estimated $\theta(h)$ function in the sub-pixel 6 ($a_{1,3}$) somewhat corresponded well to the observed $\theta(h)$ function, but the water retention curve in the sub-pixel 3 ($a_{2,3}$) deviated from the observed curve. Mostly, the simulated ET estimates matched well with the pixel-based ET with the correlations (R^2: 0.368 ~ 0.990) and MBE (MBE: -4.652 ~ -4.171). But, pixel-based ET estimation of sub-pixel ($a_{3,1}$) on DOY 158 is considerably higher than other values estimated by the S-SEBI model (Fig. 4.13), because of a noise within the pixel and the presence of a small pond. The presence of pond is ignored during hydrologic model simulation. In summary, the DDA approach has uncertainties for extracting the soil ID values of sub-pixels within the RS soil moisture footprints. However, the $\theta(h)$ functions, soil moisture dynamics $\theta(t)$, and ET (t) estimates of individual sub-pixels matched reasonably well with the observations and demonstrated the applicability of our approach at the field-scale.

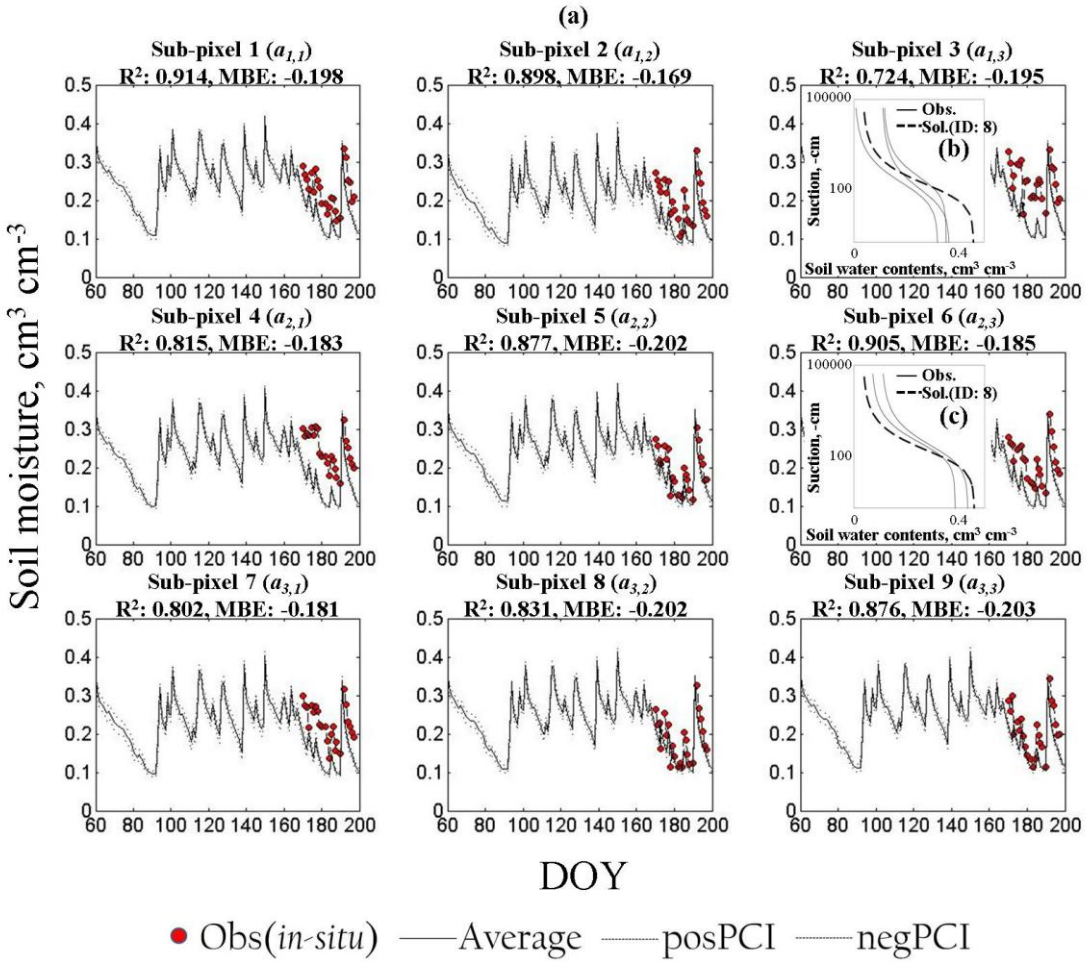

Figure 4.12: (a) Soil moisture dynamics (0-5 cm) of sub-pixels downscaled by the deterministic downscaling algorithm using the genetic algorithm (EMOGA) at the LW 13 site in Oklahoma, (b) $\theta(h)$ functions of the observation and solutions (soil ID: 8), (c) $\theta(h)$ functions of the observation and solutions (soil ID: 8)

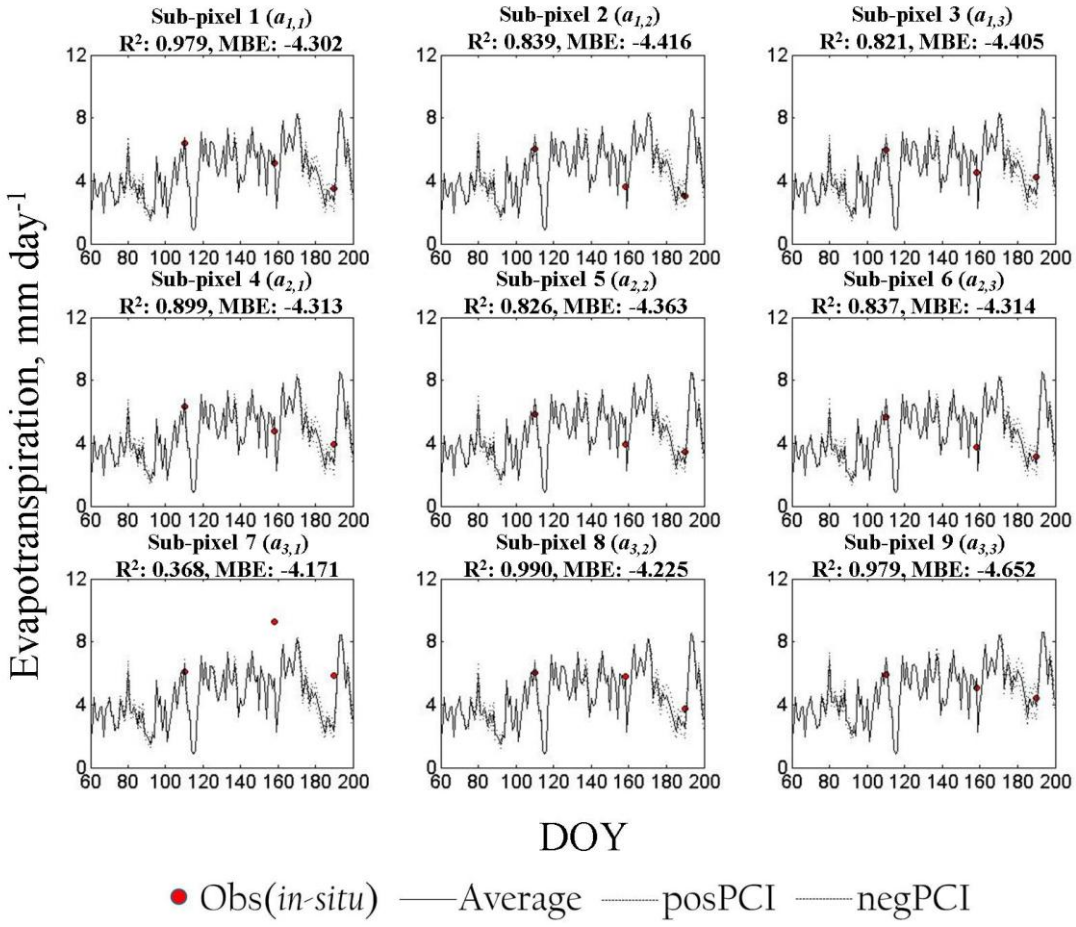

Figure 4.13: Evapotranspiration of sub-pixels by the deterministic downscaling algorithm using the genetic algorithm (EMOGA) at the LW 13 site in Oklahoma

4.5 Conclusions

We developed a deterministic downscaling algorithm using the ensemble multiple operator genetic algorithm (EMOGA) for estimating the (sub-pixel) finer-scale soil moisture from the remotely sensed (RS) soil moisture (SM) and evapotranspiration (ET) products based on the inverse modeling. We extracted the pixel-based soil ID values (representing soil textures) of sub-pixels within the RS pixel and simulated the long-term SM and ET dynamics through the hydrological model using the searched (soil ID) results of sub-pixels. Synthetic experiments were conducted under various (simple, relatively complex, and complex) land surface conditions with different vegetation covers (wheat, soybean, grass, and maize) and ground water tables (-100, -150, and -200 cm), respectively. Then, we validated the applicability of this approach with the *in-situ* soil moisture measurements and $\theta(h)$ curves derived by the soil hydraulic parameters obtained from the soil core samples collected at the field (LW 13 and 21) sites in Oklahoma.

The synthetic cases show the robustness of the algorithm for extracting the soil ID values of sub-pixels. Under the simple, relatively complex, and complex land surface conditions, this approach fairly identified the solutions of sub-pixels, although uncertainties were included in the derived solutions. Land covers have more impacts on the model performance than the presence of ground water tables. In the field experiments, the downscaled soil moisture estimates of sub-pixels (from ESTAR) show the moderate correlation (R^2: 0.724 ~ 0.914, MBE: -0.203 ~ -0.169 for the LW 13; R^2: 0.343 ~ 0.865, MBE: -0.165 ~ -0.122 for the LW 21) with the *in-situ* measurements.

Mostly, the DDA searched silt and silt loam soils (soil ID: 8, 9, 12) as dominant soils at the individual sub-pixels compared to the observations (predominantly silt loam) at the LW 13 and 21 sites. Although there exists uncertainties due to the non-uniqueness of solutions (e.g., similarities of soil hydraulic responses among the soil textures included in the physical soil texture database, co-linearity of covariates, inherent weakness of hydrological model structures, and errors in measurements and initial/boundary conditions) and few pixel-based ET measurements available, results show good performance of the approach. In other words soil moisture estimates downscaled by the deterministic downscaling algorithm matched well with the generated observations under synthetic conditions and field measurements. The good match of observed (field-observed soil hydraulic properties) and simulated (derived-soil ID values) $\theta(h)$ functions supports the robustness of our approach further in downscaling the RS products at the airborne or spaceborne footprint scales. On the basis of these findings, we suggest that the deterministic downscaling algorithm with the genetic algorithm (EMOGA) is useful for downscaling the remotely sensed soil moisture products at the spatio-temporal scales using the pixel-based evapotranspiration.

CHAPTER V

DEVELOPMENT OF DOWNSCALING AND UPSCALING ALGORITHMS FOR REMOTELY SENSED SOIL MOISTURE IN COMPLEX TERRAIN AT MULTIPLE SCALES

5.1 Synopsis

More recently spatial scaling algorithms have been developed to improve the availability of remotely sensed (RS) soil moisture for hydrologic applications. Existing approaches still have limitations, i.e., application in complex terrain, complexity of coupling down- and up-scaling schemes, etc. In this study, we developed a scaling (joint downscaling and upscaling) algorithm for remotely sensed and *in-situ* soil moisture measurements. Our new scheme can downscale RS soil moisture products as well as upscale *in-situ* measurements simultaneously across a topographically complex regional area. The approach is based on an inversion model using a genetic algorithm. Normalized digital elevation model (NDEM) and normalized difference vegetation index (NDVI) representing the heterogeneity of topography and vegetation covers were used to characterize the complexity of land surface. First, the approach quantified soil hydraulic parameters from RS and *in-situ* soil moisture data at multiple scales. Soil moisture predictions were derived by these estimated hydraulic parameters using the Soil-Water-Atmosphere-Plant (SWAP) model. As model simulated soil moisture predictions were generated for different NDEM and NDVI values across the spatial domain at a finer-scale (30 m $\times$ 30 m), downscaled and upscaled values were obtained at the scale of interest. The Little Washita watershed in Oklahoma was selected to validate this

methodology at the airborne- and satellite-footprint scales. New scaling approach performed well in several topographically complex footprints. The newly developed scaling (joint downscaling and upscaling) algorithm could improve the availability of RS soil moisture for agriculture and water resources management efficiently.

5.2 Introduction

Remote sensing (RS) techniques provide routine land surface soil moisture (SM) data at regional and global scale such as airborne-sensing (Electronically Scanned Thinned Array Radiometer, ESTAR [*Jackson et al.*, 1999], Polarimetric Scanning Radiometer, PSR [*Bindlish et al*, 2006]), and space-borne (Soil Moisture and Ocean Salinity, SMOS [*Kerr et al.*, 2001], Advanced Microwave Scanning Radiometer - Earth Observing System, AMSR-E [*Njoku*, 2008], and upcoming Soil Moisture Active and Passive, SMAP [http://smap.jpl.nasa.gov], among others). However, RS soil moisture data have a limitation in application at regional scales due to the scale discrepancy between the RS footprint size and model grid resolution [*Engman*, 1991; *Entekhabi et al*., 1999]. In order to overcome this drawback, down-/up-scaling algorithms are required to improve the quality of available RS soil moisture for agriculture and water resources management.

A few downscaling methods [*Das et al*., 2008b; *Kerr et al*., 2001; *Kim and Barros*, 2002; *Merlin et al*., 2005] have been explored for extracting fine-scale soil moisture information within a large RS pixel. Satellite-based soil moisture products were downscaled to obtain surface soil dielectric values approximating to soil water contents with a soil dielectric inversion model [*Crow et al.*, 2000]. An interpolation approach was

developed for passive microwave data based on fine-scale active microwave data [*Kim and Barros,* 2002; further *Das et al*., 2008b; *Das et al*., 2011]. A downscaling algorithm was developed by *Merlin et al.,* [2005] using fine-scale optical data during the SMOS mission for scaling down large-scale soil moisture footprints. However, downscaling schemes still have been limited in applications across complex landscapes at various scales.

Upscaling algorithms have been developed with considerable attention for extracting land surface information contained within RS soil moisture footprints by calibration of coarse-scale land surface models. In upscaling algorithms, *in-situ* soil moisture measurements at point-scale should be representative of the heterogeneity across the land surface comprising of various soil textures, vegetation covers, and topography [*Crow et al*., 2005; *Jana and Mohanty*, 2012a,b,c]. *De Lannoy et al*., [2007] conducted a data assimilation experiment for the upscaling approach indicating that soil moisture data measured from a site that can represent a study region improved the upscaled results. *Das et al.* [2008a] developed a Markov Chain Monte Carlo (MCMC) based data assimilation algorithm where moisture evolution was modeled at a spatial scale comparable to the AMSR-E soil moisture product. The basic hypothesis for their study was that the characterization of soil microwave emissions and their variations with space and time on soil surface within the remote sensing footprint can be represented by an ensemble of upscaled soil hydraulic parameters. These upscaled soil hydraulic parameters could somewhat represent the hydrologic complexities due to topography, vegetation, soil, and other landscape features within the RS footprint. However, *in-situ*

soil moisture measurements are necessarily required for evaluating the upscaling performance indicating that this approach is also limited in areas where observation information is not available.

From the above point of view, downscaling and upscaling algorithms have not only complementary features, but also retain their own limitations of application (due to the scale discrepancy, applications in complex terrain, requirement of *in-situ* soil moisture measurements etc.). Most of all, the use of both downscaling and upscaling approaches could reduce their drawbacks, but at the same time may require a large number of input parameters for hydrological models using different optimization schemes and cause complexity in its application. To our knowledge, no studies have been conducted with respect to down-/up-scaling approaches simultaneously with an easy application to soil moisture variability. Furthermore, topographic complexity of landscapes still has not been considered for downscaling RS soil moisture footprints at various scales.

In this study, we developed a new scaling algorithm for estimating finer-scale soil moisture in complex terrains at the airborne- and satellite-footprint scales. The approach uses a simulation-optimization scheme, Soil-Water-Atmosphere-Plant (SWAP) model with genetic algorithm (GA) based inverse modeling. The objectives are two-fold; 1) to develop a joint downscaling and upscaling algorithm for remotely sensed and *in-situ* soil moisture data and 2) to analyze the impacts of land surface conditions (topography and vegetation covers) in the model performance at multiple scales. This approach could provide an attractive alternative and easy application for improving the

availability of RS soil moisture products at different scales and efficient water resources management in the real world.

5.3 Methods and Materials

5.3.1 Conceptual Framework for the Joint Downscaling and Upscaling Algorithm

Topography [*Jana and Mohanty*, 2012a] and variable vegetation covers [*Clark and Arritt*, 1995] represented by elevation from digital elevation models (DEM) and normalized difference vegetation index (NDVI) respectively, significantly influence soil moisture dynamics near the land surface. The influence of DEM [*Jana et al*., 2008; *Jana and Mohanty*, 2012a,b,c] and NDVI [*Farrar et al*., 1994; *Liu and Kogan*, 1996; *Adegoke and Carleton*, 2002; *Wang et al*., 2007] on soil moisture can be used as physical controls for downscaling RS soil moisture as well as upscaling *in-situ* measurements. Thus, we hypothesize that the inclusion of topography (DEM) and vegetation (NDVI) in spatial scaling algorithms of soil moisture will lead to better estimates of soil moisture. Influence of topography and vegetation can be incorporated in the scaling algorithm by including the relative spatial distribution of elevation and vegetation cover in the domain.

Although the moisture redistribution process could be quite complex with multiple parallel surface and near-surface flow mechanisms contributing to soil moisture status in topographically complex landscapes, generally, surface runoff causes the soil moisture values at lower elevations to become higher than those at higher elevations under homogeneous land surface conditions [*Jana and Mohanty*, 2012a,b,c]. In a relative sense it may mean that DEM enables the routing of soil water across a three-dimensional

(3-D) topographically complex landscape and thus can be treated as a proxy for near-surface soil water status. In the proposed algorithm, normalized DEM (NDEM) values in Eq. (5.1) have been used as an indicator of the relative topographical position of a point. The NDEM values range from 0.0 to 1.0 with 0 representing the highest point on the slope (Fig. 5.1a). If soil moisture at a given elevation on the slope is known then the soil moisture at another elevation can be determined based on its NDEM.

$$\mathrm{NDEM}_{i,j} = \frac{\mathrm{DEM}_{i,j} - \mathrm{DEM}_{\max}}{\mathrm{DEM}_{\max} - \mathrm{DEM}_{\min}} \forall i \forall j \tag{5.1}$$

where, *i*: the row number of sub-pixels within a RS soil moisture pixel, *j*: the column number of sub-pixels within a RS soil moisture pixel, DEM: the digital elevation model (m), NDEM: the normalized DEM, DEM_{max}: the maximum DEM, and DEM_{min}: the minimum DEM within the RS pixel.

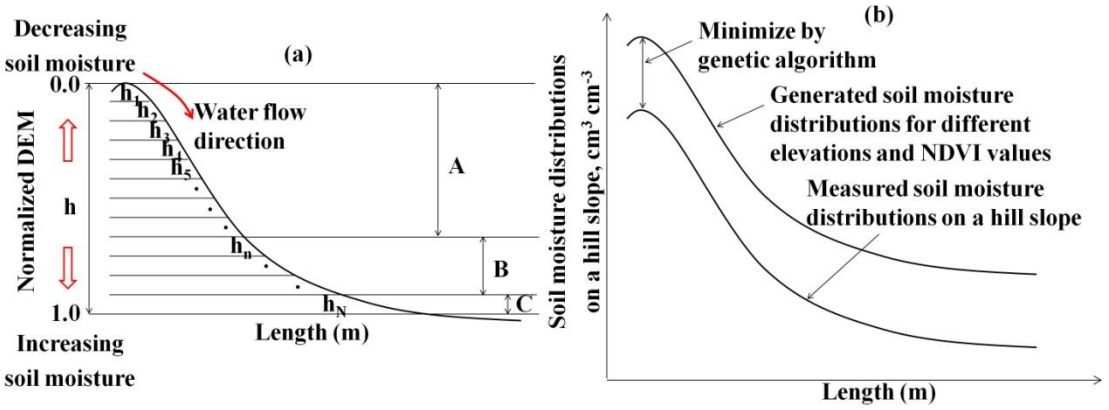

Figure 5.1: (a) Normalized digital elevation model (NDEM), (b) schematic of the downscaling and upscaling algorithm based on NDEM and NDVI using the genetic algorithm

Normalized difference vegetation index (NDVI) can be used to represent the degree of land surface that is covered by vegetation. It implies regions with high NDVI values (representing a higher vegetation cover) can hold more water than those with low NDVI [*Wang et al*., 2007]. This also indicates that root zone soil moisture has a significant correlation with NDVI under semi-arid and humid regions [*Wang et al*., 2007, *Sharma et al*., 2006, *Farrar et al*., 1994, *Liu and Kogan*, 1996]. Thus, we used fine-scale NDVI estimates (Eq. 5.2) obtained from RS images (provided from LANDSAT5-TM and LANDSAT7-ETM) for this approach. Considering daily variations of fine-scale NDVI is practically impossible due to the sparsely available time series of RS datasets. Furthermore, hydrological models already include physically-based crop growth models (i.e., WOFOST, *Supit et al.*, [1994]) that consider plant growth across the time series. For these reasons, we used NDVI taken for a single day during the simulation period and considered the relative impact of areal heterogeneity of vegetation covers at the RS pixel-scale. Analogous to NDEM, as soil moisture at a given NDVI is known then soil moisture at another NDVI can be obtained based on its NDVI.

$$\mathrm{NDVI}_{\boldsymbol{i,j}} = \frac{\mathrm{r}_{\mathrm{p4},\boldsymbol{i,j}} - \mathrm{r}_{\mathrm{p3},\boldsymbol{i,j}}}{\mathrm{r}_{\mathrm{p4},\boldsymbol{i,j}} + \mathrm{r}_{\mathrm{p3},\boldsymbol{i,j}}} \; \forall \boldsymbol{i} \forall \boldsymbol{j} \tag{5.2}$$

where, r_{p3} and r_{p4}: the RS images (band 3 and 4 from LANDSAT5-TM, 1997 and LANDSAT7-ETM, 2003, 30 m × 30 m).

For this approach, we adapted a near-surface (0-1/0-5 cm) soil moisture data assimilation scheme [*Ines and Mohanty*, 2008a] using a simulation-optimization framework. The simulation is carried out by running the one-dimensional (1-D)

hydrological model, Soil-Water-Atmosphere-Plant (SWAP, *van Dam et al.*, [1997]), and the outputs are optimized through a genetic algorithm (GA). The data assimilation scheme aims for quantifying soil water retention ($\theta(h)$) and hydraulic conductivity ($K(h)$) by optimizing 'effective' soil hydraulic parameters {$\mathbf{P}=(\alpha, n, \theta_{res}, \theta_{sat}, K_{sat})$}. Soil hydraulic parameters **k** (Eq. 5.3) are determined by minimizing the difference (*e*) between observed and simulated soil moisture estimates using the objective function ($Z(\mathbf{k})$) in Eq. (5.4). Genetic algorithms (GAs) are powerful search algorithms for solving the optimized solutions based on the precept of natural selection mechanism [*Holland,* 1975; *Goldberg,* 1989]. In this study, we used the modified-*micro*GA [*Ines and Mohanty,* 2008a] for this approach based on an inversion model. The GA has been successfully used in quantifying the soil hydraulic properties in the unsaturated zone. The readers are referred to *Ines and Mohanty,* [2008a,b, 2009] and *Shin et al*., [2012] for more details of the modified-*micro*GA.

A remote sensing footprint assigns an average soil moisture value to the entire pixel, irrespective of the heterogeneity of topography and vegetation covers the pixel is comprised of. For the downscaling algorithm, our proposed approach quantified effective soil hydraulic parameters of the RS soil moisture pixel. These estimated hydraulic parameters have been used to simulate pixel-scale soil moisture estimates ($\theta_p(\mathbf{k},t)$) using the SWAP model and that in turn were used to generate distributed (subpixel-scale) soil moisture estimates ($\hat{\theta}_{d,i,j}(\mathbf{k},t)$, Eq. 5.5) at different elevations and vegetation covers based on the NDEM and NDVI with the weighted formulation as in Eq. (5.4). The GA repeatedly estimates hydraulic parameters in the given generations

until the difference between the RS pixel and average of distributed (subpixel-scale) soil moisture predictions was minimized (Fig. 5.1b and Eq. 5.4). Thus distributed soil moisture estimates optimized by the GA for different elevations and heterogeneous vegetations across the land surface ($\hat{\theta}_{d,\boldsymbol{i,j}}$ (**k**,t)) are obtained.

For the upscaling algorithm, *in-situ* soil moisture data at a particular elevation (m) within a RS footprint were used. Analogous to the downscaling algorithm, this approach quantifies soil hydraulic parameters (**k**) from *in-situ* ($\theta^{\boldsymbol{in-situ}}(t)$) data and simulates soil moisture dynamics (θ_{in}(**k**,t)) with the SWAP model. *In-situ* data already incorporated the impacts of elevation (NDEM) and vegetation cover (NDVI) at field-scales. By using NDEM and NDVI information, we assumed that simulated soil moisture data (θ_{in}(**k**,t)) at the *in-situ* data sampling location can be mathematically solved as shown in Eq. (5.6). Note that the soil moisture θ_{ref}(**k**,t) defined as the reference soil moisture indicates that the effects of elevation and vegetation were mathematically excluded from simulated soil moisture data. As we extract the reference soil moisture, distributed soil moisture ($\hat{\theta}_{u,\boldsymbol{i,j}}$ (**k**,t)) for another sub-pixel can be determined by adding NDEM and NDVI to the reference value (Eq. 5.7). In order to assign different weights to NDEM and NDVI in the objective function (Eq. 5.4), we used the weighting factor (f) in Eqs. (5.5-5.6). A weighting factor of f=0 indicates that the entire weight is given to NDVI whereas a weighting factor of f=1 means that the entire weight is given to NDEM.

$$\mathbf{k} = \{\mathbf{P}, \lambda = 0.5\} \tag{5.3}$$

$$Z(\mathbf{k}) = e = \min \begin{cases} \{\frac{1}{T}\frac{1}{N}\frac{1}{M}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{M}((((|\theta^{RS}(t) - \hat{\theta}_{d,i,j}(k,t)|)))\} & \textit{for downscaling} \\ \{\frac{1}{T}\sum_{t=1}^{T}(|\theta^{\textit{in-situ}}(t) - \theta_{in}(k,t)|)\} & \textit{for upscaling} \end{cases} \quad (5.4)$$

$$\hat{\theta}_{d,\boldsymbol{i,j}}(k,t) = \theta_p(k,t) + f \times (\theta_p(k,t) \times |\,\mathrm{NDEM}_{i,j}\,|) + (1-f) \times (\theta_p(k,t) \times \mathrm{NDVI}_{\boldsymbol{i,j}})\ \forall t \quad (5.5)$$

$$\theta_{in}(k,t) = \theta_{ref}(k,t) + f \times \theta_{ref}(k,t) \times \mathrm{NDEM}_{ref} + (1-f) \times \theta_{ref}(k,t) \times \mathrm{NDVI}_{ref}\ \forall t \quad (5.6)$$

$$\hat{\theta}_{u,\boldsymbol{i,j}}(k,t) = \theta_{ref}(k,t) + \theta_{ref}(k,t) \times |\,\mathrm{NDEM}_{i,j}\,| + \theta_{ref}(k,t) \times \mathrm{NDVI}_{\boldsymbol{i,j}}\ \forall \boldsymbol{i} \forall \boldsymbol{j} \forall t \quad (5.7)$$

where, t: the (running) index for time, T: the time domain, M: the number of rows (*i*), N: the number of columns (*j*), θ(t): the observed soil moisture with the time (t), Z: the (minimizing) objective function, *e*: the difference between the observed and simulated results, f: the weighting factor ($0 \le f \le 1$), and DEM_{ref} and $NDVI_{ref}$: the reference elevation and vegetation (indicating NDEM and NDVI at the *in-situ* soil moisture sampling location), respectively. Note it is assumed that lateral flow between sub-pixels incorporated the DEM near the land surface. The fitness of searched solutions (representing the soil parameters) is assessed using the fitness function in Eq. (5.8). The Fitness(**k**) function is maximized as Z(**k**) is minimized.

$$\mathrm{Fitness}(\mathbf{k}) = \mathrm{maximize}[Z(\mathbf{k})]^{-1} \quad (5.8)$$

The Noah Land Surface Model (Noah LSM, *NCEP*, [2005]) and Community Land Model [CLM, *Oleson et al*., 2010] can also simulate vertical soil moisture dynamics in the soil profile as well as SWAP. These models have different model structures and parameters such as input parameters (climatic forcing datasets, soil textures, vegetation covers, initial and bottom boundary conditions, soil layering, etc.)

indicating that each model has its own performance capability and drawbacks. In order to verify our methodology, we additionally tested these models and analyzed the impacts of different model structures for the newly developed scaling algorithm. These hydrological models are robust-validated models under various meteorological and environmental conditions [*Wesseling and Kores*, 1998; *Droogers and Bastiaanssen*, 2002; *van Dam et al*, 1997; *Ahmad et al.*, 2002; *NCEP*, 2005; *Oleson et al*., 2010, among other]. The soil water flow along the soil profile can be described based on the Richards' equation. The SWAP, Noah LSM, and CLM models are well known in the hydro-climate modeling literature and, thus, not elaborated here. The search spaces of input parameters for three models were shown in Table 5.1.

In this study, we only considered the near-surface (0-1/0-5 cm) soil moisture data corresponding to the penetration depths of passive microwave C-/L-band sensors (6.9 and 1.4 GHz) in the unsaturated zone (representing that vertical soil water flow is dominant). For these reasons, it is assumed that the soil column is homogeneous in the model performance. The soil profile was discretized by 33 computational layers for the SWAP model. The soil column at the near-surface of 10 cm (1-10th layer) had intervals of 1 cm. And the sub-surface (11-20th and 20-32nd layers) 10-60 and 60-200 cm from the soil surface were discretized at the intervals of 5 cm and 10 cm (20 cm for the 33rd layer), respectively. The soil column (total depth: 200 cm) for the Noah LSM model was discretized by seven computational layers with thickness of 5, 5, 10, 10, 20, 30, and 120 cm. The CLM model has the soil profile discretized into 10 computational layers and the

layer thicknesses increase with depth such as 1.75, 2.76, 4.55, 7.5, 12.36, 20.38, 33.60, 55.39, 91.33, and 113.7 cm from the soil surface (total depth: 343 cm).

5.3.2 Description of Study Sites and Data

We tested the performance of this approach for the Little Washita (LW) watershed in Oklahoma at multi-scales (LW 13 and 21 sites for the airborne-scale and LW site for the satellite-scale) during the crop growing season (March 1 to July 31), as shown in Fig. 5.2. The *in-situ* and airborne-based soil moisture (0-5 cm) estimates were taken during the Southern Great Plains experiment (SGP97: June 18 - July 18, 1997) [*Mohanty et al*., 2002]. The *in-situ* near-surface (0-5 cm) soil moisture measurements (24 days for the LW 13 and 17 days for the LW 21, *Mohanty et al*., [2002]) were collected at the 49 (7×7) sampling points across the field sites. The RS pixel-based (0-5 cm) soil moisture products (airborne Electronically Scanned Thin Array Radiometer, ESTAR) were taken for 17 days at the LW 13 and 21 sites [*Jackson et al.*, 1999]. The satellite-based (Advanced Microwave Sensing Radiometer, AMSR-E, [*Njoku*, 2004]) soil moisture products (46 days) during 2003 were used for the LW site.

Table 5.1: Parameter constraints used in the genetic algorithm for various hydrological models

Numerical scenarios	Numerical scenarios	Parameters	Search space		No. of bit (L)	Binary(2^L)
			Min. values	Max. values		
Scenarios 1* and 2**	SWAP	a^a	0.006	0.033	5	2^5 (32)
		n^b	1.2	2.1	6	2^6 (64)
		$\theta_{res}{}^c$	0.03	0.16	7	2^7 (128)
		$\theta_{sat}{}^d$	0.3	0.55	5	2^5 (32)
		$K_{sat}{}^e$	1.84	130	10	2^{10} (1024)
Scenarios 1*	Noah LSM	maxsmc[f]	0.35	0.55	5	2^5 (32)
		psisat[g]	0.1	0.65	6	2^6 (64)
		satdk[h]	1.0E-06	1.0E-05	6	2^6 (64)
		B[i]	4	10	6	2^6 (64)
		quartz[j]	0.1	0.82	5	2^5 (32)
		vrcmin[k]	40	400	8	2^8 (256)
		rgl[l]	30	100	7	2^7 (128)
		Hs[m]	36	55	6	2^6 (64)
		Z0[n]	0.01	0.1	5	2^5 (32)
		LAI[o]	0.1	5	6	2^6 (64)
Scenarios 1*	CLM	WATSAT[p]	0.33	0.66	5	2^5 (32)
		HKSAT[q]	1.00E-05	0.1	8	2^8 (256)
		BCH[r]	3	10	6	2^6 (64)
		SMPSAT[s]	-750	-30	7	2^7 (128)
		WATDRY[t]	0.02	0.3	5	2^5 (32)

*airborne-scale; **satellite-scale

SWAP: [a]Empirical shape parameter (cm^{-1}), [b]Empirical shape parameter (-), [c]Residual water content (cm^3 cm^{-3}), [d]Saturated water content (cm^3 cm^{-3}) and [e]Saturated hydraulic conductivity (cm d^{-1}).

Noah LSM: [f]Maximum volumetric soil moisture (m^3 m^{-3}); [g]Saturated soil matric potential (m m^{-1}); [h]Saturated soil hydraulic conductiviey (m s^{-1}); [i]Clapp-Hornberger b parameter; [j]Quartz content; [k]Minimal stomatal resistance (s m^{-1}); [l]Parameter used in solar radiation term of canopy resistance; [m]Parameter used in vapor pressure deficit term of canopy resistance; [n]Roughness length (m); [o]Leaf area index.

CLM: [p]Volumetric soil water content at saturation; [q]Saturated soil hydraulic conductivity (mms^{-1}); [r]Clapp-Hornberger b parameter; [s]Saturated soil matric potential (mm mm^{-1}); [t]Soil water content

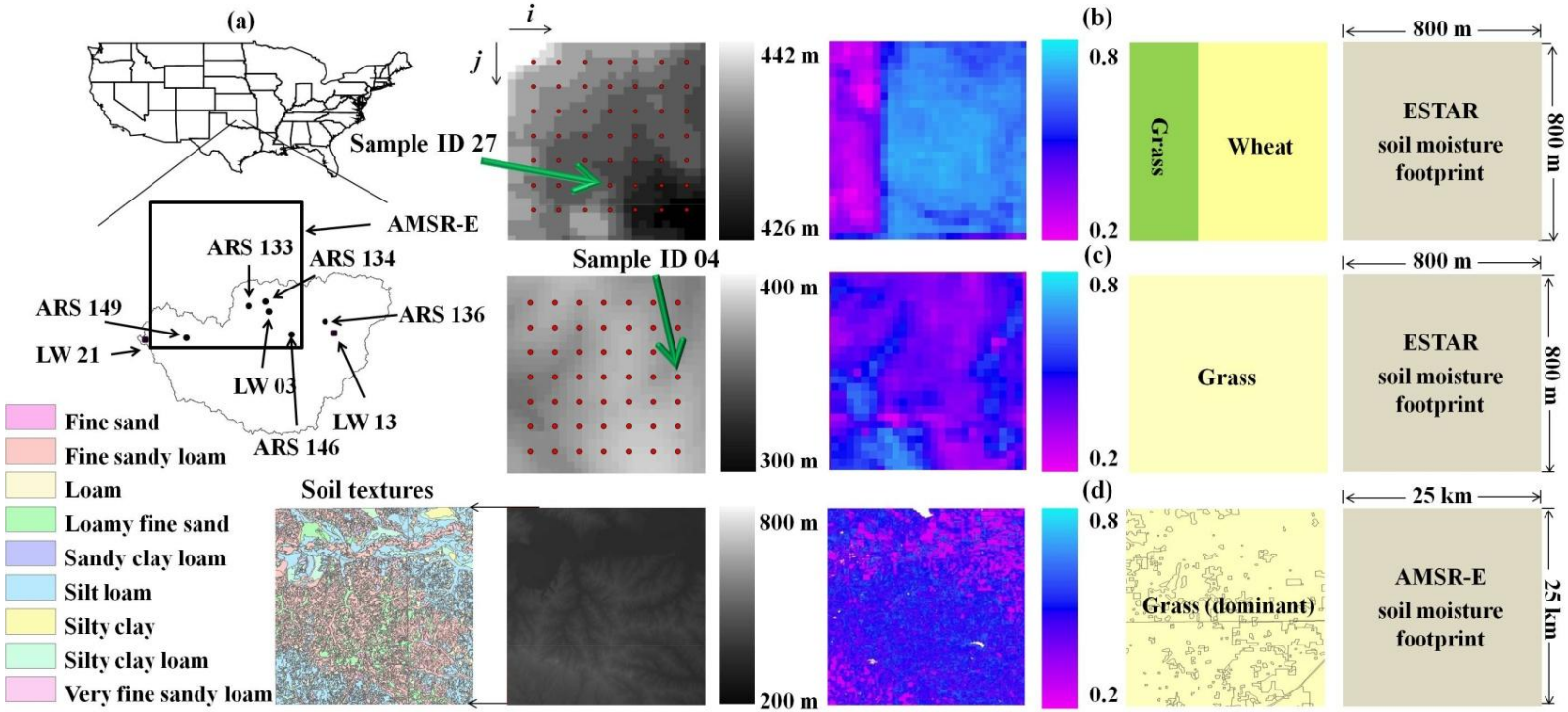

Figure 5.2: Study area; (a) Little Washita (LW) watershed in Oklahoma, (b) LW 21, (c) LW 13, and (d) LW sites including the digital elevations (49 *in-situ* soil moisture measuring points), NDVI, vegetation covers, and remotely sensed (ESTAR and AMSR-E) soil moisture footprints

The LW 13 and 21 sites are composed of a (predominant) silt loam soil [*Mohanty et al*., 2002]. The LW 21 site has two different vegetation covers (winter wheat for the two thirds and the other third for the short native grass). The wheat cover area has a flat slope (less 1%) and the grass cover region is relatively rolled (3-12%). The LW 13 site has a hill slope with the grass cover. The LW site at the satellite-scale has various soil textures with different vegetation covers (with grass being dominant). The land surface information for the LW study site is shown in Fig. 5.2d. Daily weather datasets such as precipitation, wind speed, maximum and minimum temperature, and solar radiation for the input parameters of hydrological (SWAP, Noah LSM, and CLM) models were collected from the USDA-Agricultural Research Service (ARS 136 for the LW 13 site, ARS 149 for the LW 21 site, and ARS 133, 134, 146, and 149 for the LW site; http://ars.mesonet.org/) Micronet weather station in Oklahoma. We used the digital elevation model (http://seamless.usgs.gov/website/seamless/viewer.htm) and NDVI obtained from LANDSAT5-TM (taken on DOY 110, 1997) and LANDSAT7-ETM (taken on DOY 119, 2003) images with a finer resolution (30 m×30 m) at the study sites (Table 5.2). The field-observed hydraulic parameters [as shown in Table 5.3 from *Mohanty et al*., 2002] obtained from the soil core samples at the soil depth of 3-9 cm were used for validation.

Table 5.2: Description of the fine-scale LANDSAT images

Image Character	Values	
Sensor	LANDSAT5-TM	LANDSAT7-ETM
Path/Row	28/36	28/36
Acquisition date (time)	20 April 1997 (DOY 110)	29 April 2003 (DOY 119)
Reference system	UTM-14N	UTM-14N
Resolution	Band 3 and 4 (30m×30m)	Band 3 and 4 (30m×30m)

Table 5.3: Field-observed soil hydraulic properties derived by the soil core samples collected at the LW 03, 13, and 21 sites in Oklahoma

Sites	Soil depth	Soil samples	α^a	n^a	$\theta_{res}{}^a$	$\theta_{sat}{}^a$	$K_{sat}{}^a$
LW 03	3-9 cm	1	0.010	1.787	0.039	0.370	46.051
		2	0.012	1.381	0.089	0.341	4.216
		3	0.014	2.085	0.030	0.380	95.645
		4	0.014	2.057	0.046	0.361	88.128
LW 13	3-9 cm	1	0.012	1.262	0.106	0.435	129.427
		2	0.009	1.387	0.068	0.387	22.550
LW 21	3-9 cm	1	0.006	1.581	0.117	0.429	31.795
		2	0.009	1.734	0.115	0.432	17.885

[a]Field observations [*Mohanty et al*., 2002]

5.3.3 Field Experiment

We conducted the field studies for testing the scaling (joint downscaling and upscaling) algorithm at multi-scales. The field experiments were comprised of two scenarios under the rain-fed conditions: i) scenario 1: airborne-scale and ii) scenario 2: satellite-scale, respectively.

In scenario 1, we tested various weighting factors (f) ranging from 0.1 to 0.9 at the LW 13 site for evaluating the impacts of topography (NDEM) and vegetation covers (NDVI) in the model performance and selected the one providing the best fitness. With the selected weighting factor, the RS (ESTAR) and *in-situ* soil moisture data were down- and up-scaled by this approach at the study sites. Then, downscaled and upscaled results were validated with the *in-situ* soil moisture data. The soil moisture measurements of sub-pixels at the LW 13 (sample ID 04 with the elevation 372 m) and 21 (sample ID 27 with the elevation 437 m) sites were upscaled (Fig. 5.2b-c). Additionally, Noah LSM and CLM models were used to test our approach in only upscaling *in-situ* soil moisture data at the LW 13 site based on the selected weighting factor. Three hydrological models have different model settings for the initial and bottom boundary conditions. In order to compare the impacts of different model structures, we assumed that the bottom boundary condition was governed under free drainage (Table 5.4). Note that the initial soil moisture conditions for Noah LSM and CLM were determined through spinning up of models in the preceded term before the simulation period. Also, we only used the fitted parameters searched by the GA for each model in the comparison study.

Table 5.4: Modeling conditions of the genetic algorithm and various hydrological models

	Field experiments		
	SWAP	Noah LSM	CLM
GA parameters:			
No. population	10	10	10
No. seed	-1000	-1000	-1000
No. generation	500	500	500
P_{creep}	0.05	0.05	0.05
P_{cross}	0.5	0.5	0.5
P_{mutate}	0.05	0.05	0.05
Modeling Conditions:			
Top boundary	Time dependent flux/head	Time dependent flux/head	Time dependent flux/head
Bottom boundary	Free drainage (FD), ground water (GW) table		
Initial conditions	depths (-100, -150, -200 cm) h(z,t=0)=-200, -300, -500 cm for FD Equilibrium with the bottom boundary conditions for GW	Initial soil moisture condition was determined through spinning up of model	Initial soil moisture condition was determined through spinning up of model

For the scenario 2, we downscaled the satellite-scale (AMSR-E) soil moisture (0-1 cm) products across the LW site using the joint NDEM and NDVI. Downscaled results were validated using the fine-scale (30 m×30 m) temperature (°C) obtained from LANDSAT7-ETM data, because of no *in-situ* soil moisture data available. The LW 13 and 21 sites have relatively small areas (800 m×800 m) and predominant silt loam soils [*Mohanty et al*., 2002] indicating that a silt loam soil can represent these study sites. Therefore, this approach can upscale *in-situ* soil moisture data measured from a single location across each study site at the airborne-scale. However, the upscaling approach at the satellite-scale needs *in-situ* datasets measured from multiple locations representing a

large area to be verified. In this study, the *in-situ* data in a single location (LW 03 site) is only available for validation within a large-scale RS pixel. Hence, the upscaling algorithm at the satellite-scale was excluded and will be verified with available *in-situ* measurements in a future work. Additionally, we compared water retention curves ($\theta(h)$) derived by the field-observed (Table 5.3) and estimated soil hydraulic parameters to support the robustness of this approach at the airborne- and satellite-scales.

Generally, the initial and bottom boundary conditions significantly influence the performance of physically-based hydrological model in estimating hydraulic parameters, but these conditions were unknown at field-scales. We tested the impacts of various initial and bottom boundary conditions in modeling under free-drainage and presence of ground water table depths (-200, -150, and -100 from the soil surface) in Table 5.4. Then, we conducted uncertainty analysis under the different initial and bottom boundary conditions. The solutions (maximum fitness) of various conditions searched by the GA were averaged and classified them above and below the average. We only selected the solutions (above average) and calculated the statistics (correlation-R^2 and root mean square error-RMSE) of chosen solutions in Eqs. (5.9-5.10). The variability of solutions selected above average was shown as the 95 percent confidence interval ($\pm$ 95PCI) in Eq. (5.11).

$$R^2 = \frac{\boldsymbol{Cov}(X,Y)}{\boldsymbol{Var}(X)\boldsymbol{Var}(Y)} \tag{5.9}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}(X_t - Y_t)^2}{T}} \tag{5.10}$$

$$\pm 95PCI = Avg. \pm 1.96 \times SD \quad (5.11)$$

where R^2: the Pearson's correlation, X: the observations, Y: the output results, *Cov*: the covariance of X and Y, *Var*: the variance of X or Y, +95PCI: the upper boundary, –95PCI: the lower boundary, Avg.: the average of output results, and SD: the standard deviation of output results.

5.4 Results and Discussion

5.4.1 Scenario 1: Airborne-Scale

We tested the impacts of various weighting factors ($0.1 \leq f \leq 0.9$) for the model performance at the LW 13 site as shown in Fig. 5.3. Overall, the correlations (R^2: 0.408~0.495) with various weighting factors were not considerably changed. However, as the more weights were given to the NDEM (f=0.9) or NDVI (f=0.1), the upscaled results (RMSE: 0.097 and RMSE: 0.091) had more uncertainties compared to the *in-situ* data. Especially, as the weights were given to the NDEM, the distributed soil moisture predictions were similarly shown across topography with more uncertainties. The results (R^2: 0.489 and RMSE: 0.087) under the joint NDEM and NDVI (f=0.5) matched the *in-situ* data with good correlation and reduced uncertainties. Thus, we selected the weighting factor of f=0.5 for the downscaling and upscaling approaches.

Figure 5.4 shows the impacts of various initial and bottom boundary conditions for the model performance at the LW 13 site. We compared the averaged near-surface soil moisture predictions of all sub-pixels downscaled by this approach with the observations (ESTAR and *in-situ* measurement).

The results have large variations in modeling along the time series. Since the rainfall events were generated, the near-surface soil moisture predictions under free-drainage were considerably increased up compared to those of ground water tables, especially with the initial condition of -500 cm (from the soil surface). As the variations became stable during the dry days, the soil moisture dynamics with the presence of shallow ground water tables tend to gradually increase more than those of the free-drainage conditions. It may indicate that soils near the surface absorbed rainfall water quickly into the soil matrix under the free-drainage conditions. On the other hand, when the bottom boundary condition is governed by the ground water table, moisture is constantly delivered from ground water to the root zone by the capillary upward flow phenomena and root activities.

Generally, the soil water content near the soil surface is significantly affected by the atmospheric forcings at field-scale. However, these findings showed that the initial and bottom boundary conditions also have a large impact on soil moisture estimates as much as the weather conditions.

Therefore, we used the statistics (average and ±95 PCI) of derived soil moisture dynamics under various initial and bottom boundary conditions for uncertainty analysis at the spatio-temporal scales.

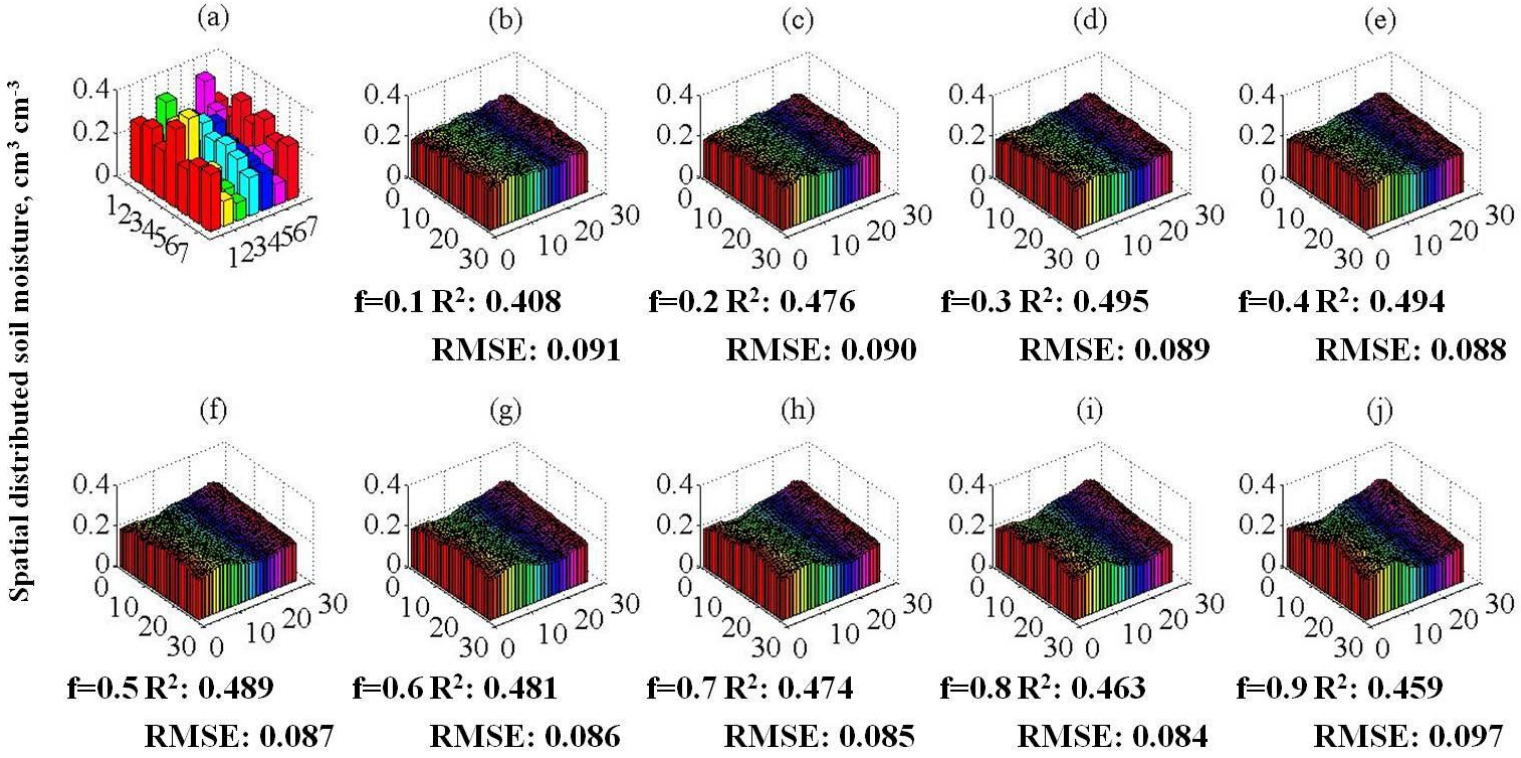

Figure 5.3: (a) *In-situ* and (b-j) upscaled near-surface (0-5 cm) soil moisture with the various weighting factors (0.1≤f≤0.9) on DOY 178 at the LW 13 site

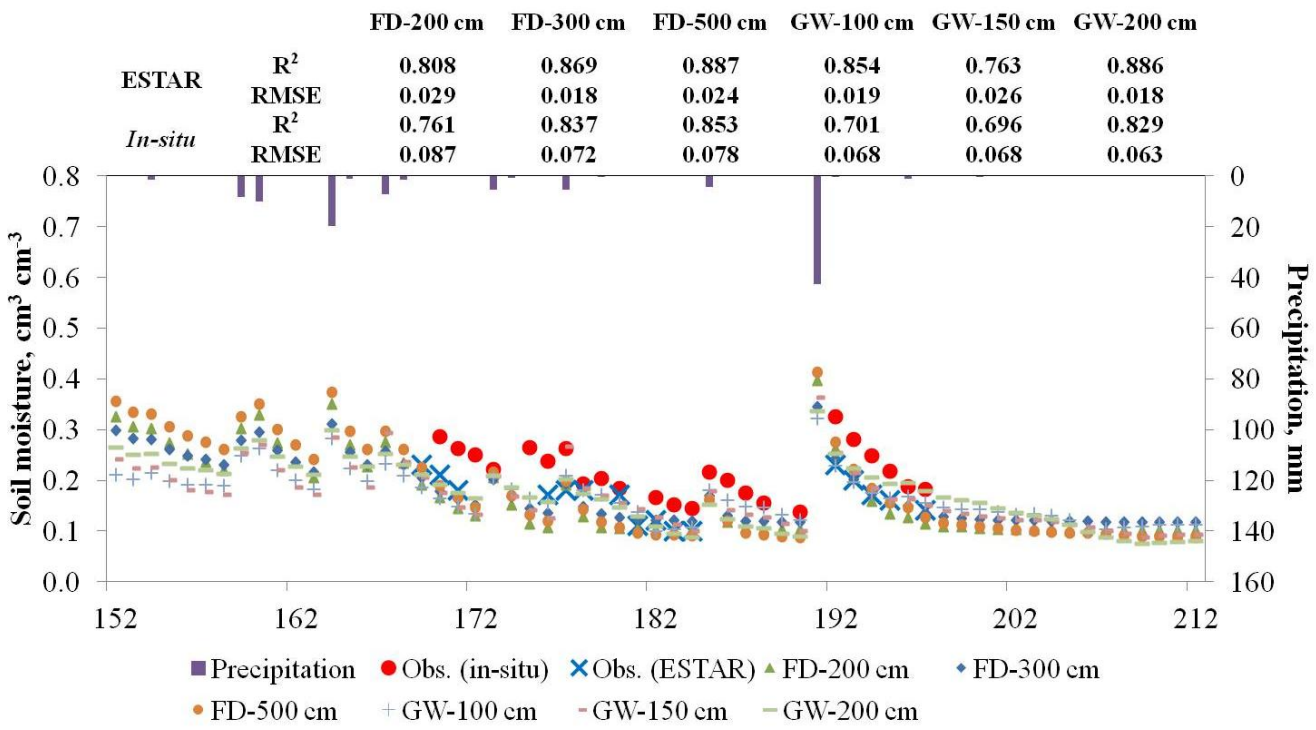

Figure 5.4: The observed (ESTAR and *in-situ*) and averaged (for the downscaled sub-pixels) near-surface (0-5 cm) soil moisture dynamics at the time series under various initial and bottom boundary conditions

FD: free-drainage with the initial condition of -500, -300, and -200 cm from the soil surface

GW: ground water tables of -200, -150, and -100 cm (the initial conditions were in equilibrium with the GW)

Figure 5.5 presents the measured (at the 49 measuring points) and predicted (downscaled and upscaled) near-surface (0-5 cm) soil moisture estimates on DOY 178 and 182 for the LW 13 site. We showed the results of two days only for the sake of brevity. Mostly, the downscaled and upscaled near-surface soil moisture estimates (R^2: 0.307~0.500 and RMSE: 0.064~0.102) matched the *in-situ* measurements across a hill slope, although some uncertainties exist. However, when we compared the *in-situ* measurements on DOY 178 to those on DOY 182, the measured soil moisture estimates were not matched exactly across the land surface (topographically). Soil moisture values could be different based on the rainfall amounts, but the spatial distributions of soil moisture on DOY 178 and 182 would be similarly shown at least. It indicated that uncertainties between the estimated results and *in-situ* soil moisture data might be due to the measurement errors (i.e., determination of sampling points near plants, operator errors, etc.). The downscaled (R^2: 0.406 and RMSE: 0.040) and upscaled (R^2: 0.406 and RMSE: 0.074) soil moisture estimates on DOY 173 at the LW 21 site matched well with the measurements (Fig. 5.6), but the results (R^2: 0.204 and RMSE: 0.064 for the downscaling, R^2: 0.204 and RMSE: 0.091 for the upscaling) on DOY 180 were less identifiable. As shown in the results of LW 13 site, the *in-situ* data at the sub-pixels (i=1, j=5,…,7) on DOY 173 and 180 even had uncertainties each other indicating that the relatively low statistics of results on DOY 180 might be affected by the measuring errors. When we considered that the LW 21 site has the grass (i=1,…,3, j=1,…,7) and wheat (i=4,…,7, j=1,…,7) covers as shown in Fig. 5.2b, the downscaled and upscaled soil moisture estimates were similarly distributed with both topography and vegetations.

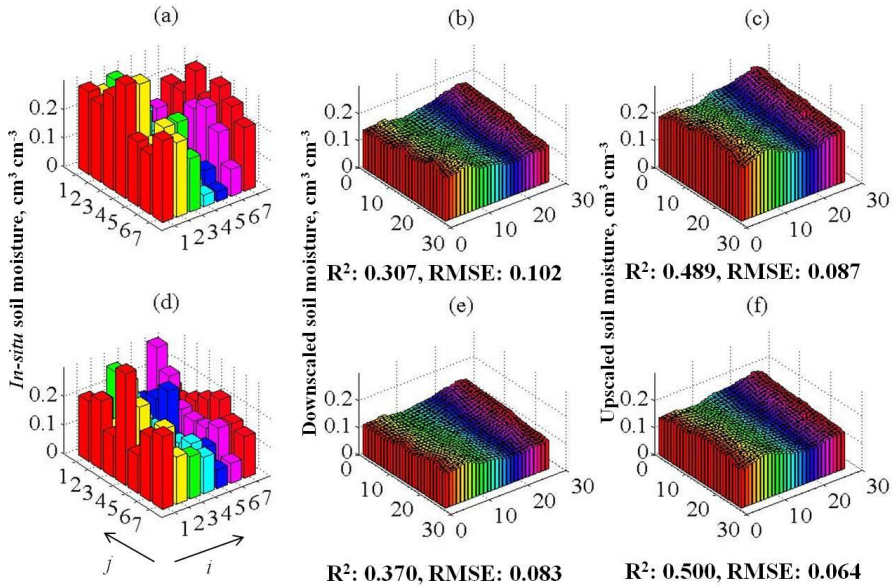

Figure 5.5: *In-situ*, downscaled, and upscaled near-surface (0-5 cm) soil moisture; (a-c) DOY 178 and (d-f) DOY 182 on a hill slope at the LW 13 site

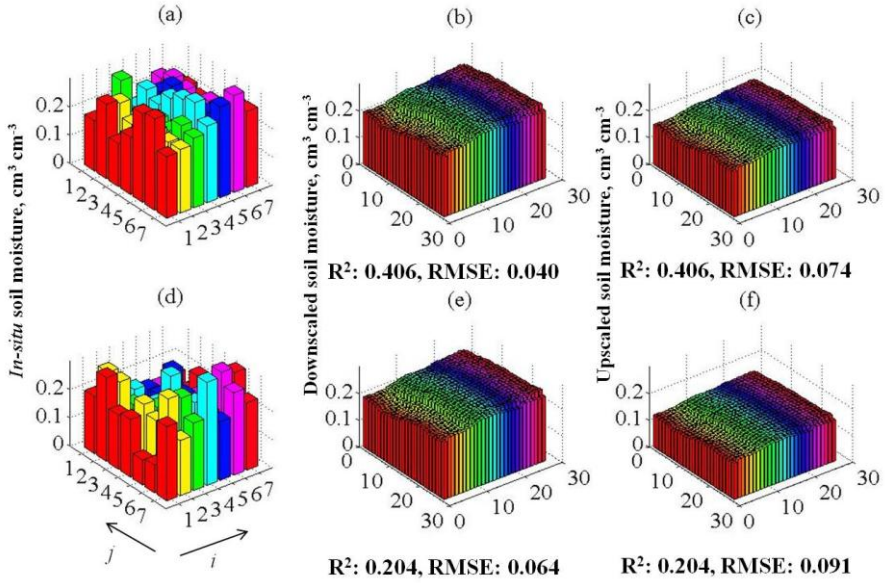

Figure 5.6: *In-situ*, downscaled, and upscaled near-surface (0-5 cm) soil moisture; (a-c) DOY 173 and (d-f) DOY 180 on a hill slope at the LW 21 site

We compared the averaged near-surface (0-5 cm) soil moisture dynamics of all sub-pixels downscaled from the ESTAR products with the *in-situ* measurements at the LW 13 site along the time series (Fig. 5.7). The results were only shown from June 1 to July 31 comparable to the *in-situ* data. The results fairly matched the ESTAR products with the correlation (R^2: 0.881) and root mean square error (RMSE: 0.019). However, the downscaled results (R^2: 0.836 and RMSE: 0.072) were underestimated when compared to the *in-situ* measurements. The RS pixel-based soil moisture products tend to be underestimated from the *in-situ* measurements on a hill slope. This tendency contributed to the underestimation of downscaled soil moisture estimates. We derived the water retention curves θ(h) using the searched (Table 5.5) and field-observed (Table 5.2) hydraulic parameters and validated our approach. The obtained θ(h) curves have some uncertainties (Fig. 5.7b), but the average curve matched the observations.

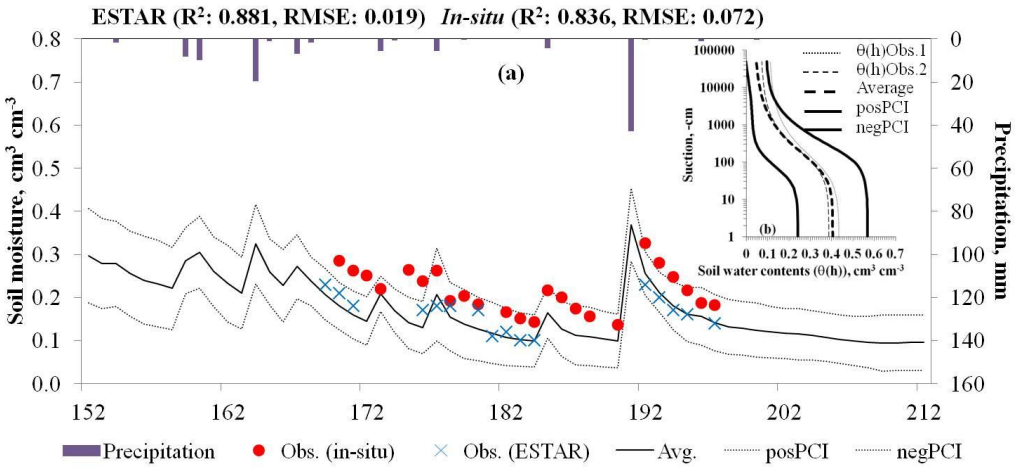

Figure 5.7: (a) The precipitation, remotely sensed (ESTAR), *in-situ*, and average (for the downscaled sub-pixels) near-surface (0-5 cm) soil moisture dynamics at the LW 13 site and (b) comparison of the field-observed and estimated curves (θ(h))

Table 5.5: The effective soil hydraulic properties derived by the downscaling and upscaling algorithms at the Little Washita watershed in Oklahoma

Scaling	Soil hydraulic parameters	Scenario 1 (Airborne-scale)				Scenario 2 (Satellite-scale)	
		LW 13		LW 21		LW	
		Average	SD	Average	SD	Average	SD
Down*	α	0.012	0.007	0.022	0.011	0.015	0.010
	n	1.655	0.197	1.948	0.205	1.999	0.098
	θ_{res}	0.033	0.031	0.030	0.014	0.068	0.006
	θ_{sat}	0.405	0.084	0.376	0.047	0.414	0.073
	K_{sat}	111.828	19.191	72.586	35.800	94.983	44.038
Up**	α	0.012	0.008	0.024	0.010	-	-
	n	1.572	0.158	2.024	0.141	-	-
	θ_{res}	0.043	0.039	0.059	0.060	-	-
	θ_{sat}	0.494	0.063	0.368	0.031	-	-
	K_{sat}	80.525	29.471	78.114	43.641	-	-

*Pixel-scale soil hydraulic parameters derived from a remotely sensed soil moisture footprint by our proposed approach.

**Soil hydraulic parameters derived from *in-situ* soil moisture data measured at the sub-pixels (sample ID 04 with the elevation 372 m for the LW 13 site and sample ID 27 with the elevation 437 m for the LW 21 site).

Figure 5.8 illustrated the results of upscaling performance at the LW 13 site. The simulated soil moisture dynamics (R^2: 0.918 and RMSE: 0.042 for the ESTAR and R^2: 0.880 and RMSE: 0.027 for the *in-situ*) identified well with the *in-situ* measurements while the derived $\theta(h)$ curves (Fig. 5.8b) had small uncertainties. This result indicated that the upscaling algorithm is more suitable for representing the soil moisture estimates on a hill slope.

In the LW 21 site, the downscaled soil moisture dynamics identified well with the ESTAR soil moisture products (R^2: 0.903 and RMSE: 0.035) and *in-situ* measurements (R^2: 0.706 and RMSE: 0.084) in Fig. 5.9. The *in-situ* datasets during DOY 173 to 177 were not matched with the trend of ESTAR soil moisture products.

When we considered that there were no rainfall events for three days since it rained on DOY 174 (6.7 mm), the relatively high soil moisture value (0.228 cm^3 cm^{-3}) on DOY 177 may be caused by the measuring errors as mentioned above. The estimated θ(h) curves were also underestimated in the range of uncertainties (±95 PCI) when compared to the field observations in Fig. 5.10b. However, these results still show good match with the RS products and *in-situ* measurements. In the upscaling algorithm, the estimated soil moisture (R^2: 0.941 and RMSE: 0.042 for the ESTAR and R^2: 0.742 and RMSE: 0.083 for the *in-situ*) also matched well with the ESTAR and *in-situ* data as shown in the downscaled results (Fig. 5.10), although the estimated curves (θ(h)) were biased compared to the observations.

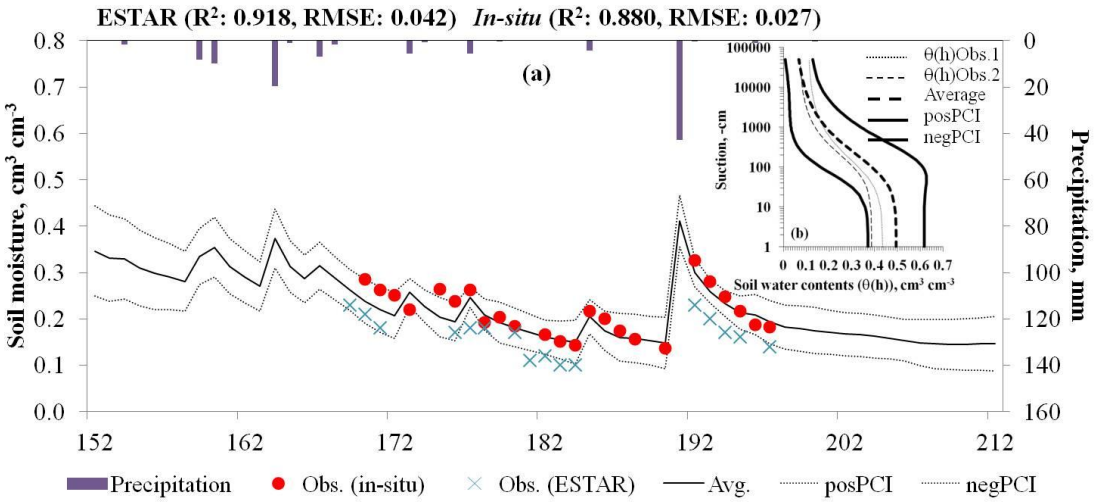

Figure 5.8: (a) The precipitation, remotely sensed (ESTAR), *in-situ*, and average (for the upscaled sub-pixels) near-surface (0-5 cm) soil moisture dynamics at the LW 13 site and (b) comparison of the field-observed and estimated curves (θ(h))

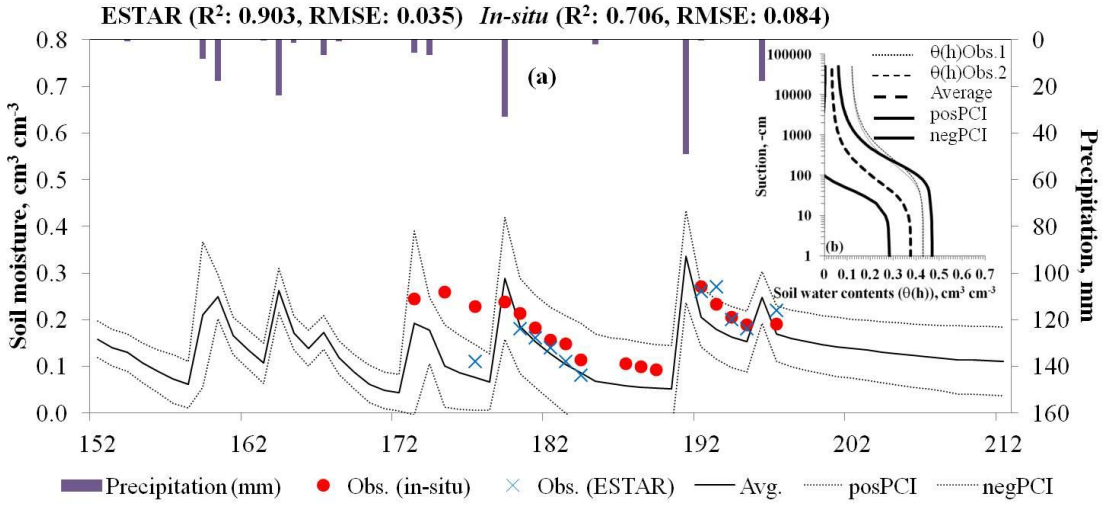

Figure 5.9: (a) The precipitation, remotely sensed (ESTAR), *in-situ*, and average (for the downscaled sub-pixels) near-surface (0-5 cm) soil moisture dynamics at the LW 21 site and (b) comparison of the field-observed and estimated curves (θ(h))

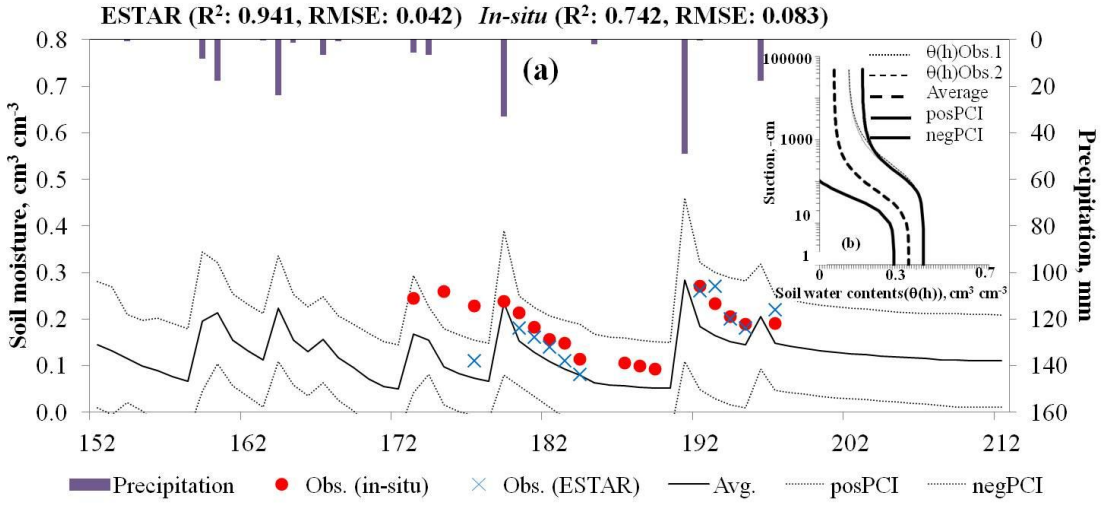

Figure 5.10: (a) The precipitation, remotely sensed (ESTAR), *in-situ*, and average (for the upscaled sub-pixels) near-surface (0-5 cm) soil moisture dynamics at the LW 21 site and (b) comparison of the field-observed and estimated curves (θ(h))

Additionally, the RS soil moisture products taken across a relatively flat topography (LW 21 site) show the similar trends (soil moisture quantities and pattern along the time series) with the *in-situ* measurements. As mentioned in the results of LW 13 site, the RS soil moisture datasets over a hill slope had relatively lower values than the *in-situ* data indicating that topography significantly affects the quality of RS products. Consequentially, these underestimations propagate to the downscaled soil moisture estimates directly. But our approach proposed in this study can improve a drawback of RS products in topographically complex terrains by comparing downscaled and upscaled soil moisture estimates simultaneously. Thus, these findings support the robustness of this approach at the field-scales.

Fig. 5.11 showed the upscaled (average for all sub-pixels) soil moisture dynamics using various hydrological models with the selected weighting factor (f=0.5) at the LW 13 site along the time series. Note that the estimated parameters for each model were not shown, because these models have different parameters (incomparable between models) for simulating soil moisture dynamics. Overall, the upscaled soil moisture dynamics estimated by the SWAP (R^2: 0.913 and RMSE: 0.029), Noah LSM (R^2: 0.877 and RMSE: 0.058), and CLM (R^2: 0.927 and RMSE: 0.051) models showed good match with the measurements along the time scenarios whereas the results of Noah and CLM models were overestimated. In the given condition of this study, the SWAP model provided slightly better predicted soil moisture dynamics than others for the upscaling approach.

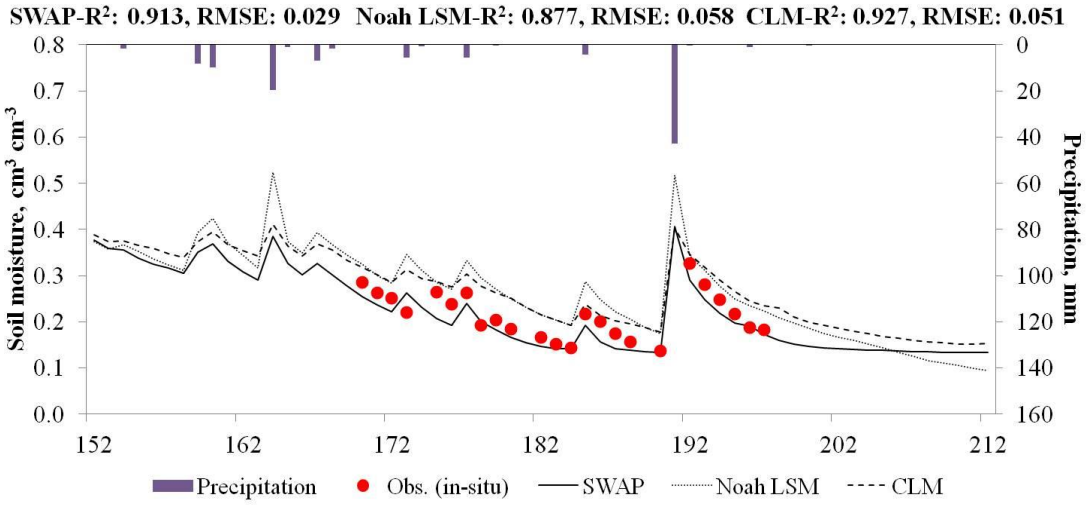

Figure 5.11: The precipitation, *in-situ*, and average (for the upscaled sub-pixels) near-surface (0-5 cm) soil moisture dynamics for various hydrological models (SWAP, Noah LSM, and CLM) at the LW 13 site

5.4.2 Scenario 2: Satellite-Scale

In the scenario 2, we downscaled the satellite-based (AMSR-E, 25 km×25 km) soil moisture products at the LW site. The downscaled near-surface (0-1 cm) soil moisture estimates on DOY 119 were validated with the pixel-based (LANDSAT7-ETM) temperature (30 m×30 m, °C) at the spatial coverage in Fig. 5.12. The simulated soil moisture estimates at the sub-pixels with relatively low temperature were higher than those, which have high temperature. Because moisture at the sub-pixels with high temperature dries quickly due to active generations of soil evaporation near the soil surface. Overall, the downscaled soil moisture distributions (R^2: 0.627 on DOY 119) fairly matched the fine-scale distributed temperature. Also, the results corresponded well to the NDEM and NDVI across the study site. As the elevation becomes lower as shown

in the lower corner of the left side, the downscaled soil moisture estimates were increased along the lower elevations (topography). Further, although the elevations were relatively lower than other regions in the upper corner on the right side, these sub-pixels have various (high and low) soil moisture ranges indicating that soil moisture estimates at the sub-pixels were also affected by vegetation covers (NDVI). It demonstrated that this approach could map the near-surface soil moisture estimates based on the NDEM and NDIV at the satellite-scale, although this approach needs to be verified with the *in-situ* measurements in a future study.

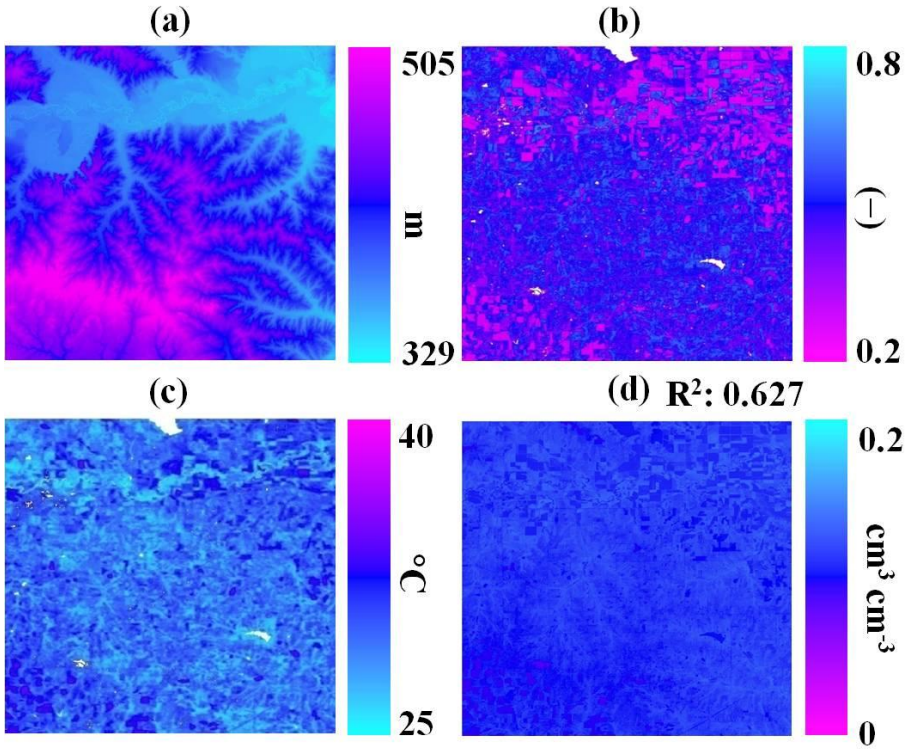

Figure 5.12: (a) Digital elevation model (DEM), (b) pixel-based (LANDSAT7-ETM) normalized difference vegetation index (NDVI, 30 m×30 m), (c) pixel-based (LANDSAT7-ETM) temperature (30 m×30 m), and (d) downscaled (30 m×30 m) soil moisture (DOY: 119, 2003) for the LW site at the satellite-scale

Figure 5.13 shows the satellite (AMSR-E)-based and downscaled (average for all sub-pixels) near-surface soil moisture dynamics along the time steps. The downscaled soil moisture dynamics were considerably overestimated compared to the AMSR-E products during the early rainy period (DOY: 152-180) with the correlation (R^2: 0.351) and uncertainties (RMSE: 0.090), but the results during the dry days (DOY: 181-212) relatively matched well. It may indicate that the AMSR-E soil moisture products have been tempered by C-band frequencies contaminated with anthropogenic radio frequency interference (RFI) causing underestimation [*Jackson et al.*, 2005]. Also, the rainfall event (16.7 mm) on DOY 177 was generated, but the AMSR-E soil moisture value was not corresponded. It might show that the scale discrepancy between the weather station-based rainfall and AMSR-E products caused uncertainties in modeling across the regional site. Overall, this approach estimated well the near-surface soil moisture dynamics at the spatio-temporal scales with existing uncertainties.

We compared the water retention curves ($\theta(h)$) estimated by this approach and the field-observations at the LW 03 site located within the boundary of LW site in Fig. 5.13b. Although we could not validate these results with *in-situ* measurements directly at the spatial domain, good match of the derived and field-observed $\theta(h)$ curves (Fig. 5.13b) supported the robustness of this approach. Our findings showed that this approach provides realistic downscaled near-surface soil moisture dynamics at the satellite-scale.

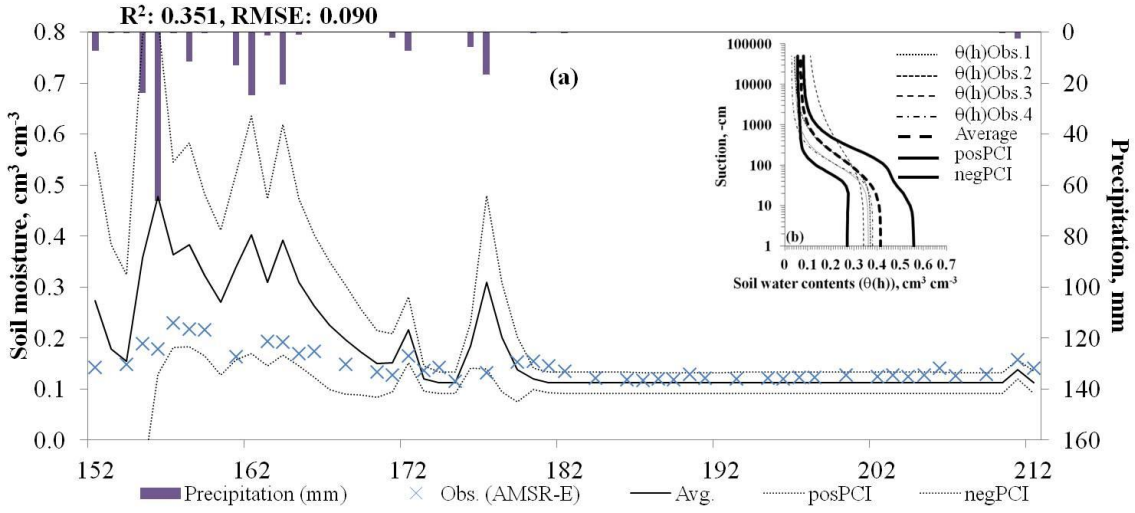

Figure 5.13: (a) The precipitation, remotely sensed (AMSR-E), and average (for the downscaled sub-pixels) near-surface (0-1 cm) soil moisture dynamics at the LW site and (b) comparison of the field-observed and estimated curves (θ(h))

5.5 Conclusions

In this study, we developed a new genetic algorithm-based spatial scaling algorithm that can downscale remotely sensed (RS) soil moisture footprints as well as upscale *in-situ* measurements across a complex regional area. Normalized digital elevation model (NDEM) and normalized difference vegetation index (NDVI) were used to characterize the heterogeneity of topography and vegetation covers within a RS pixel. The newly developed approach quantified soil hydraulic parameters using the soil-water-atmosphere-plant (SWAP) model with a genetic algorithm (GA) based on an inversion model. Near-surface soil moisture estimates were driven by the SWAP model using the derived hydraulic parameters in a forward mode. Then, the soil moisture predictions for different elevations and NDVI values were generated across field sites. Finally, the

distributed soil moisture predictions were used as downscaled and upscaled results. The Little Washita (LW) watershed in Oklahoma was selected for validating our proposed methodology including two scenarios; i) scenario 1: airborne-scale for the LW 13 and 21 sites and ii) scenario 2: satellite-scale for the LW site.

In the scenario 1, we tested the impacts of various initial and bottom boundary conditions (see Table 5.4) for the model performance at the LW 13 site. This result showed that the initial and bottom boundary conditions significantly influence root zone soil moisture dynamics in modeling. The downscaled and upscaled soil moisture predictions had more uncertainties when more weight was given to the NDVI (f=0.1) or NDEM (f=0.9), respectively. However, the joint NDEM and NDVI (f=0.5) reduced uncertainties for the scaling performance. The downscaled (R^2: 0.204~0.406 and RMSE: 0.040~0.102) and upscaled (R^2: 0.204~0.500 and RMSE: 0.064~0.091) soil moisture predictions at the LW 13 and 21 sites had good match with the *in-situ* measurement across the regional areas, although uncertainties exist due to the measuring errors, model structure, a lag between sampling points and pixel-based DEM and NDVI, etc.

The averaged near-surface (0-5 cm) soil moisture dynamics of all sub-pixels for the downscaled (R^2: 0.706~0.903 and RMSE: 0.019~0.084) and upscaled (R^2: 0.742~0.941 and RMSE: 0.027~0.083) results matched well with both the airborne sensing (ESTAR) and *in-situ* soil moisture data across the LW 13 and 21 sites. Also, the airborne-based (ESTAR) soil moisture products taken over the LW 21 (relatively flat topography) matched well with the *in-situ* measurements, but the RS datasets at the LW 13 site were considerably lower than *in-situ* data indicating bias included within pixel-

based soil moisture products taken across a hill slope. Also, the upscaled results by using various hydrological models (SWAP, Noah LSM, and CLM) matched well with the *in-situ* soil moisture data across the LW 13 site, while the SWAP model estimated slightly better than others.

At the satellite-scale, this approach downscaled well the near-surface (0-1 cm) soil moisture distributions on DOY 119 (R^2: 0.627) at the LW site compared to the fine-scale (LANDSAT7-ETM, 30 m $\times$ 30 m) land surface temperature (°C), although our approach was limited in validating with the *in-situ* measurements due to no datasets available at the satellite-scale. The averaged (for all sub-pixels) soil moisture dynamics along the time steps also appeared to be identifiable with the AMSR-E footprints with the statistics (R^2: 0.351 and RMSE: 0.090), although there exists uncertainties in the early simulation period. Good match of field-observed and estimated water retention curves ($\theta(h)$) at the airborne-/satellite-scales supported the robustness of our proposed methodology. This approach provides easy applications for downscaling RS soil moisture products as well as upscaling *in-situ* measurements simultaneously with the given observations. Also, the scaling algorithm can characterize the variability of topography and vegetation covers across topographically complex terrains within a RS pixel based on the NDEM and NDVI obtainable easily at multiple scales. Thus, we suggest that our proposed approach could be very useful to improve the availability of remotely sensed and *in-situ* soil moisture data for water resources management efficiently in the real world.

CHAPTER VI

DEVELOPMENT OF NON-PARAMETRIC EVOLUTIONARY ALGORITHM FOR PREDICTION OF ROOT ZONE SOIL MOISTURE

6.1 Synopsis

Prediction of root zone soil moisture is critical for water resources management. In this study, we explored a multivariate framework for prediction of root zone soil moisture from a time series of spatially-distributed rainfall across multiple weather locations under two different hydro-climatic regions. A new non-parametric evolutionary algorithm (genetic algorithm-based hidden Markov model, HMMGA) was developed to estimate long-term root zone soil moisture dynamics at different soil depths. Also, we analyzed rainfall occurrence probabilities and dry/wet spell lengths reproduced by this approach. The HMMGA was used to estimate the optimal state sequences (surface soil wetness) based on the precipitation history. Historical root zone soil moisture statistics were then determined based on the soil wetness conditions. To test the new approach, we selected two different soil moisture fields, Oklahoma (130 km × 130 km) and Illinois (300 km × 500 km), during 1995 to 2009 and 1994 to 2010, respectively. We found that the newly developed framework performed well in predicting root zone soil moisture dynamics at both the spatial scales. Also, the reproduced rainfall occurrence probabilities and dry/wet spell lengths matched well with the observations at the spatio-temporal scales. Since the proposed algorithm requires only precipitation and historical soil moisture data from existing, established weather

stations, it can serve an attractive alternative for predicting root zone soil moisture in the future using climate change scenarios and root zone soil moisture history.

6.2 Introduction

Root zone soil moisture is an important component of many hydro-climatic processes. Accurate measurements of root zone soil moisture are required for water resources management, understanding rainfall and runoff processes, irrigation scheduling, water quality monitoring, determining the partitioning of sensible and latent heat fluxes, etc. In general, *in-situ* soil moisture measurements are fairly accurate, but these point-scale measurements may not be representative at the field scale. Since remote sensing (RS) soil moisture retrievals have been developed and improved, remote sensing techniques are better suited for estimating root zone soil moisture with respect to spatial and temporal coverages across the world [*Ines and Mohanty,* 2008a,b, 2009]. The major concern regarding RS products is, however, the question of whether their resolution is too coarse to represent the scale at which hydrological processes occur [*Engman*, 1991; *Entekhabi et al*., 1999]. This indicates that RS data suffer on account of the scale discrepancy between observation (RS) and modeling resolution.

Various physically-based hydrological models such as Soil-Plant-Atmosphere-Water (SWAP) [*van Dam et al*., 1997], Community Land Model (CLM) [CLM3.0 User's Guide, *Oleson et al*., 2004], Variable Infiltration Capacity (VIC) Macroscale Hydrologic Model [*Liang et al*., 1996], the U.S. Department of Agriculture Hydrograph Laboratory (USDAHL) model [*Holtan*, 1961], the Sacramento Soil Moisture Accounting (SAC-SMA) Model [*Peck*, 1976], among others were developed for estimating root zone

soil moisture dynamics. Besides the above mentioned deterministic models, several other approaches have been developed for soil moisture estimations. A sequential assimilation approach was developed [*Mahfouf*, 1990] and improved [*Bouttier et al*., 1993a,b] for soil moisture estimations from atmospheric temperature and relative humidity. A genetic algorithm-based assimilation scheme was developed for quantifying effective soil hydraulic properties based on inverse method (IM) [*Ines and Mohanty*, 2008a,b, 2009; *Shin et al*., 2012]. In addition, statistical models have also been developed for forecasting soil moisture estimates. *Liou et al*., [2001] explored an error propagation learning back propagation (EPLBP) neural network for retrieving soil moisture dynamics using the brightness temperature. Neural networks using a backscattering coefficient were used to retrieve surface roughness and soil moisture with various radar configurations (VV 23°, HH 39°, and HH 47°) [*Baghdadi et al*., 2002]. However, these approaches are limited by the number and complexity of required physical parameters in modeling.

In the case of weather prediction, various stochastic models such as parametric empirical-statistical models [*Stern and Coe*, 1984; *Woolhiser and Roldan*, 1982], non-parametric scheme [*Young*, 1994; *Lall et al*., 1996], etc. exist for representing daily weather sequence. *Wilks* [1998a,b] conditioned daily stochastic precipitation at multiple sites on total monthly precipitation. *Young* [1994] described a multivariate chain model for simultaneously simulating temperatures and daily precipitation amounts. *Kirshner* [2005] used hidden Markov models (HMMs) to model discrete multivariate time series

data of rainfall. However, till date, no stochastic methods have been designed for forecasting root zone soil moisture dynamics across the land surface.

In this study, we explored a multivariate framework for forecasting long-term daily root zone soil moisture dynamics using a stochastic optimization approach on a multivariate time series of rainfall datasets obtained from a network of weather stations. The objective of our study was to predict root zone soil moisture dynamics at various spatial domains using a newly developed non-parametric evolutionary algorithm (genetic-algorithm based hidden Markov model, HMMGA) under different hydro-climate regions. Additionally, rainfall occurrence probabilities and dry/wet spell lengths estimated by the HMMGA were analyzed at the spatio-temporal scales. Overall the study is to develop a HMMGA algorithm for surface and root zone soil moisture prediction based on rainfall history and dry/wet pattern.

6.3 Methods and Materials

6.3.1 Conceptual Multivariate Framework

In this study, we developed a novel methodology for daily root zone soil moisture predictions at various spatial-scales using multi-site precipitation data. The approach comprised of three steps; i) development of a hidden Markov model using a genetic algorithm (HMMGA) for estimating the optimal state sequences (representing daily surface soil wetness) using precipitation; ii) estimation of statistics (mean and standard deviation) of categorized root zone soil moisture measurements based on the surface soil wetness, and iii) root zone soil moisture predictions using the derived

surface soil wetness and statistics of categorized soil moisture history. The methodology is schematically shown in Fig. 6.1.

Stochastic approaches coupled with physically-based hydrological models have been commonly used in hydrological predictions such as optimum interpolation-based sequential assimilation technique for estimating soil moisture [*Bouttier et al*., 1993a,b], genetic algorithm-based near-surface [*Ines and Mohanty*, 2008a,b, 2009] and layer-specific [*Shin et al.*, 2012] soil moisture data assimilation, among others. However, these coupled models are parameter intensive and are considerably influenced by the initial parameter values. In order to address this limitation of existing schemes, we developed a new algorithm by integrating a stochastic hidden Markov model (HMM) with genetic algorithm. The hypothesis of our proposed approach is that a HMM can provide better trained parameters for obtaining the surface soil wetness through optimized initial parameters derived by a GA. This coupling eliminated the need of using a parameter intensive physical model and is immune to user-based errors in estimating initial parameter values.

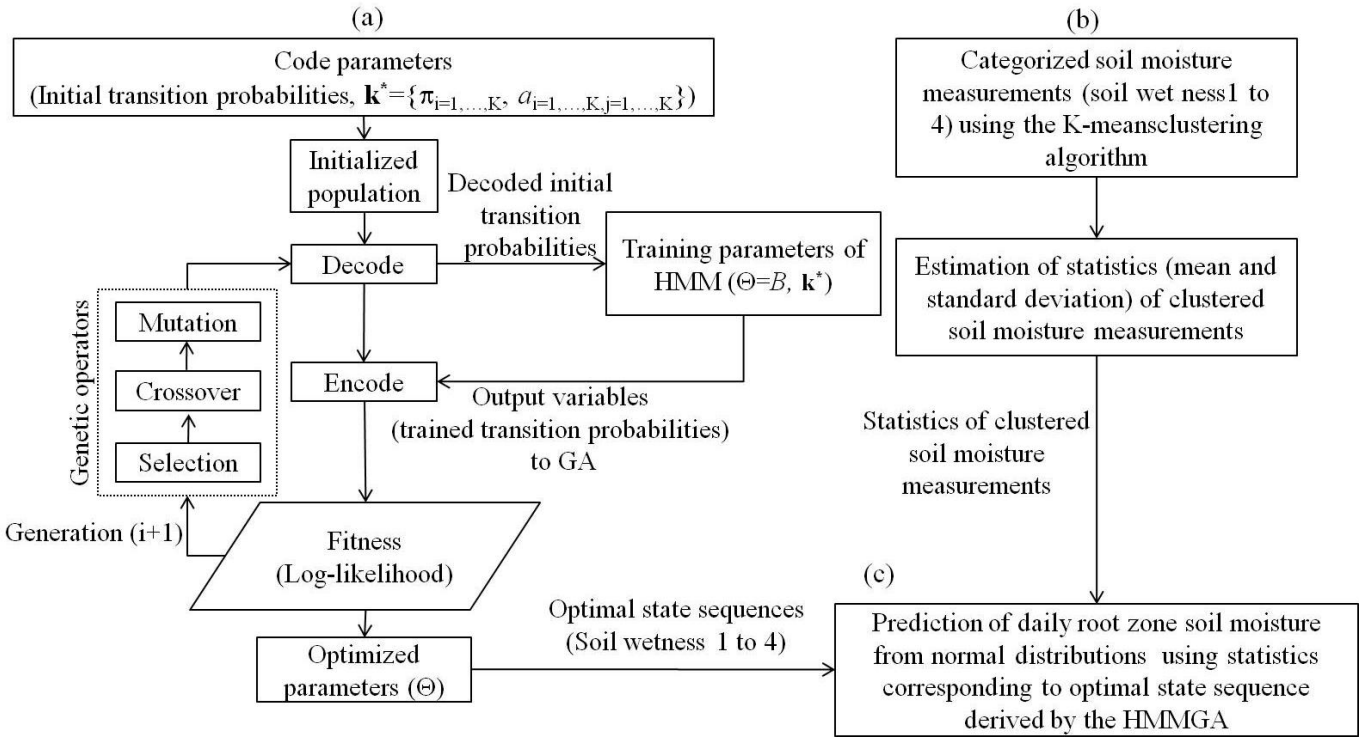

Figure 6.1: Flow chart of the multivariate framework for forecasting root zone soil moisture dynamics; (a) hidden Markov model based on a genetic algorithm (HMMGA), (b) categorized soil moisture measurements using the K-means clustering algorithm, and (c) predictions of root zone soil moisture dynamics

State 1: wet condition; State 2: relatively wet condition; State 3: relatively dry condition; State 4: dry condition

6.3.2 Hidden Markov Model (HMM) Based on Genetic Algorithm

In a regular Markov model, state transition probabilities are directly visible to the observer. However, as its name suggests, a hidden Markov model (HMM) has invisible (or hidden) state transition probabilities. A HMM can be defined as a joint distribution on a rainfall occurrence observation and hidden state of rainfall along the time (t) based on two conditional independence assumptions [*Hughes et al.,* 1999; *Robertson et al*., 2004; *Kirshner*, 2005]. Note that we denoted all random variables as capital letters (e.g., O and S) and values of random variables by lowercase letters (e.g., o and s):

(i) the random vector of rainfall occurrence O_t (O_t^1 ,…, O_t^M) from multiple weather stations (M) at the time (t) is independent of other occurrences (O_{t-1} or O_{t+1}), conditional only on the hidden states S_t (S_1,…,S_T) with the number K (representing the transition matrix of K × K) of hidden states at the time (t) in Eq. (6.1). The observed rainfall value $o_t^m = 1$ indicates "rain" on day t and $o_t^m = 0$ means "no rain" on day t at weather station m (m=1,…,M), respectively. Note that this assumption only represents independence of rainfall events across time. The spatial dependence of rainfall is accounted for by the state variable in the model,

$$P(O_t|S_{1:t}, O_{1:t-1}) = P(O_t|S_t) \quad (6.1)$$

(ii) the hidden (rainfall) state process (S_t) is first-order Markovian (representing that Markov process is homogeneous with the transition matrix of K×K in time) in Eq. (6.2).

$$P(S_t|S_{1:t-1}) = P(S_t|S_{t-1}) \quad (6.2)$$

It is also assumed that the rainfall observations across weather stations (m) at the time (t) are spatially independent (conditional on the hidden state) [*Robertson et al*., 2004; *Kirshner*, 2005]. We estimated $P(O_t|S_t)$ as,

$$P(O_t|S_t)=\prod_{m=1}^{M} P(O_t^m = o \mid S_t = s) = \prod_{m=1}^{M} p_{s,o}^m \quad (6.3)$$

where the rainfall occurrences at each weather station (m) were denoted as "o" ($p_{s,o=0}^m$ means no rain and $p_{s,o=1}^m$ indicates rain). The readers are referred to *Rabiner* [1989], *McDonald and Zucchini* [1997], *Robertson et al*. [2004], and *Kirshiner* [2005] for a more detailed explanation of the HMM.

Sequences of daily rainfall occurrences (O_t) and hidden rainfall states (S_t) are shown in Fig. 6.2. In this study, a GA searches the optimized first (initial) hidden state parameters ($\pi_{i=1,\ldots,K}$) and transition matrix ($a_{i=1,\ldots K,j=1,\ldots,K}$) composed of parameters ($\mathbf{k}^*=\{\pi_{i,}\, a_{i,j}\}$) to be trained by a HMM, because GAs are powerful search algorithms for solving complex problems (parameters) through the survival of the fittest mechanism [*Holland*, 1975; *Goldberg*, 1989]. Thus, a HMM can train a set of parameters (Θ, Eq. 6.4) for estimating the surface soil wetness conditions using the optimized initial hidden states ($\mathbf{k}^*$) provided by a GA,

$$\Theta = (B, \mathbf{k}^*) = (A, B, \pi_i) \quad (6.4)$$

where, A is the transition probabilities (a_{ij}) on row state i and column state j, B is the observation (output) probability, π_i is the first state probability at the time (t=1) on row state i.

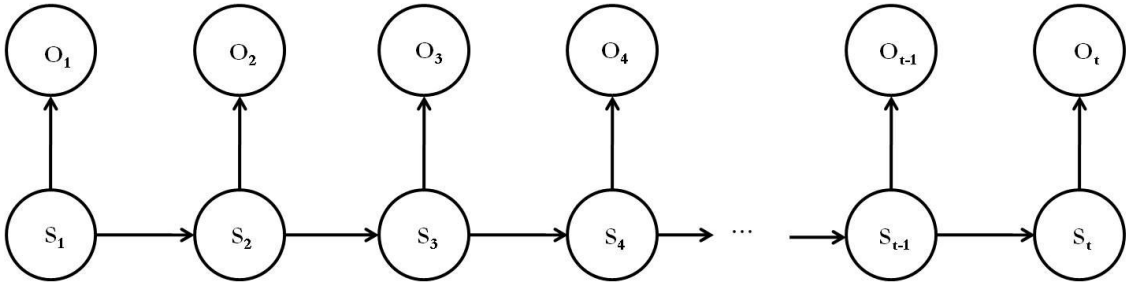

Figure 6.2: Graphical model representation of a hidden Markov model
O_t: the observations (outputs) of rainfall occurrences with the time series (t)
S_t: the hidden states with the time series (t)

The appropriate number of transition probabilities (K=2,…,N) are necessary to predict accurate rainfall occurrences. It must be considered that a high N value will increase the number of elements in the transition matrix (which is of size $K \times K$) causing excessive repetitions of simulations. There are no set guidelines for determining the optimum orders of K. *Robertson et al.,* [2004] tested the HMM with K ranging from 2 to 6 and suggested that four hidden states (K=4) provide the most reasonable results. Therefore, we tested different orders of K (2,…,6) for selecting the optimum order K of hidden states, which gives better results.

We used the log-likelihood (*l*) of the datasets to improve the computational process for this statistical model in Eq. (6.5),

$$l=\log P(O)=\log \sum_{S}[P(S_1)\prod_{t=2}^{T}P(S_t \mid S_{t-1})][\prod_{t=1}^{T}P(O_t \mid S_t)] \quad (6.5)$$

To compute the log-likelihood of the data $P(O_{1:T}|S_{1:T})$, the *Forward-Backward* procedure [*Rabiner*, 1989] was adopted as a recursive method. For observations

($O_1,\ldots,O_T$) and optimal state sequences ($S_1,\ldots,S_T$) to reproduce the observations is given by the recurrence relations for the *Forward* algorithm,

$$V_1(i)=P(O_1|S_1)\pi_i \tag{6.6}$$

$$V_t(j)= \max_{1\le i\le K} P(O_t \mid S_t)\boldsymbol{a_{ij}} V_{t-1}(\boldsymbol{i}) \tag{6.7}$$

Here V_t is the probability of optimal state sequences (Viterbi path, *Viterbi*, [1967]) representing the daily surface soil wetness for the observation at time t+1. In the *Backward* procedure, the optimal state sequences can be retrieved by saving back pointers that remember which state S_t was used,

$$S_T= \max_{1\le i\le K} V_T(\boldsymbol{i}) \tag{6.8}$$

$$S_t=S_{t+1} \arg\max_{1\le i\le K} \boldsymbol{a_{ij}} V_t(\boldsymbol{i}) \tag{6.9}$$

Note that "arg max" meant the argument for selecting the maximum value in the given data sets (i=1,…,K). Fig. 6.3 shows the log-likelihood values and correlations (R^2, see section 2.3) between the observed rainfall occurrence probabilities and surface soil wetness conditions derived by the HMMGA for various orders of hidden states (K=2,…,6) in the Oklahoma domain (see section 2.2). Note that a log-likelihood value closer to zero denotes better performance. The log-likelihood values considerably decreased from K=2 to K=5, but the decreasing trend from K=5 to K=6 was relatively gentle. The correlation was highest at K=4 indicating that four hidden states can represent the soil wetness conditions better than others as shown in the findings of *Robertson et al.*, [2004]. Thus, we selected the four hidden states (K=4) for estimating surface soil wetness conditions. Four hidden states represent wet (state 1), relatively wet (state 2), relatively dry (state 3), and dry (state 4) conditions, respectively.

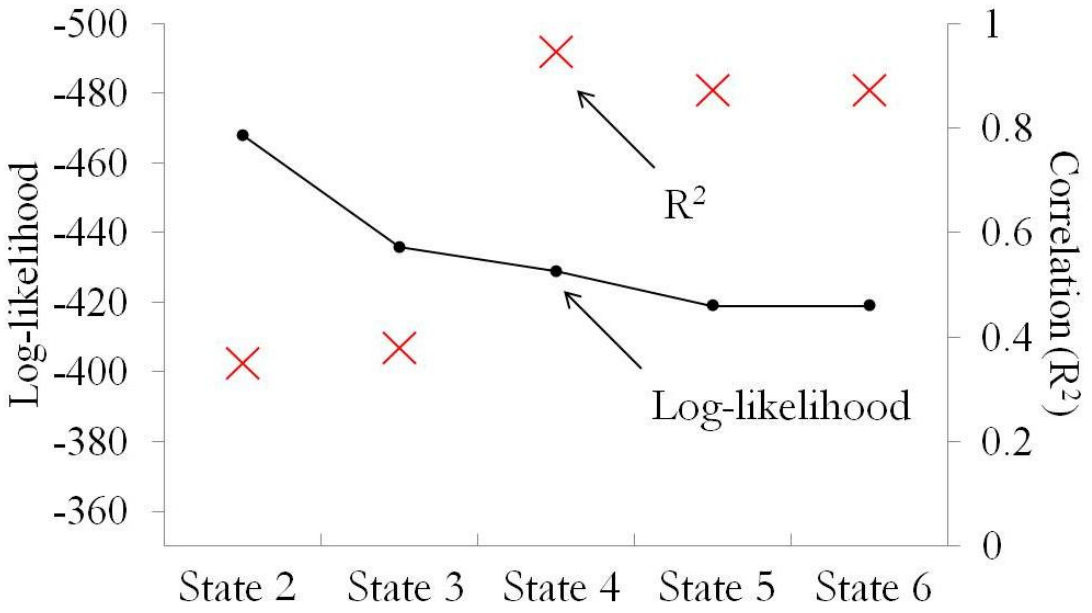

Figure 6.3: Comparison of log-likelihood values and correlations (between the observed rainfall occurrences and HMMGA-based optimal state sequences) with different orders (K=2,…,6) of hidden states for the Oklahoma domain

State 2: K=2; State 3: K=3, State 4: K=4, State 5: K=5; State 6: K=6

In this study, a modified-*micro*GA [*Ines and Honda*, 2005; *Ines and Mohanty*, 2008a] was used to search parameters ($\mathbf{k}^*$) along the generations. The parameters are composed of genes, which are arranged in an array of binary values called a chromosome. The modified-*micro*GA is a GA variant that uses a micro-population for searching parameters [*Ines and Droogers*, 2002; *Carroll*, 1996; *Goldberg*, 2002; *Krishnakumar*, 1989]. It has the unique ability to restart when the searched chromosomes of the micro-population are similar in structure (90 %) to each other before the generations are completed. This improves the algorithm computationally and makes it more efficient. Restarting allows the modified-*micro*GA to search the solutions in the global space. In order to infuse new genes, creep (at the decimal level, D. L. Carroll, Fortran genetic algorithm (GA) driver, available at www.cuaerospace.com/carroll/ga.html) and intermittent jump [*Ines and Mohanty*, 2008a] mutations are implemented in the modified-*micro*GA. Furthermore, a time saving mechanism used in the GA remembers the elite (fittest) chromosome in the previous generation (g-1) [*Ines and Honda*, 2005]. Then the elite chromosome is reproduced in the next generation (g) without computation (note that the modified-*micro*GA always reproduces the elite chromosome in the next generation). The search spaces of variables and probabilities of selection, crossover, and mutation used in a GA for the different study domains are shown in Table 6.1.

Table 6.1: The search spaces of variables and probabilities of selection, crossover, and mutation used in the genetic algorithm

Locations	Parameters	Search space of parameters		No. of bit (L)	2L	Population size	Seed Number	Number of generation	P_{creep}	P_{mutate}	P_{cross}
		Minimum values	Maximum Values								
Oklahoma domain	π_i	0.0	1.0	7	128	10	3000	30	0.05	0.05	0.05
	a_{ij}	0.0	1.0	7	128						
Illinois domain	π_i	0.0	1.0	7	128	10	3000	30	0.05	0.3	0.05
	a_{ij}	0.0	1.0	7	128						

$\pi_{i=1,\ldots,K}$: initial state probabilities of being in state *i* (K is the number of transition probabilities)
a_{ij}: transition probabilities (K × K) of transitioning from state *i* to state *j*
P_{creep}: creep probability
P_{mutate}: mutation probability
P_{cross}: crossover probability

6.3.3 Categorizing Root Zone Soil Moisture

In this study, the HMMGA predicted the surface soil wetness conditions using only the observed rainfall occurrences. However, soil moisture values are usually affected by the precipitation, as well as by the land surface conditions (i.e., soil textures, vegetation covers, topography, etc.) in different regions. Thus, corresponding to the four-hidden states fitted by the HMMGA in section 2.1.1, we classified historical root zone soil moisture data into four categories (K=4) ranging from wet to dry (represented by 1, 2, 3, 4 in order of wetness) at each weather station using the K-means clustering algorithm [*Hartigan and Wong*, 1979] based on the Euclidean distance,

$$d(x,y) = \frac{2}{L}\sum_{q=1}^{Q}(x_q - y_q)^2 \qquad (6.10)$$

where d(x,y) is the Euclidean distance between two points of x_q and y_q on the order of *q*, x_q and y_q are the root zone soil moisture measurement on the order of *q*, *q* is the running index, and Q is the total number of root zone soil moisture measurements, respectively.

Initially, the K-means clustering algorithm assigns the soil moisture data randomly to a cluster. This clustering approach repeats a two-step process in the given number (n=100) of runs as follows; i) the mean soil moisture data of individual cluster is estimated and ii) the soil moisture data are reallocated to each cluster whose mean value is closest to the soil moisture data. For more details on the popular K-means clustering algorithm refer to *Jain et al*., 1999. In order to incorporate site specific root zone soil moisture dynamics, the clustering of soil moisture data was done for each weather station (m) separately. Then, we calculated mean ($\mu_{i=1,\ldots K,m=1,\ldots M}$) and standard deviation (

$\sigma_{i=1,\ldots K,m=1,\ldots M}$) of the classified root zone soil moisture measurements for each weather station. The soil wetness ranges (maximum and minimum) for four wetness categories were determined based on clustered soil moisture measurements at individual weather station.

6.3.4 Prediction of Root Zone Soil Moisture

Based on the derived optimal state sequences ($V_t=\{S_1,\ldots,S_T\}$, section 2.1.1) and estimated statistics ($\mu_{i,m}$ and $\sigma_{i,m}$, the section 2.1.2) of historical soil moisture data, we generated normal distributions ($N(\mu_{i,m}, \sigma^2_{i,m})$) of root zone soil moisture for each individual surface soil wetness state along the time series. The predicted daily root zone soil moisture (θ_t^m) value was determined by randomly selecting a value from the generated normal distribution corresponding to the optimal state sequences selected by the HMMGA (Eq. 6.11). Each optimal state value indicates the surface wetness condition (S_t) at the time (t). Note that the function of f( ) returns the randomly generated soil moisture values for each weather station (m) along the time series (t).

$$\theta_t^m = f(V_t, N(\mu_{i,m}, \sigma^2_{i,m})) \quad \forall m \ \forall t \tag{6.11}$$

6.3.5 Study Sites and Data Description

We Since the proposed approach is sensitive to the domain size, two domains of different sizes were chosen to evaluate the applicability of this approach under two hydro-climatic regions; i) the Oklahoma domain (130 km × 130 km) with humid subtropical climate and ii) the Illinois domain (300 km × 500 km) with humid continental climate, respectively. For the Oklahoma domain, we selected seven Mesonet

weather stations (site 1-ELRE, site 2-SPEN, site 3-MINC, site 4-ACME, site 5-WASH, site 6-PAUL, and site 7-BYAR, http://www.mesonet.org/index.php) as shown in Fig. 6.4a. The weather stations have elevations in the range of 265.8 m ~ 421.2 m above MSL. Sites 1, 3, and 4 in the west have higher elevations (404.3 m ~ 421.2 m) than those located in the east (265.6 m ~ 349.5 m). For the Illinois domain, we selected seven weather station sites from the International Soil Moisture Network (ISMN: http://www.ipf.tuwien.ac.at/insitu/) in Illinois (site 1-Bondville, site 2-Dekalb, site 3-Freeport, site 4-Monmouth, site 5-Olney, site 6-Orr Center, and site 7-Peoria) as shown in Fig. 6.4b. The weather stations located in the northern portion of this domain had slightly higher elevations. The Oklahoma and Illinois areas generally have a flat topography.

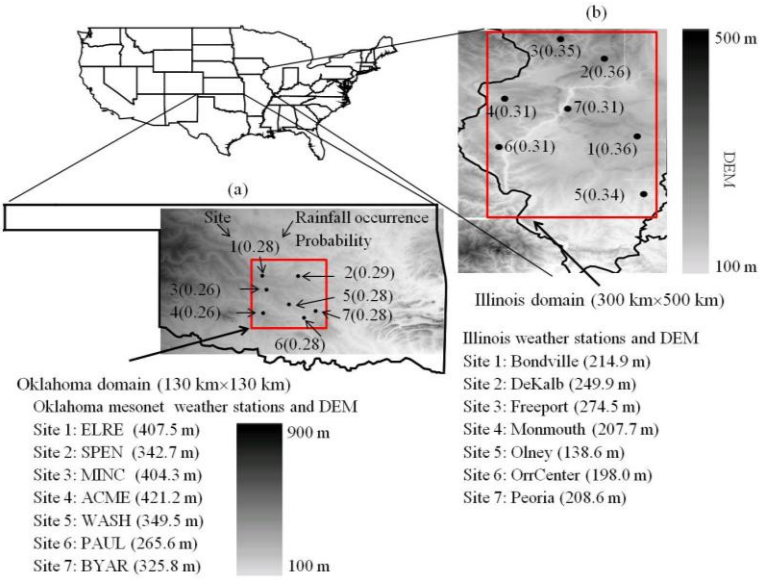

Figure 6.4: Network of multiple locations (sites 1 to 7), digital elevation model (DEM), and observed rainfall occurrence probabilities under different hydro-climatic regions; (a) Oklahoma domain and (b) Illinois domain

Daily precipitation datasets were obtained from the Oklahoma mesonet stations (1995 to 2009) in Oklahoma and the ISMN (1994-2010) in Illinois during the crop growing season (90 days from March 1 to May 29). At the Oklahoma sites, daily soil temperature readings (available from 2003 to 2009) measured by Campbell scientific 229-L sensors were converted to root zone (0-5 cm) volumetric soil moisture with the empirical coefficients provided from the mesonet stations [see details, *Illston et al.*, 2008]. At the Illinois locations, biweekly volumetric soil moisture data at different soil depths (0-10, 10-30, 30-50, and 0-50 cm) were measured from 1994-2004 using Troxler neutron surface and depth probes by the Illinois State Water Survey [*Hollinger and Isard*, 1994]. Note that soil moisture at 0-50 cm was obtained by averaging soil moisture measurements at the soil depths of 0-10, 10-30 and 30-50 cm. Descriptive statistics (mean, standard deviation, minimum, and maximum) of root zone soil moisture measurements for the Oklahoma and Illinois sites were shown in Table 6.2 and 6.3.

Precipitation data from the weather stations were available from 1995-2009 in Oklahoma and 1994-2010 in Illinois. However, the soil moisture data were measured for a shorter duration (2003 to 2009 for Oklahoma and 1994 to 2004 for Illinois). The optimal (hidden) state sequences for both study domains were determined by the HMMGA using the entire range of precipitation data available. However, in order to generate the normal distributions for the root zone soil moisture, we used only the estimated optimal state sequences from 2003 to 2009 for Oklahoma and 1994 to 2004 for Illinois. The measurements used for the calibration were taken from 1994 to 2000 and the validation was conducted from 2001 to 2004 at the Illinois sites.

Table 6.2: Statistics of categorized soil moisture measurements at the soil depth (0-5cm) using the K-means clustering algorithm in Oklahoma (2003-2009)

Soil depth	Sites	Oklahoma[a]															
		State 1 (Wet condition)				State 2 (Relatively wet condition)				State 3 (Relatively dry condition)				State 4 (Dry condition)			
		Mean	SD	Min.	Max.	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
0-5 cm	1	0.391	0.014	0.379	0.418	0.350	0.013	0.328	0.361	0.293	0.015	0.273	0.314	0.248	0.011	0.234	0.266
	2	0.221	0.001	0.221	0.227	0.211	0.002	0.208	0.216	0.201	0.003	0.195	0.205	0.187	0.003	0.183	0.193
	3	0.364	0.014	0.346	0.391	0.321	0.007	0.316	0.331	0.290	0.012	0.260	0.302	0.224	0.011	0.215	0.253
	4	0.246	0.004	0.243	0.257	0.235	0.002	0.232	0.237	0.221	0.005	0.213	0.227	0.200	0.005	0.194	0.210
	5	0.294	0.003	0.293	0.304	0.281	0.003	0.274	0.283	0.253	0.009	0.238	0.265	0.217	0.011	0.201	0.233
	6	0.401	0.005	0.400	0.423	0.370	0.010	0.355	0.377	0.314	0.015	0.297	0.334	0.253	0.026	0.209	0.282
	7	0.204	0.000	0.202	0.206	0.198	0.002	0.195	0.199	0.186	0.003	0.181	0.191	0.173	0.004	0.169	0.179

[a] Campbell scientific 229-L sensor (daily soil temperature estimates were measured by this sensor and converted to daily soil moisture dynamics using the equation of volumetric water content (VWC: $cm^3\ cm^{-3}$) with the empirical coefficients provided from the mesonet stations [*Illston et al.*, 2008].)

Table 6.3: Statistics of categorized soil moisture measurements at the different soil depths (0-10, 10-30, 30-50, and 0-50 cm) using the K-means clustering algorithm in Illinois (1994-2000)

Soil depths	Sites	Illinois (Troxler Neutron Surface and Depth Probes [*Hollinger and Isard*, 1994])															
		State 1 (Wet condition)				State 2 (Relatively wet condition)				State 3 (Relatively dry condition)				State 4 (Dry condition)			
		Mean	SD	Min.	Max.	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
0-10 cm	1	0.493	0.037	0.456	0.558	0.414	0.018	0.384	0.439	0.338	0.020	0.312	0.363	0.273	0.024	0.218	0.300
	2	0.444	0.025	0.415	0.503	0.380	0.016	0.354	0.405	0.316	0.016	0.298	0.343	0.260	0.015	0.244	0.280
	3	0.448	0.016	0.434	0.481	0.405	0.010	0.389	0.421	0.339	0.016	0.304	0.361	0.246	0.015	0.221	0.265
	4	0.464	0.032	0.424	0.506	0.380	0.020	0.352	0.417	0.311	0.021	0.281	0.335	0.240	0.029	0.191	0.262
	5	0.531	0.025	0.509	0.581	0.474	0.017	0.450	0.494	0.420	0.014	0.399	0.445	0.352	0.036	0.306	0.384
	6	0.448	0.023	0.417	0.482	0.374	0.017	0.348	0.406	0.310	0.021	0.271	0.339	0.202	0.037	0.167	0.240
	7	0.515	0.044	0.473	0.580	0.404	0.020	0.373	0.441	0.335	0.018	0.304	0.361	0.252	0.030	0.206	0.287
10-30 cm	1	0.428	0.012	0.416	0.447	0.385	0.011	0.369	0.400	0.349	0.011	0.331	0.365	0.285	0.015	0.267	0.302
	2	0.407	0.021	0.386	0.444	0.350	0.019	0.326	0.376	0.297	0.020	0.262	0.321	0.182	0.035	0.147	0.217
	3	0.370	0.016	0.350	0.413	0.323	0.010	0.306	0.336	0.285	0.019	0.254	0.301	0.171	0.010	0.161	0.181
	4	0.403	0.015	0.383	0.420	0.361	0.010	0.348	0.380	0.325	0.013	0.301	0.340	0.273	0.019	0.242	0.289
	5	0.341	0.003	0.337	0.347	0.332	0.002	0.328	0.336	0.323	0.003	0.318	0.327	0.303	0.010	0.291	0.309
	6	0.408	0.015	0.389	0.438	0.358	0.010	0.343	0.377	0.310	0.020	0.282	0.327	0.243	0.014	0.234	0.264
	7	0.515	0.044	0.473	0.580	0.404	0.020	0.373	0.441	0.335	0.018	0.304	0.361	0.252	0.030	0.206	0.287
30-50 cm	1	0.421	0.011	0.404	0.439	0.384	0.007	0.376	0.400	0.366	0.006	0.350	0.373	0.316	0.014	0.300	0.328
	2	0.427	0.014	0.409	0.453	0.386	0.010	0.374	0.404	0.359	0.009	0.344	0.371	0.318	0.031	0.272	0.337
	3	0.347	0.011	0.336	0.363	0.324	0.008	0.311	0.333	0.297	0.011	0.283	0.306	0.219	0.010	0.209	0.229
	4	0.391	0.012	0.376	0.401	0.360	0.007	0.350	0.371	0.340	0.007	0.326	0.347	0.309	0.013	0.285	0.320
	5	0.374	0.004	0.368	0.382	0.356	0.003	0.353	0.361	0.347	0.003	0.341	0.350	0.333	0.005	0.327	0.338
	6	0.431	0.008	0.421	0.446	0.408	0.007	0.398	0.418	0.377	0.009	0.359	0.388	0.337	0.017	0.320	0.354
	7	0.432	0.008	0.425	0.444	0.415	0.006	0.404	0.421	0.387	0.006	0.376	0.398	0.349	0.009	0.340	0.358
0-50 cm	1	0.433	0.017	0.413	0.469	0.388	0.009	0.370	0.401	0.346	0.009	0.331	0.359	0.311	0.012	0.294	0.327
	2	0.411	0.015	0.383	0.434	0.355	0.013	0.339	0.371	0.314	0.012	0.293	0.332	0.221	0.010	0.211	0.231
	3	0.384	0.008	0.374	0.398	0.342	0.013	0.319	0.361	0.289	0.007	0.280	0.297	0.210	0.010	0.200	0.220
	4	0.423	0.006	0.416	0.430	0.367	0.015	0.344	0.392	0.318	0.010	0.305	0.337	0.273	0.011	0.261	0.282
	5	0.412	0.010	0.400	0.433	0.384	0.007	0.375	0.394	0.361	0.006	0.353	0.371	0.325	0.016	0.309	0.341
	6	0.432	0.012	0.417	0.453	0.392	0.010	0.374	0.410	0.349	0.010	0.333	0.363	0.281	0.023	0.247	0.297
	7	0.437	0.018	0.416	0.464	0.387	0.012	0.368	0.403	0.344	0.010	0.330	0.360	0.296	0.015	0.279	0.313

6.3.6 Analysis Method

The challenge of this study is to model and predict the rainfall occurrence probabilities and root zone soil moisture dynamics by using historical precipitation data and statistics of soil moisture measurements across the multiple locations (sites 1 to 7) for the Oklahoma and Illinois domains. In this study, the spatial dependence of rainfall occurrence was analyzed during given time periods in different study domains. Also, long-term daily precipitation records in any region can be analyzed by evaluating the probability of dry and wet spells based on a Markov chain model [*Barron, et al*., 2003; *Krishnamurti, et al*., 1995; *Wilks and Wilby*, 1999]. Thus, long-term frequency behaviors of dry and wet spell lengths were considered in the modeling. Here, we defined a dry spell length as consecutive dry days preceded and followed by a wet day (more than 0 mm rainfall) while a wet spell length means consecutive wet days preceded and followed by a dry day (no rainfall).

To verify the accuracy of our suggested HMMGA method, we compared root zone soil moisture predictions derived by this approach with those from a near-surface soil moisture assimilation scheme based on inversion model [*Ines and Mohanty*, 2008a] in Illinois. The data assimilation scheme determined effective soil hydraulic properties (α, n, θ_{res}, θ_{sat}, K_{sat}) by tuning a physical-based hydrological (Soil Water Atmosphere Plant: SWAP) model with a genetic algorithm [see details in *Ines and Mohanty*, 2008a]. The root zone soil moisture dynamics were estimated by the SWAP model [*van Dam et al.*, 1997] using the soil hydraulic parameters derived by GA.

In the proposed HMMGA process, we generated multiple realizations (r) using a re-sampling technique [*Efron*, 1982] for uncertainty analysis of daily root zone soil moisture dynamics. Also, the multi-populations (from initial random seed numbers of -1000, -950, and -750) were used for assessing uncertainties of the near-surface soil moisture assimilation scheme. The Pearson's correlation (R^2) and root mean square error (RMSE) were used for evaluating uncertainties,

$$R^2 = \frac{\boldsymbol{Cov}(X, Y)}{\boldsymbol{Var}(X)\boldsymbol{Var}(Y)} \tag{6.12}$$

$$RMSE = \sqrt{\frac{\sum_{z=1}^{n}(X_z - Y_z)^2}{n}} \tag{6.13}$$

where R^2 is the Pearson's correlation, X is the observed dataset, Y is the simulated dataset, *Cov* is the covariance of X and Y, *Var* is the variance of X or Y, n is the number of datasets, and z is the running index.

6.4 Results and Discussion

Figure 6.4 shows the observed rainfall occurrence probabilities across the multiple weather stations with the digital elevation models (DEM) for the Oklahoma (130 km × 130 km) and Illinois (300 km × 500 km) domains. The rainfall occurrence probabilities (0.31~0.36) in Illinois were usually higher than those (0.26~0.29) in Oklahoma. The rainfall occurrences were evenly spread within each study area in both domains. The topographies of the two domains were relatively flat and thus, the rainfall occurrences were not influenced as much by topography of the domains. The observed rainfall occurrence probabilities at the weather stations in the Oklahoma (relatively

smaller) domain were highly correlated among the weather stations (with $R^2 > 0.615$), but the Illinois domain had relatively lower correlations (R^2: 0.223~0.592) between sites, as shown in Fig. 6.5. It showed that spatial dependence of rainfall occurrences were significantly affected by the distance between the stations. In other words, the capability of HMM in simulating the multivariate time series of rainfall occurrences is influenced by how closely the weather stations are located across the spatial domain.

The plots in Fig. 6.6 represent the performance (maximum log-likelihood) of HMMGA for the Oklahoma and Illinois domains. The step-wise improvements in searched solutions show that the log-likelihood maximum fitness derived by the HMMGA was instantly converged during the initial generation (max: 30). This result demonstrated that HMM can train the parameter set (Θ from Eq. 6.3) well through the optimized initial parameters by the GA.

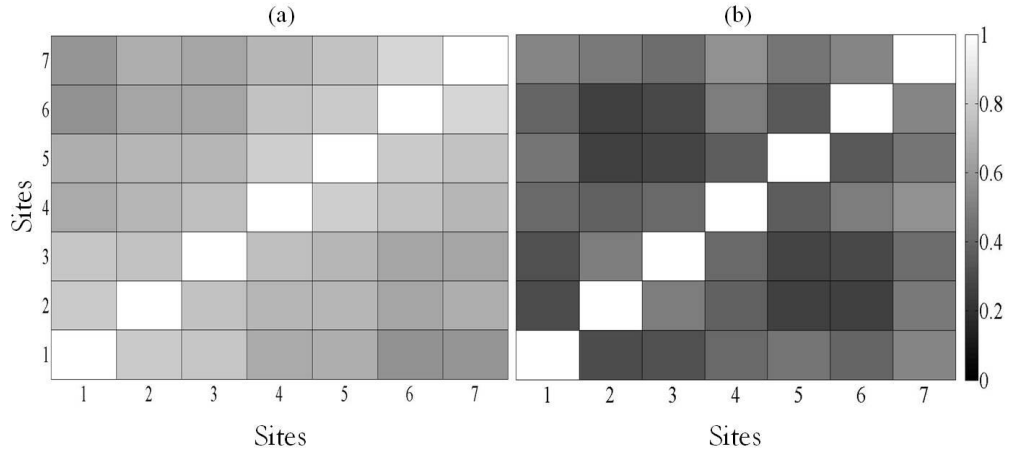

Figure 6.5: Correlations of observed rainfall occurrence probabilities at the network of multiple locations (site 1 to 7); (a) Small domain in Oklahoma, (b) Large domain in Illinois

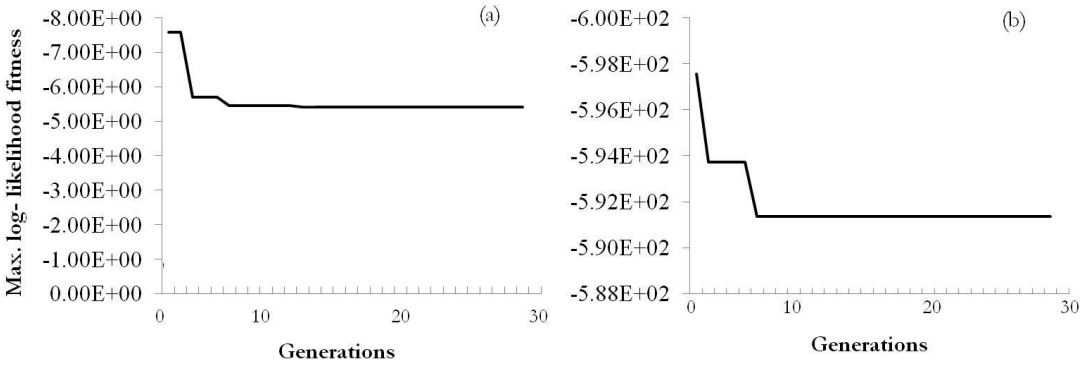

Figure 6.6: Fitness of hidden Markov model based on a genetic algorithm (HMMGA); (a) Small domain in Oklahoma, (b) Large domain in Illinois

6.4.1 Oklahoma Domain

The (hidden) surface wetness state transition probabilities for the Oklahoma domain (130 km × 130 km) are given in Table 6.4. The rows (*i*) represent the present wetness conditions (state 1: wet, state 2: relatively wet, state 3: relatively dry, and state 4: dry conditions) at time t whereas the columns (*j*) represent the expected wetness conditions at time t+1.

The Given a wet condition at time t, the probability of observing a wet condition at time t+1 is highest (at 0.455) while the probability of transitioning to a dry state is 0.318. A relatively wet condition at time t also has the relatively high probabilities of observing wet (0.346) and dry conditions (0.375) at time t+1 whereas relatively dry and dry conditions at time t has the high transition probabilities with a dry condition (0.679 and 0.753) at time t+1. Overall, the surface wetness conditions at time t had the relatively high transition probabilities with wet and dry conditions at time t+1. It is inferred that rainfall events for a single day (or few consecutive days) were usually generated rather than for a few (more than three or four) consecutive days during the given period. It indicated that moisture on the soil surface at these sites quickly run dry since the short-term rainfall durations. Further, the direct transitions between states 2 and 3 were rare, which meant that states 2 and 3 play the role of intermediaries.

Table 6.4: Predicted transition probabilities for surface wetness states under different hydro-climatic regions

Study sites	State at time t	State at time t+1			
		State 1	State 2	State 3	State 4
Oklahoma domain	State 1	0.455	0.140	0.087	0.318
	State 2	0.346	0.094	0.185	0.375
	State 3	0.167	0.138	0.016	0.679
	State 4	0.134	0.077	0.036	0.753
Illinois domain	State 1	0.377	0.131	0.341	0.151
	State 2	0.246	0.294	0.067	0.393
	State 3	0.191	0.091	0.358	0.360
	State 4	0.096	0.090	0.165	0.649

State 1: wet condition; State 2: relatively wet condition; State 3: relatively dry condition; State 4: dry condition

Fig. 6.7 shows the comparison of observed rainfall occurrence probabilities and optimal state sequences (surface soil wetness) derived by the HMMGA during the simulation period (90 days from March 01 for each year during 1995 to 2009). Overall, the optimal state sequences had good match (R^2: 0.945) with the observed rainfall occurrence probabilities along the time series. This result showed that the derived sequences were capable of representing the surface soil wetness conditions. We present the reproduced rainfall occurrence probabilities corresponding to the four soil wetness states for all the weather stations in the Oklahoma domain in Fig. 6.8. The rainfall occurrence probabilities for state 1 (wet condition) and 4 (dry condition) were slightly higher and lower, respectively than those for states 2 and 3, but there were no apparent trends for four states. Also, the individual states usually had similar probabilities across the spatial domain as shown in the observations (0.26~0.29) Fig. 6.4a.

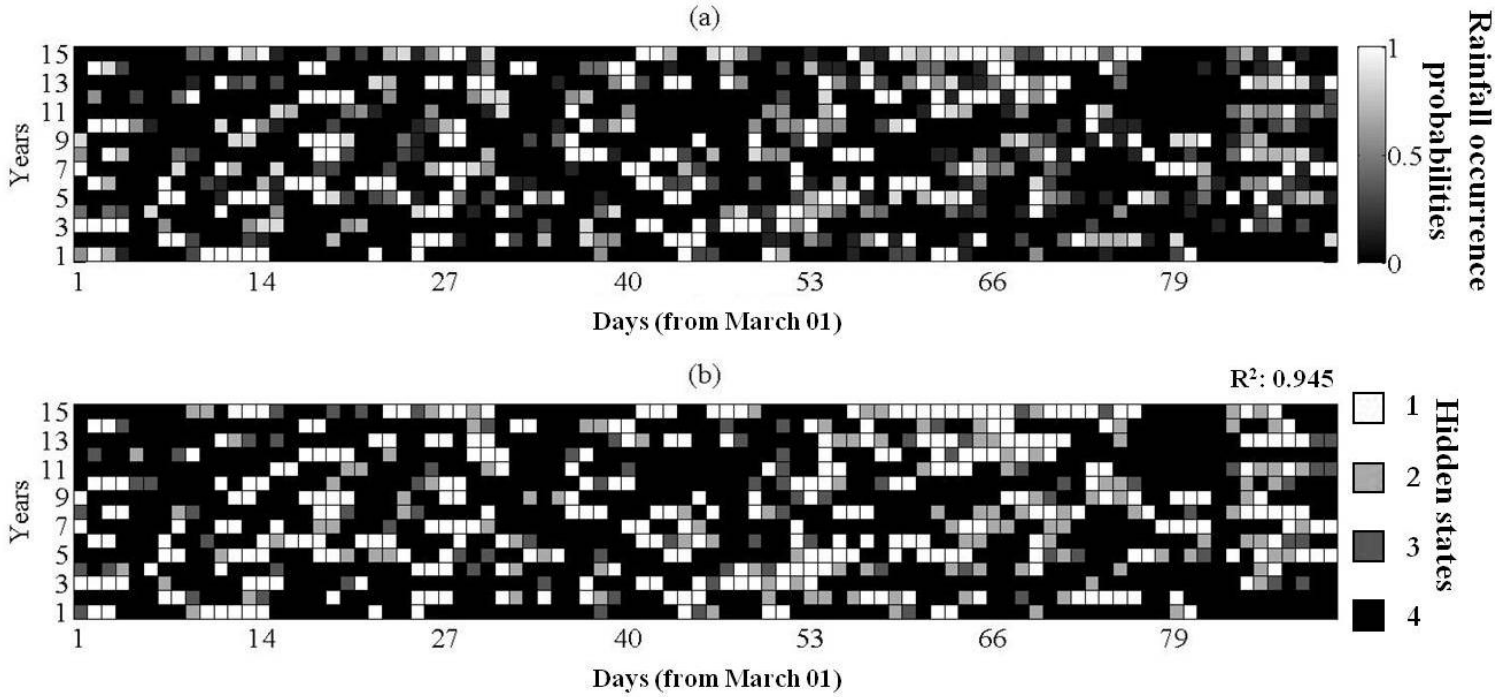

Figure 6.7: (a) The observed rainfall occurrence probabilities and (b) optimal state sequences derived by the HMMGA (representing soil wetness conditions comprised of hidden states 1 to 4) during 1995~2009 (years 1 to 15) in Oklahoma

State 1: wet condition; State 2: relatively wet condition; State 3: relatively dry condition; State 4: dry condition

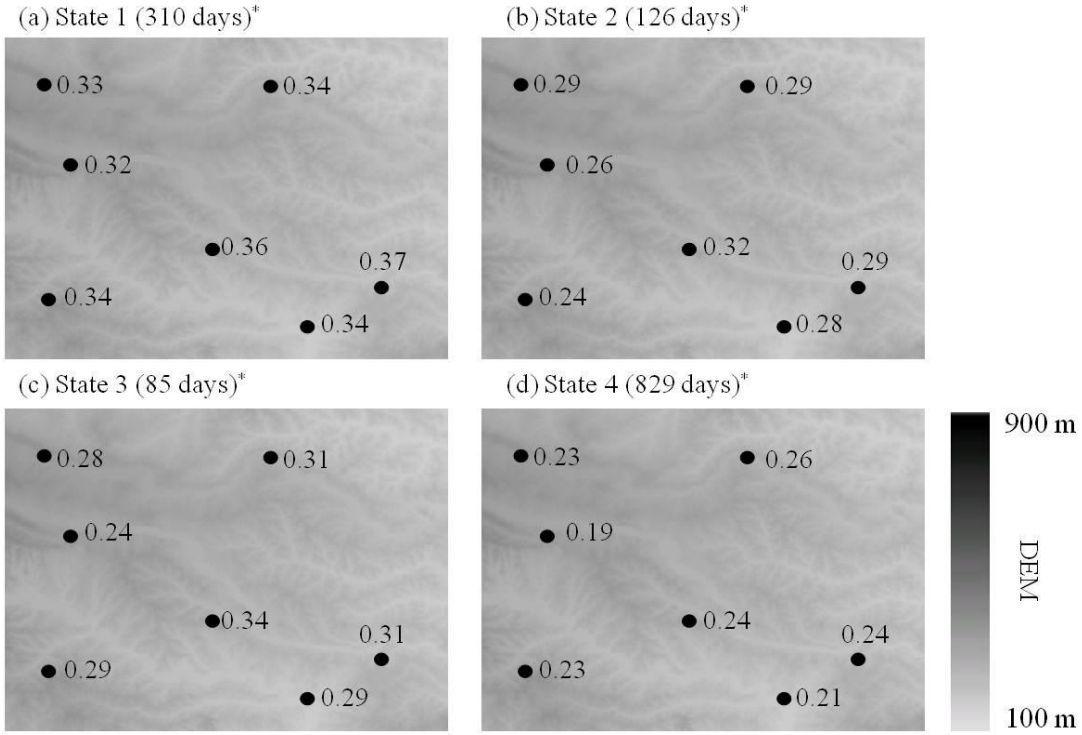

Figure 6.8: Predicted four-state rainfall occurrence probabilities together with digital elevation model (DEM) for the Oklahoma domain

State 1: wet condition; State 2: relatively wet condition; State 3: relatively dry condition; State 4: dry condition
*(the total number of days for each state)

Fig. 6.9 shows the seasonal variations (daily average) of the four hidden wetness states for all (1995-2009) years. The probabilities of dry condition for state 4 were considerably higher than those of wet, relatively wet, and relatively dry conditions indicating that the Oklahoma area had significant durations of dry conditions during the simulation period (90 days from March 01). The probabilities of wet condition for state 1 tend to be contrary to those of state 4. The probabilities of states 2 and 3 show relatively irregular trends, because relatively wet and dry conditions played a role as an intermediary of transitioning from a wet condition to a dry condition (or from dry to wet) as shown in Table 6.4. The probabilities of four states had no apparent seasonality indicating that the rainfall events were evenly generated along the time series.

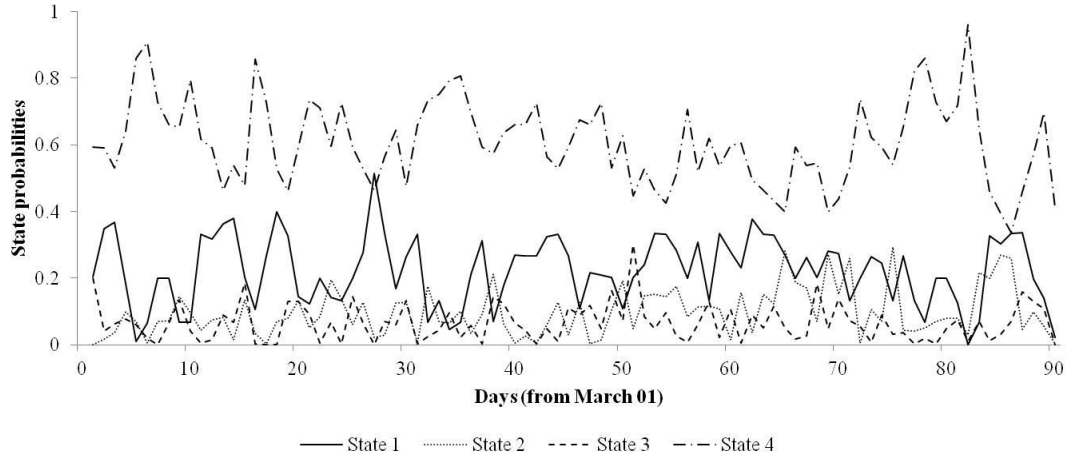

Figure 6.9: Seasonal variations (daily average from 1995 to 2009) of four hidden-states at the Oklahoma domain

State 1: wet condition; State 2: relatively wet condition; State 3: relatively dry condition; State 4: dry condition

Fig. 6.10a-g shows the comparison of yearly observed and simulated rainfall occurrence probabilities for the individual sites. The yearly simulated probabilities matched fairly well with the observed datasets (R^2: 0.460~0.676, RMSE: 0.074~0.084). However, when we compared the observed and predicted rainfall occurrence probabilities (averaged for all years) for the individual weather stations at the entire spatial domain (Fig. 6.10h), the results (R^2: 0.784 and RMSE: 0.012) were highly correlated. This implies that our proposed scheme could predict long term rainfall occurrences well at both the spatial and temporal scales, especially better for the spatial domain. In Fig. 6.11, the frequency distributions of dry (R^2: 0.931~0.993 and RMSE: 6.387~9.980) and wet (R^2: 0.945~0.995 and RMSE: 6.458~14.381) spell-lengths estimated by the HMMGA matched well with the observations, although small uncertainties exist. Overall, the correlations (R^2) of dry and wet spell lengths for each weather station were similarly shown, but the wet spell lengths had more uncertainties than those of the dry spell lengths.

Similarly, we predicted the time averaged (2003 to 2009) root zone soil moisture (0-5 cm) for the individual sites, as shown in Fig. 6.12. The predicted root zone soil moisture matched quite well with the measurements (R^2: 0.999 and RMSE: 0.005). The observed and estimated soil moistures for the sites 2, 4, 5, and 7 were relatively lower than those of sites 1, 3, and 6. When we compared the (observed/simulated) rainfall occurrence probabilities with the root zone soil moisture predictions, the soil moisture estimates did not correspond to the rainfall patterns. This supported that the soil moisture

was affected by not only the weather conditions, but also by the land surface conditions (i.e., soil textures, vegetation covers, topography, etc.).

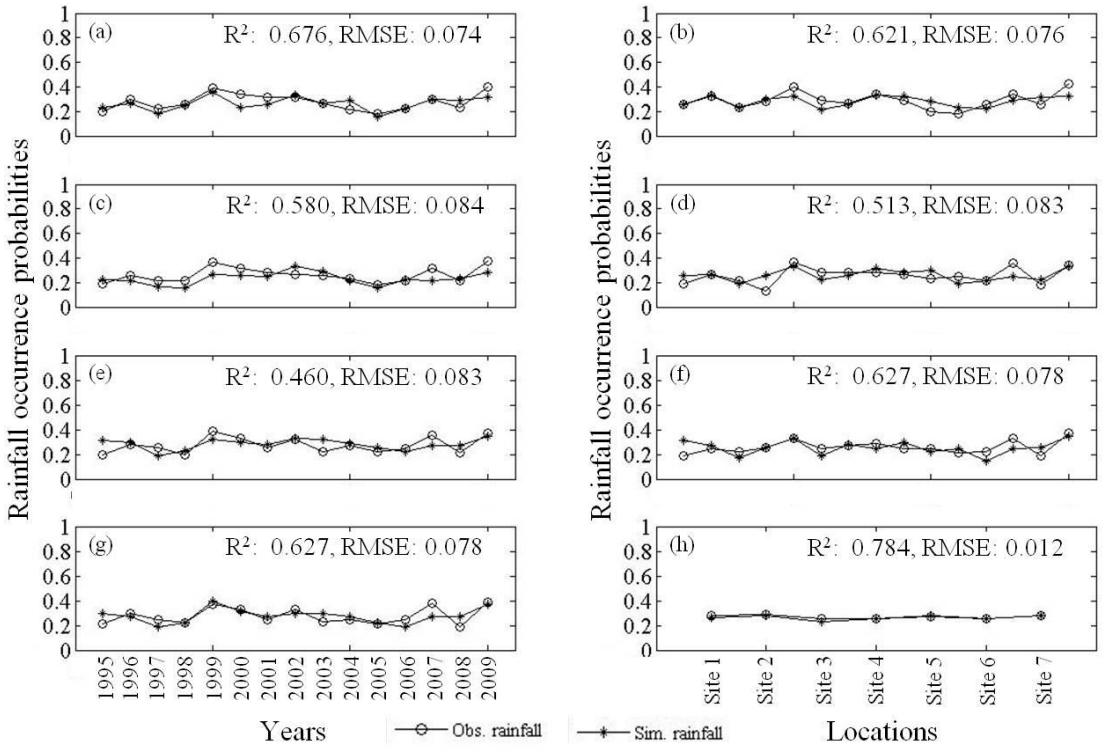

Figure 6.10: (a-g) Yearly observed and simulated rainfall occurrence probabilities (1995-2009) for each weather station (sites 1 to 7) at the temporal domain and (h) average probabilities of observed and simulated rainfall occurrences for all years at the spatial domain in the Oklahoma domain

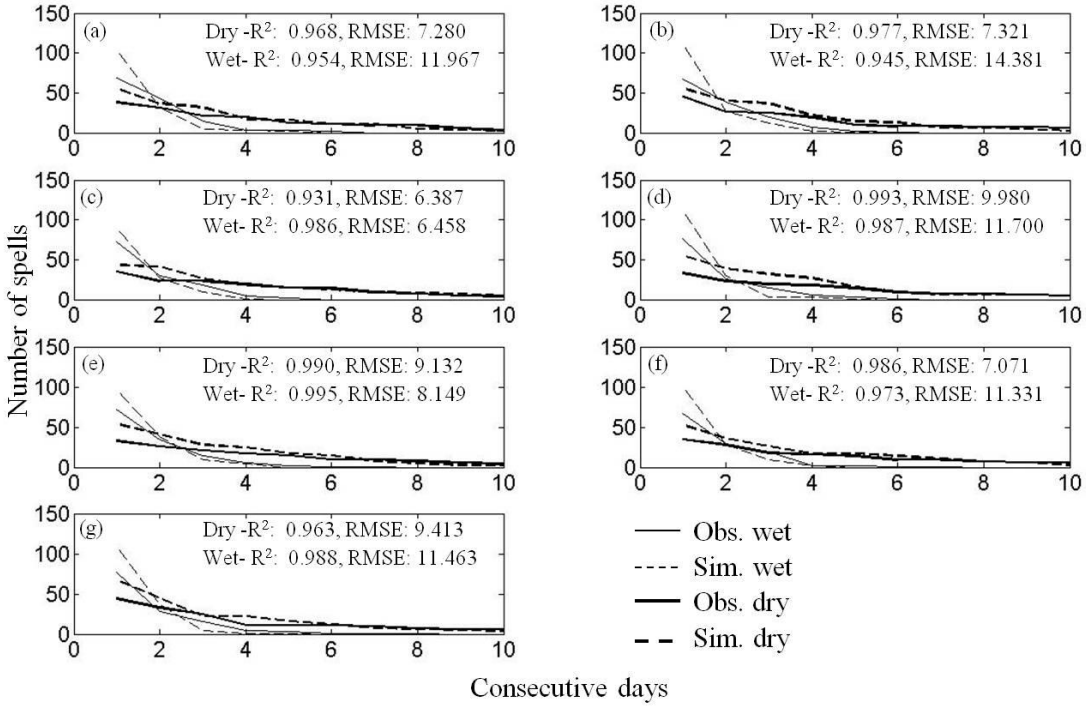

Figure 6.11: Observed and simulated dry/wet spell lengths in the network of multiple weather locations (a-g: sites 1 to 7) for the Oklahoma domain (a-g: sites 1 to 7) during 1995-2009

Total number of days: 1350 (15 years × 90 days).

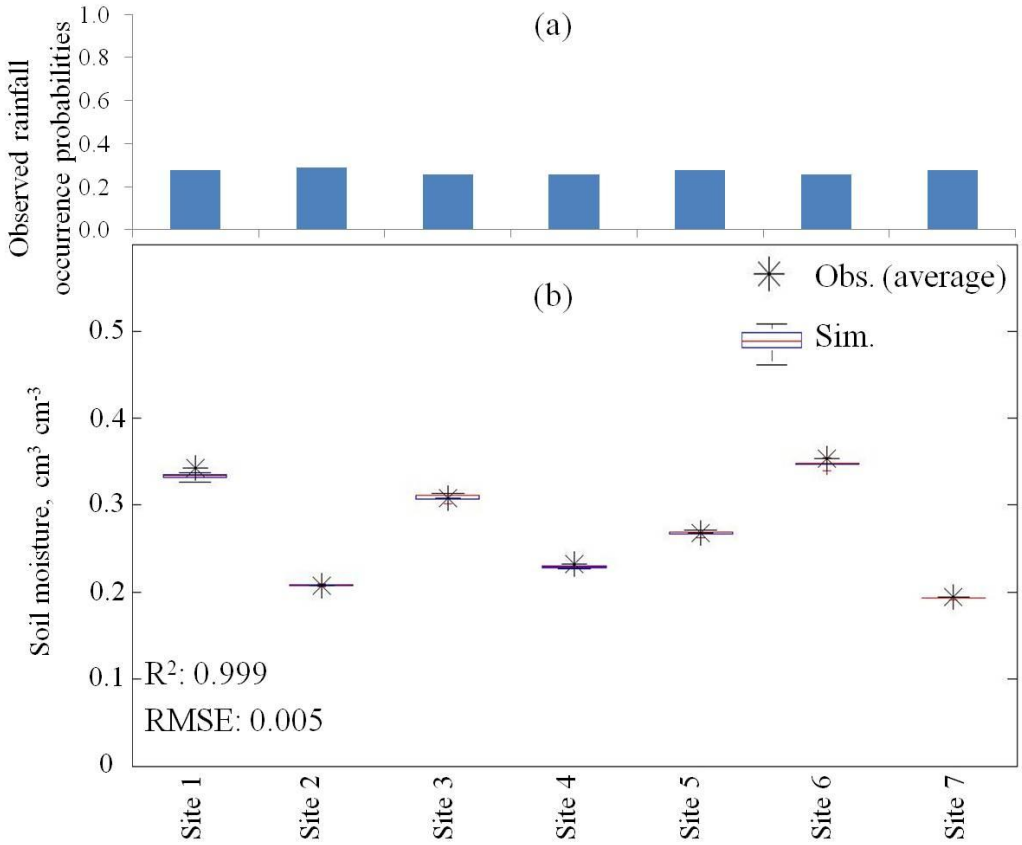

Figure 6.12: (a) Observed rainfall occurrence probabilities (averaged for all years) and (b) measured and predicted root zone soil moisture values* (averaged for all years, cm^3 cm^{-3}) at the soil depth of 0-5 cm in the network of multiple weather locations for the Oklahoma domain during the simulation period (2003-2009)

*On each box indicating the simulated soil moisture, the central line is the median value of data, the edges of the box are the 25th and 75th percentiles, and the whiskers extend to the most extreme data points.

6.4.2 Illinois Domain

We tested the HMMGA approach for the Illinois domain (300 km × 500 km). The (hidden) state transition probabilities given in Table 6.4 show slightly different behavior compared to the results of the Oklahoma domain.

A wet condition at time t had relatively high probabilities of observing wet (0.377) and relatively dry (0.341) conditions at time t+1. A relatively wet condition at time t was evenly transitioned to other conditions (0.246~0.393) at time t+1, except of a relatively dry condition (0.067). Given a relatively dry condition at time t, the probabilities of observing relatively dry (0.358) and dry (0.360) conditions at time t+1 were higher than others, but a dry condition at time t still had the highest self-transition probability (0.649) at time t+1. Overall, the probabilities of observing a dry condition at time t+1 in Illinois were lower than those of Oklahoma indicating that the Illinois sites had more rainfall events.

Fig. 6.13 exhibits the comparison of observed rainfall occurrence probabilities and estimated optimal state sequences. The estimated optimal state sequences matched quite well (R^2: 0.912) with the observed rainfall occurrence probabilities, similar to the results of the Oklahoma domain (shown in Fig. 6.7). This result indicated that the HMMGA performs well in predicting the surface soil wetness for the Illinois (relatively larger) domain also.

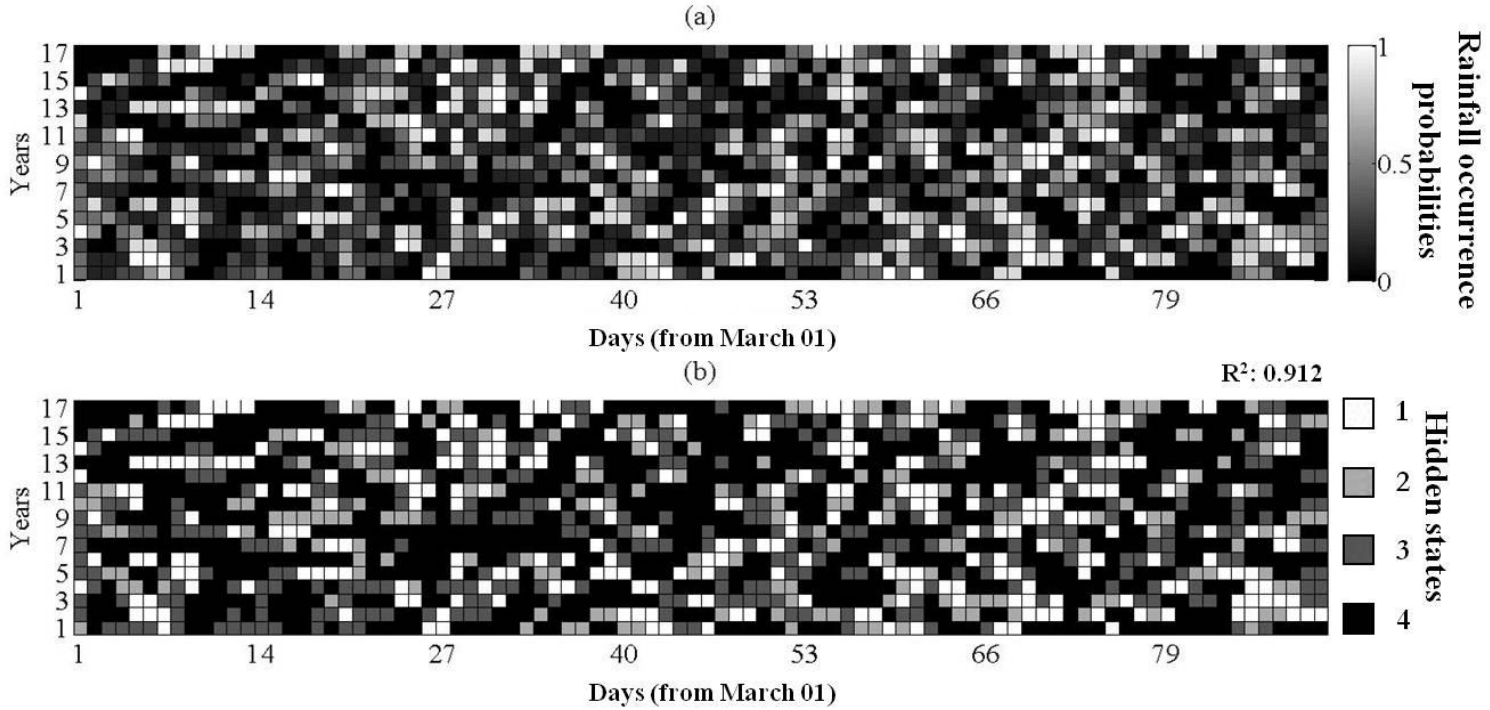

Figure 6.13: (a) The observed rainfall occurrence probabilities and (b) optimal state sequences derived by the HMMGA (representing soil wetness conditions comprised of hidden states 1 to 4) during 1994~2010 (years 1 to 17) over the multiple weather locations in Illinois

State 1: wet condition; State 2: relatively wet condition; State 3: relatively dry condition; State 4: dry condition

The rainfall occurrence probabilities at four different states were presented for the multiple sites of the Illinois domain in Fig. 6.14. In general, the results for four different states showed similarly across the study sites. However, the probabilities of rainfall occurrences under a dry condition (state 4) were lower than for the other states, which meant that less rainfall events were generated. Fig. 6.15 showed the seasonal variations (daily average from 1994 to 2010) of four hidden states. Overall, the variation characteristics of different states (1 to 4) in Illinois were similar as shown in the results of Oklahoma (Fig. 6.9) and the seasonality of the probabilities for four states was also elusive. The probabilities of dry condition (state 4) were slightly lower than those of Oklahoma indicating that the Illinois sites had relatively more rainfall events than those of the Oklahoma sites. The comparison of observed rainfall occurrence probabilities for the Oklahoma (Fig. 6.7a) and Illinois (Fig. 6.13a) domains supported these findings. Also, the state 4 probabilities were still higher than other states (1, 2, and 3).

In Fig. 6.16a-g, the yearly predicted rainfall occurrence probabilities (R^2: 0.102~0.436 and RMSE: 0.063~0.084) for the sites 1 to 5 derived by the HMMGA were less identifiable with the observations than those (R^2: 0. 460-0.676 and RMSE: 0.074-0.084) of the Oklahoma (relatively smaller) domain in Fig. 6.10a-g, especially for the sites 6 and 7 (R^2: -0.383 and -0.077, RMSE: 0.084 and 0.077). It may indicate that the predictions were considerably influenced by the distance between the weather stations. The time averaged probabilities (R^2: 0.962 and RMSE: 0.006) for all years in the individual sites were fairly matched with the observations (Fig. 6.16h) at the spatial domain, although the predicted rainfall probabilities were slightly underestimated. In

line with the results for the Oklahoma domain, it was found that the stochastic approach may be more robust in predicting long-term rainfall across the entire spatial domain rather than predicting the short term daily time series of rainfall at individual weather stations. The statistics of rainfall predictions have some uncertainties in modeling for the Illinois domain, but overall the predicted results tend to be similar with the observations. The simulated dry/wet spell lengths (R^2: 0.958~0.992 for the dry and R^2: 0.946~0.994 for the wet) by this approach matched well with the observations in Fig. 6.17. Also, the uncertainties (RMSE: 9.274~19.209) of wet spell lengths were slightly higher than those (RMSE: 7.937~12.661) of the dry condition as shown in the results of Oklahoma domain.

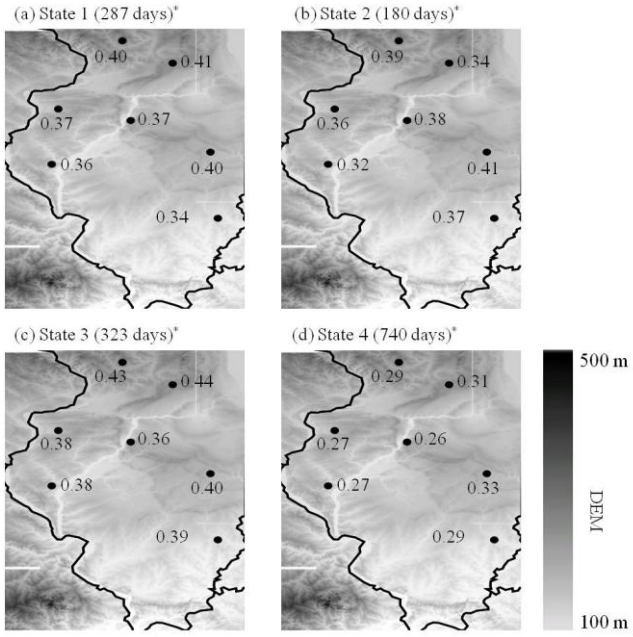

Figure 6.14: Predicted four-state rainfall occurrence probabilities together with digital elevation model (DEM) for the Illinois domain

State 1: wet condition; State 2: relatively wet condition; State 3: relatively dry condition; State 4: dry condition

*(the total number of days for each state)

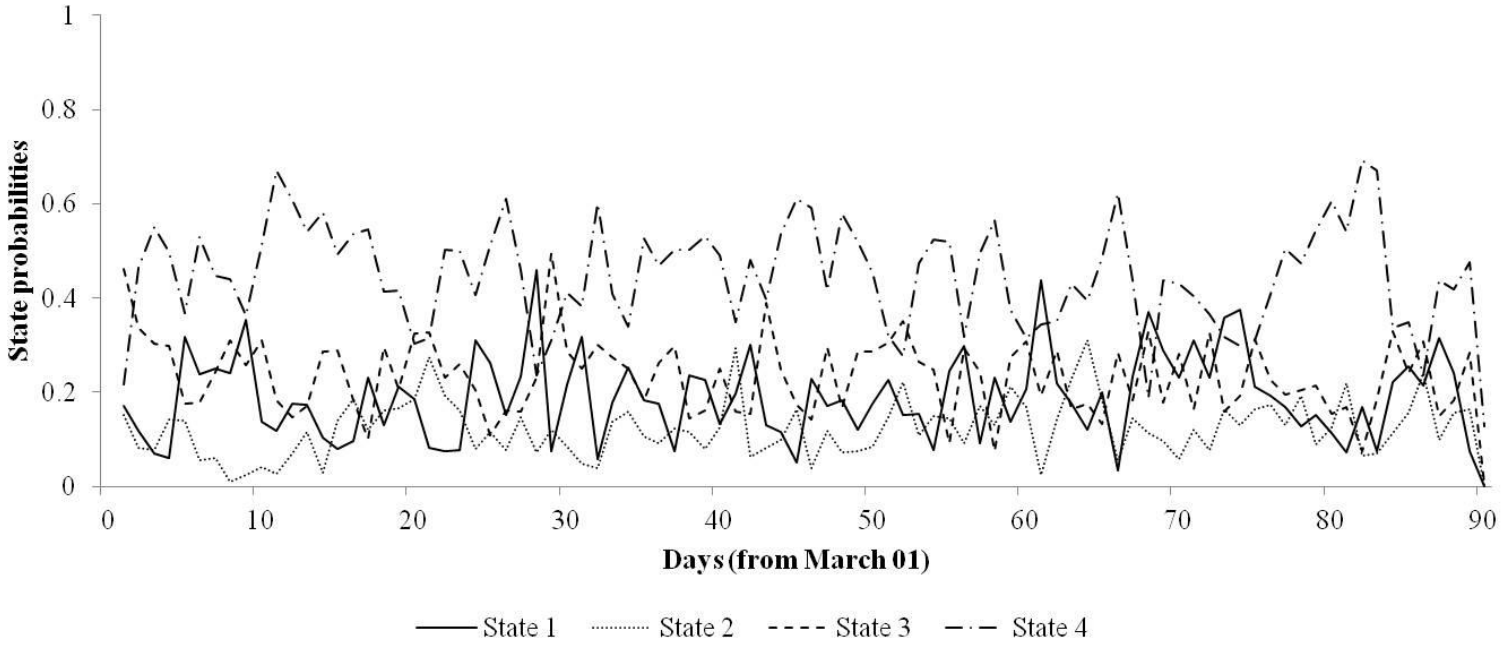

Figure 6.15: Seasonal variations (daily average from 1995 to 2009) of four hidden-states at the Illinois domain
State 1: wet condition; State 2: relatively wet condition; State 3: relatively dry condition; State 4: dry condition

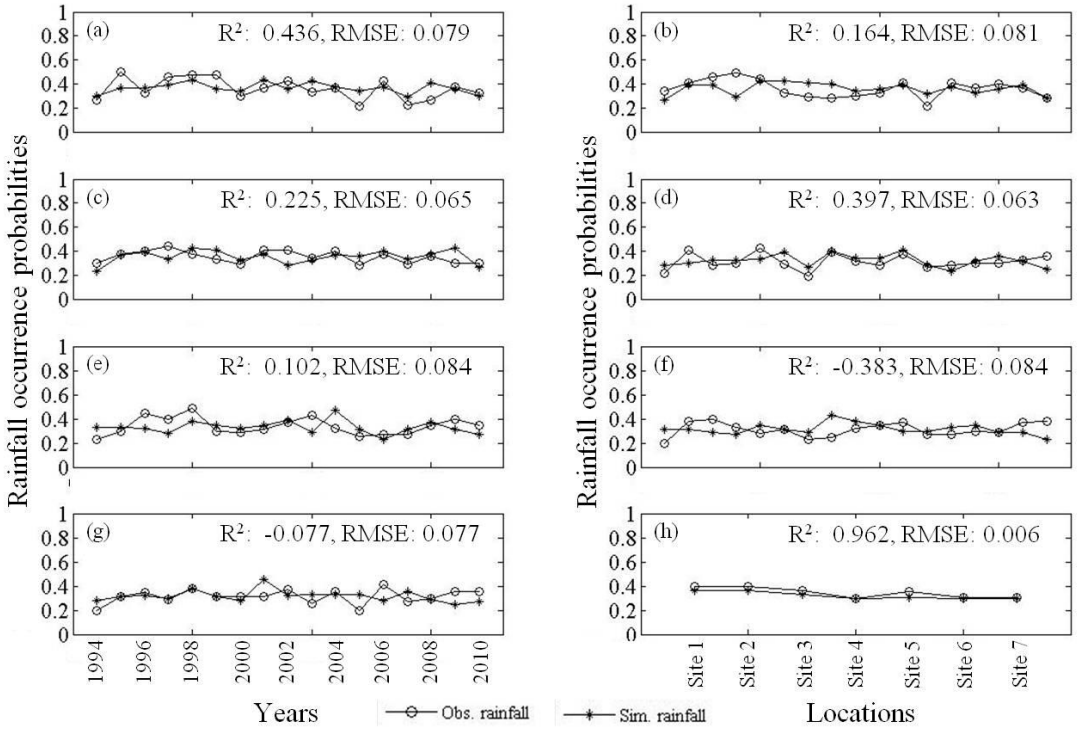

Figure 6.16: (a-g) Yearly observed and simulated rainfall occurrence probabilities (1994-2010) for each weather station (sites 1 to 7) at the temporal domain and (h) average probabilities of observed and simulated rainfall occurrences for all years at the spatial domain in the Illinois domain

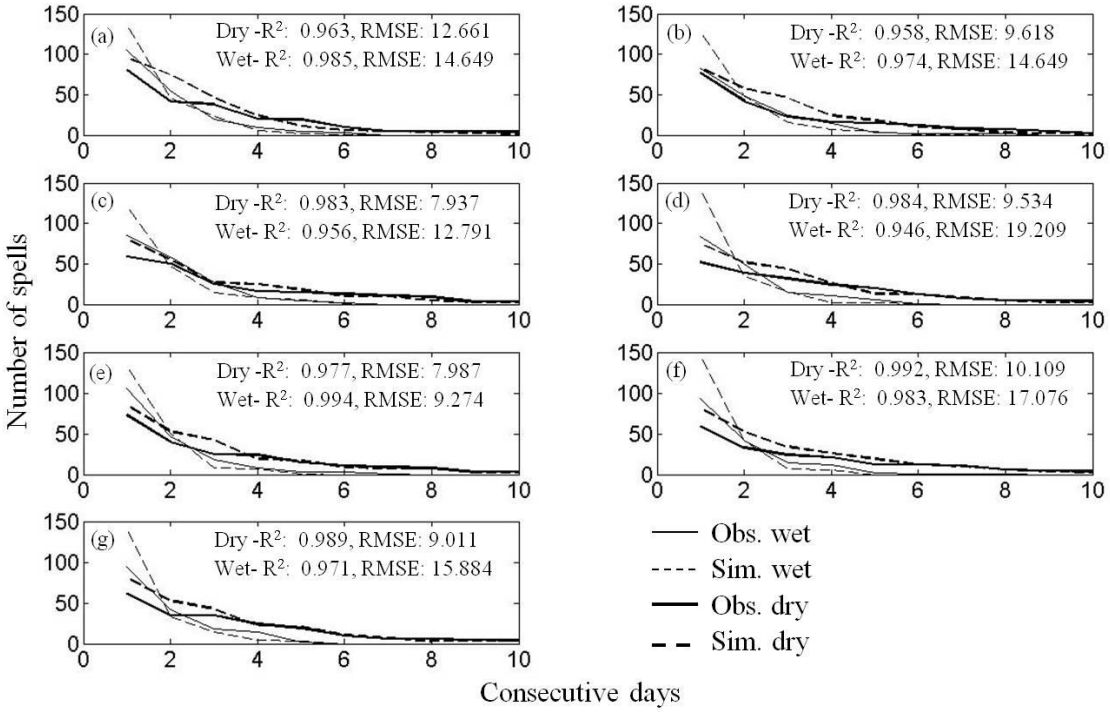

Figure 6.17: Observed and simulated dry/wet spell lengths in the network of multiple weather locations (a-g: sites 1 to 7) for the Illinois domain during 1994-2010

Total number of days: 1530 (17 years × 90 days)

Fig. 6.18 shows the daily measured and predicted root zone soil moisture estimates (cm^3 cm^{-3}) derived by our HMMGA approach and also from the data assimilation scheme [*Ines and Mohanty et al*., 2008a] at Site 1 (Bondville) in Illinois. Note that the daily comparison of observed and predicted root zone soil moisture dynamics in 1994 only is shown for the sake of brevity. The daily root zone (0-50 cm) soil moisture dynamics derived by the physical-based assimilation scheme were better estimated (R^2: 0.739 and RMSE: 0.027) than the HMMGA-based predictions (R^2: 0.433 and RMSE: 0.058). Although the stochastic HMMGA soil moisture estimates had uncertainties in modeling, the predictions matched well with the observations. This result shows that the stochastic approach can provide somewhat available root zone soil moisture predictions on a daily basis.

We predicted the time averaged (1994 to 2000) soil moisture at the near-surface (0-10 cm) and in the sub-surface layers (10-30, 30-50, and 0-50 cm) at each of the weather stations (Fig. 6.19). The predicted root zone soil moistures (R^2: 0.675~0.918 and RMSE: 0.024~0.041) at the various soil depths matched well with the measurements. The near-surface (0-10 cm) soil moisture estimates were more variable than those of the sub-surface layers. Mostly, the predicted root zone soil moisture values have uncertainties compared to the measurements, especially for the sites 2 and 3. These uncertainties may be caused by the underestimations of predicted rainfall occurrence probabilities (Fig. 6.16h).

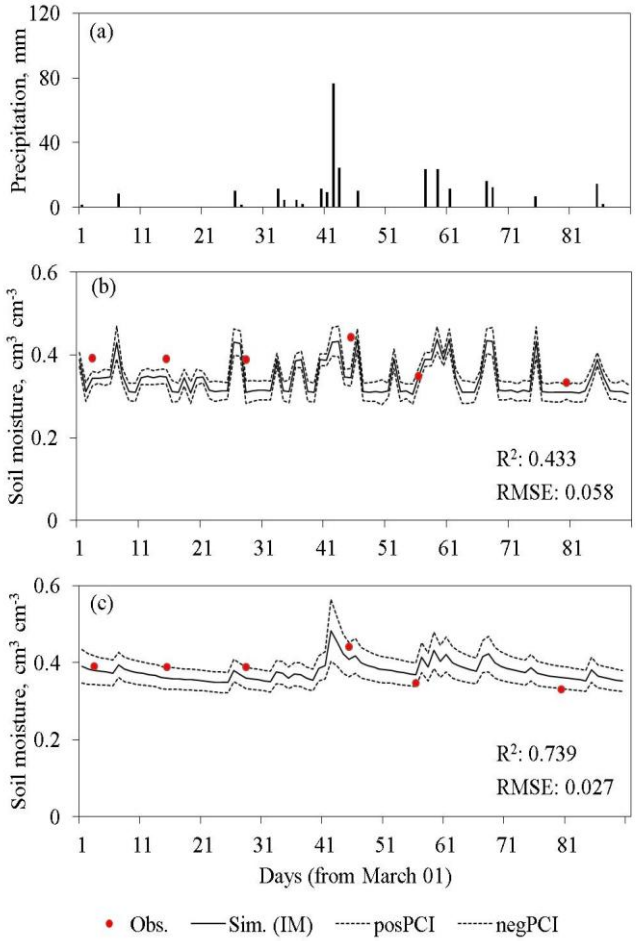

Figure 6.18: Comparison of the measured and predicted soil moisture at Illinois Site 1 in 1994 (90 days from March 01 to May 29); (a) rainfall in Bondville (site 1), (b) predicted near-surface (0-50 cm) soil moisture dynamics derived by the HMMGA approach[a], and (c) Simulated near-surface (0-50 cm) soil moisture dynamics derived by the near-surface data assimilation scheme[b]

[a]Realizations=30; Random number generator seeds (3000); Max. generation=30; Population=10
[b]Random number generator seeds (-1000, -950, and -750); Max generation=500; Population=10

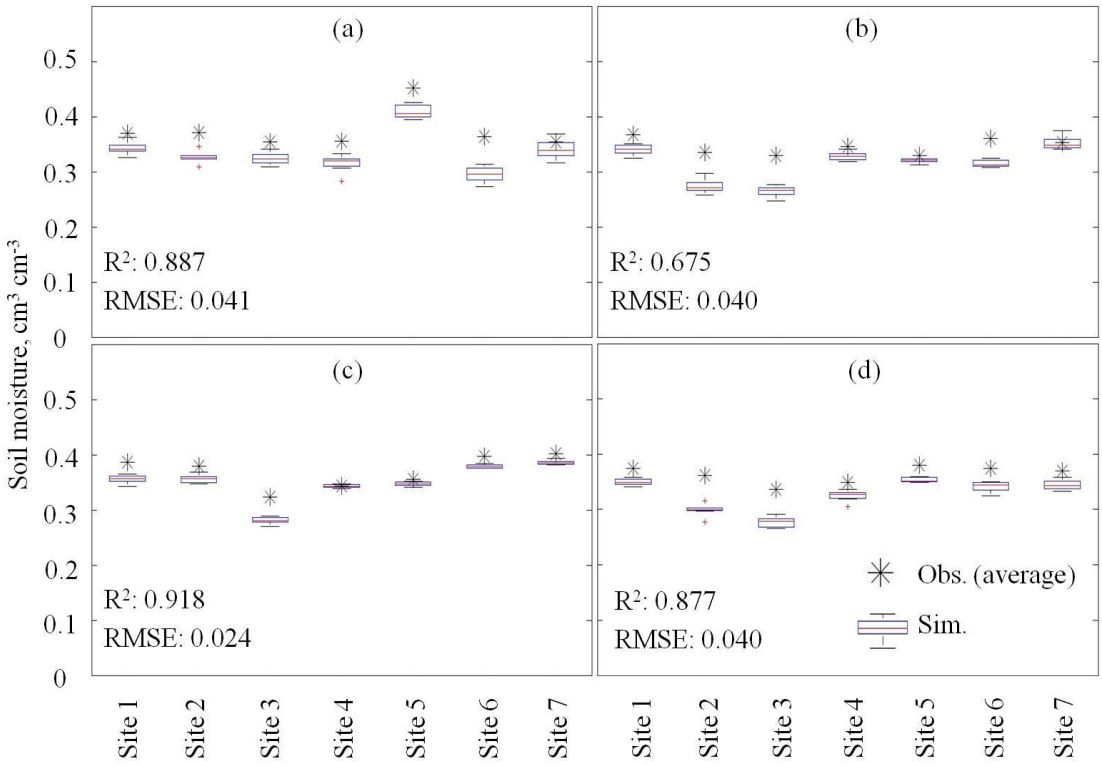

Figure 6.19: Measured and predicted root zone soil moisture dynamics* (averaged for all years, $cm^3\ cm^{-3}$) at various soil depths in the network of multiple weather locations for the Illinois domain during the simulation period (2003-2009); (a) 0-10 cm, (b) 10-30 cm, (c) 30-50 cm, and (d) 0-50 cm

*On each box indicating the simulated soil moisture, the central line is the median value of data, the edges of the box are the 25th and 75th percentiles, and the whiskers extend to the most extreme data points.

Figure 6.20 shows the calibration (1994-2000) and validation (2001-2004) results for soil moisture (the soil depth of 0-50 cm). Overall, the predicted root zone soil moisture estimates (R^2: 0.310 and RMSE: 0.051) during the validation period matched the observations as well as those (R^2: 0.877 and RMSE: 0.040) for the calibration (1994-2000), while the soil moisture predictions of sites 2 and 3 for the validation have more bias, especially for site 2 causing the relatively low correlation. These findings are quite important in this study, because our proposed methodology can provide an attractive alternative in forecasting long-term root zone soil moisture for the future using only climate change scenarios (i.e., global climate models-GCMs, regional climate models-RCMs, etc.) and historical soil moisture measurements (available). In this regard, this approach could be quite useful for developing/improving efficient agricultural/water resources management in the future.

Generally, the root zone soil moisture predictions for the Illinois domain had more uncertainties than those of the Oklahoma (relatively smaller) domain indicating that the stochastic soil moisture predictions may be considerably influenced by the spatial domain size. Although the results for the Illinois domain had more bias compared to the measurements, the HMMGA approach still predicted the root zone soil moisture well across the spatial domain with some uncertainties.

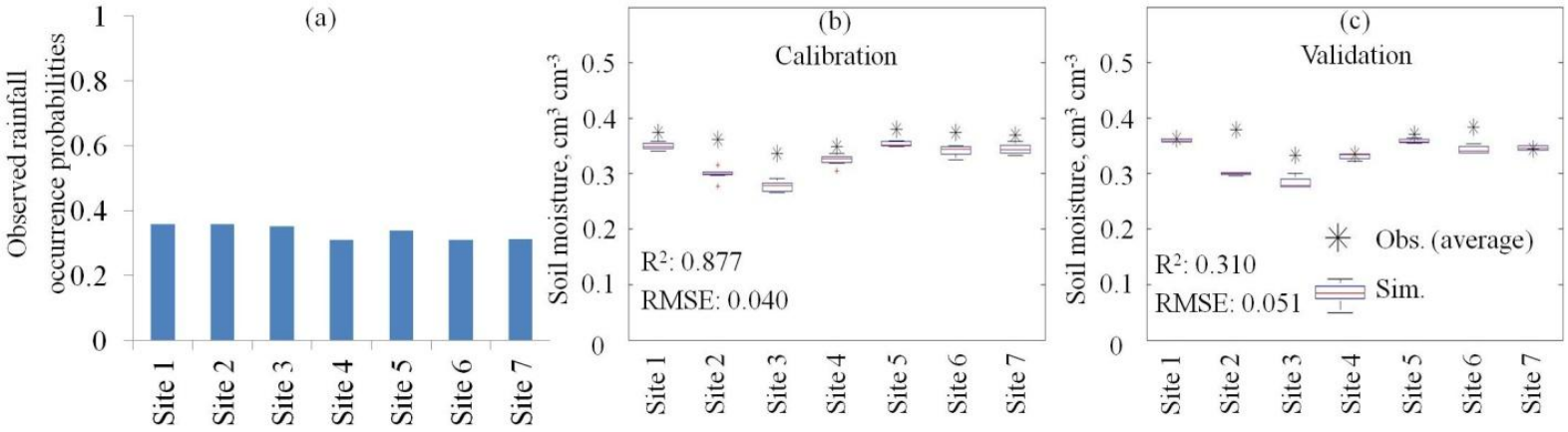

Figure 6.20: Measured and predicted root zone soil moisture dynamics* (averaged for all years, cm^3 cm^{-3}) at the soil depth (0-50 cm) in the network of multiple weather locations for the Illinois domain during the calibration and validation periods; (a) observed rainfall occurrence probabilities, (b) calibration (1994-2000), and (c) validation (2001-2004)

*On each box indicating the simulated soil moisture, the central line is the median value of data, the edges of the box are the 25th and 75th percentiles, and the whiskers extend to the most extreme data points.

The data assimilation scheme using the physical-based hydrological (SWAP) model also has uncertainties (i.e., due to the model structure, parameter uncertainties, user based errors, etc.). When we considered these limitations mentioned above and complexity of applications, our suggested approach can provide an easy application for root zone soil moisture predictions using only precipitation and historical soil moisture measurements existing in the real world conditions.

6.5 Conclusions

In this study, we developed a non-parametric evolutionary algorithm (genetic algorithm-based hidden Markov model, HMMGA) for predicting root zone soil moisture dynamics using only precipitation and historical soil moisture measurements across multiple weather stations. Rainfall occurrence probabilities and dry/wet spell lengths reproduced by the HMMGA were also analyzed.

The HMMGA was used to estimate optimal state sequences (representing surface soil wetness comprised of the four hidden states-wet, relatively wet, relatively dry, and dry conditions) based on precipitation history. Historical soil moisture data measured at different soil depths (0-5, 0-10, 10-30, 30-50, and 0-50 cm) were categorized based on the soil wetness conditions. Then, statistics (mean and standard deviation) of categorized soil moisture data were determined. Using the HMMGA-based surface soil wetness and historical categorized soil moisture statistics, root zone soil moisture dynamics were estimated.

Two different domain sizes in Oklahoma (1995-2009, 130 km × 130 km) and Illinois (1994-2010, 300 km × 500 km) under different hydro-climatic locations were selected for testing this approach.

The HMMGA estimated well the optimal state sequences for the Oklahoma (R^2: 0.945) and Illinois (R^2: 0.912) domains compared to the observed rainfall occurrence probabilities. The yearly predicted rainfall occurrence probabilities at the individual weather stations for the Oklahoma domain had a better match (R^2: 0.460~0.676, RMSE: 0.074~0.084) with observations, but the Illinois domain had more variations (R^2: -0.383~0.436, RMSE: 0.063~0.084) indicating that the HMMGA processes are dependent on the distance between the weather stations.

The average rainfall occurrence predictions for all years across the individual sites matched the observations well (R^2: 0.784 and RMSE: 0.012 for the Oklahoma domain, R^2: 0.962 and RMSE: 0.006 for the Illinois domain). This result showed that the HMMGA provided better long-term predictions across a spatial domain than short term (daily) predictions. Also, the estimated dry/wet spell lengths have some uncertainties with the observations for both the domains, but they still matched well.

The average near-surface (0-5 cm) soil moisture dynamics for all seven years at the Oklahoma domain matched almost perfectly (R^2: 0.999 and RMSE: 0.005) with the measurements. For the Illinois domain, the daily root zone (0-50 cm) soil moisture predictions (R^2: 0.433 and RMSE: 0.058) at the site 1 (Bondville, 1994) were identifiable with the observations. Also, the average root zone soil moisture dynamics at the different soil depths (0-10, 10-30, 30-50, and 0-50 cm) performed fairly well (R^2:

0.675~0.918 and RMSE: 0.024~0.041) compared to the observations, although uncertainties were evident in the predictions.

Generally, the soil moisture predictions had slightly different trends compared to the rainfall occurrence probabilities for both domains. These findings showed that the predicted root zone soil moisture estimates were influenced by not only the precipitation, but also the land surface conditions (e.g., topography, soil textures, vegetation covers, etc.). Thus, this scheme which uses rainfall and historical soil moisture data performed well in predicting root zone soil moisture.

The results of calibration and validation supported the robustness of this approach in forecasting the root zone soil moisture dynamics at multiple locations. This result indicates that our approach can be used to predict root zone soil moisture dynamics in the future using only climate change scenarios and soil moisture history existing.

Based on these findings, our proposed methodology can predict well not only the rainfall occurrences and dry/wet spell lengths, but also the root zone soil moisture estimates at the multiple sites. Further, as the predicted results at the spatial domain matched quite well with the observations, this approach can be also used for downscaling both remotely sensed precipitation [*Robertson et al.*, 2004] and soil moisture products. Thus, the newly developed algorithm could provide an attractive alternative for agricultural/water resources management in an efficient and economical manner.

CHAPTER VII

DEVELOPMENT OF A DROUGHT SEVERITY ASSESSMENT FRAMEWORK USING REMOTELY SENSED SOIL MOISTURE PRODUCTS UNDER CLIMATE CHANGE SCENARIO

7.1 Synopsis

Evaluating drought severity based on future climate scenarios plays an important role for water resources management. In this study we assessed drought severity based on soil moisture for individual soil-crop combinations. Based on the historical data, pixel-scale hydraulic parameters at finer-scales were estimated from remotely sensed (RS) soil moisture using a newly developed algorithm EMOGA (Ensemble Multiple Operators Genetic Algorithm) coupled with Soil-Water-Atmosphere-Plant (SWAP) hydrological model. These estimated hydraulic parameters along with meteorological variables obtained from general circulation models (GCMs) were used to predict soil moisture using the SWAP model. Further, drought severity was calculated using a soil moisture deficit index (SMDI) based on the projected soil moisture obtained from the SWAP model. The proposed model was evaluated based on synthetic and field data under different hydro-climates (Lubbock, Texas; Little Washita watershed, Oklahoma; Walnut Creek watershed, Iowa). Finer-scale root zone soil moisture predictions were considerably influenced by various combinations of environmental factors (soils, crops, groundwater table, etc.) along with GCM scenarios. However, these local environmental factors had relatively limited impacts (compared to precipitation dynamics) on reducing drought severity in the study region. The absolute SMDI values do indicate the

occurrence of agricultural drought during 2010-2020. Thus, our proposed approach can be used to assess drought severity at finer-scales using a RS soil moisture product for efficient agricultural/water resources management.

7.2 Introduction

Drought is one of the most severe environmental outcomes of the climate change. It impacts agriculture, ecosystem, food security, etc. leading to economic loss. According to the U.S. Federal Emergency Management Agency [*FEMA*, 1995], annual loss due to drought is approximately U.S. $ 6-8 billion. In 2011, Texas suffered under an intense drought driven by La Niña with a total damage of $ 7.6 billion [*LBB*, 2011]. Generally, droughts develop and evolve gradually with time when compared to other natural disasters (i.e., earthquake, tornado, etc.) indicating the possibility of making appropriate mitigation plans through drought prediction and monitoring systems [*Cancelliere et al*., 2007].

Drought evaluations have been conducted by using precipitation frequency and intensity based on drought indices. Many global circulation models (GCMs) have been developed/improved to simulate earth's climate change [*Roeckner et al.*, 1996; *Gong et al*., 2003; among others]. The Intergovernmental Panel on Climate Change (IPCC) provides various meteorological, socioeconomic, and environmental data based on the past and future scenarios for the assessment of climate change [*IPCC*, 2007]. Goddard Institute for Space Studies (GISS), National Aeronautics and Space Administration (NASA) also provides various GCM versions (e.g., Model E, Model AOM-GR and Model II). A number of climate models from Canadian Centre for Climate modeling and

analysis (CCCma) are available for understanding climate change and variability in various Earth system processes. Despite better understanding in climatic forecasts, hydrologists still struggle to improve coarse-scale GCM scenarios in applications (e.g., water resources planning, water management, etc.). Various studies were conducted for developing downscaling algorithms to best reproduce realistic predicted values of climate data. The widely used approaches are dynamic modeling such as a regional climate model (RCM) [*Leung et al*., 2004] within GCM scenarios, statistical or empirical transfer functions [*Hewitson and Crane*, 1996; *Wilby and Wigley*, 1997, *Wilby et al*., 1998], etc. Further, the frequency and intensity distribution of GCM daily [*Ines and Hansen,* 2006] and monthly [*Wood et al*., 2002] precipitation scenarios were corrected using a quantile-based mapping [*Cayan et al*., 2008; *Hayhoe et al*., 2004; *Maurer and Duffy*, 2005; *Maurer*, 2007; *Panofsky and Brier*, 1968].

Root zone soil moisture dynamics is a key factor in agricultural water management, rainfall-runoff processes, seasonal climate predictions, and ecosystem dynamics and alteration. Its spatio-temporal variability in a field influences runoff at the soil surface and in the sub-surface, evapotranspiration, and groundwater recharge. In this regard, importance of root zone soil moisture prediction has increased for agricultural water management, because of unbalanced water supply and demand due to agricultural drought, differences in irrigation and drainage practices, farm level water distribution and scheduling, etc.

In order to evaluate a drought condition, various drought indices such as the Palmer Drought Severity Index (PDSI) [*Palmer*, 1965], Crop Moisture Index (CMI)

[*Palmer*, 1968], Standardized Precipitation Index (SPI) [*McKee et al*., 1993], Surface Water Supply Index (SWSI) [*Shafer and Dezman*, 1982], Soil Moisture Deficit Index (SMDI) [*Narasimhan and Srinivasan*, 2005], and Evapotranspiration Deficit Index (ETDI) [*Narasimhan and Srinivasan*, 2005], among others were designed for water resources management decision-making. Two of the most widely used drought indices are the PDSI and the SPI. The PDSI is primarily a climate-based drought index for assessing long periods under abnormally wet and dry conditions. However, crops are usually influenced not only by climatic variability, but also by soil water content in the root zone governed by bottom boundary conditions (shallow ground water table) [*Jobbágy and Jackson*, 2004; *Nosetto et al*., 2009; *Pollacco et al.*, 2012]. In this regard, the SMDI based on weekly soil moisture deficit [*Narasimhan and Srinivasan* 2005] could be more suitable for assessing agricultural drought severity.

A remote sensing (RS) scheme has some advantages for mapping distributed root zone soil moisture at the spatial domain. Many airborne- and satellite-scale RS data sets have been used to observe pixel-based (~m to tens of km scale) soil moisture products [*Jackson et al*., 1999; *Jackson et al*., 2005]. However, these RS datasets suffer from their coarse spatial resolutions to effectively be used for agricultural water management at local-/field-scales. *Crow et al.,* [2000] downscaled satellite-based RS soil moisture products for obtaining soil dielectric values by using a soil dielectric inversion model. A new downscaling algorithm was developed by *Merlin et al.,* [2005] to improve RS near-surface soil moisture products at fine scale for the Soil Moisture and Ocean Salinity (SMOS) mission [*Kerr et al.,* 2001] with fine-scale optical data. Recently, *Ines et al.,*

[2012] developed a stochastic disaggregation method to derive soil moisture for various land covers (representing soil properties and area fractions of corresponding soil-vegetation combinations) within a RS soil moisture footprint using a genetic algorithm.

Historically, many studies have been conducted with respect to interaction between soil moisture variations and climate changes at the lower boundary of the atmosphere [e.g., *Charney et al*., 1977; *Walker and Rowntree*; 1977; *Srinivasan, et al*., 1995; *Yang and Lau*, 1998; *Shen et al*., 1998], sensitivity of a GCM simulation to land surface hydrology [*Stamm and Wood,* 1994], among others. However, till date studies have not been designed to evaluate agricultural drought severity using root zone soil moisture predictions based on weather forecast models. The scale incompatibility between hydrologic and weather models leads to errors at the catchment-, watershed-, and basin-scales for predicting root zone soil moisture. To address proper water balance in the root zone, we need to develop better long-term root zone soil moisture prediction schemes at localized regions by scaling down coarse-scale RS data.

In this study, we used root zone soil moisture forecast (at fine spatial scale) coupled with a local land surface parameterization approach. We adapted a grid-based disaggregation algorithm [*Ines et al*., 2012] with a new simulation-optimization scheme, Genetic Algorithm (GA) and Soil-Water-Atmosphere-Plant (SWAP) model [*van Dam et al.,* 1997] to predict root zone soil moisture for the current decade (2010-2020). The objectives of this research were two-fold: (1) to develop a drought severity assessment framework using remotely sensed soil moisture products with global circulation model (GCM) based climate forecasts under various regional hydro-climatic conditions; and (2)

to evaluate local environmental factors such as various soil textures, vegetations, soil depths, shallow ground water tables, etc. on reducing drought severity at fine-scales across local regions. This approach will greatly assist in assessing potential risk (e.g., agricultural drought) and correspondingly in determining most efficient agricultural/water management practices during the current decade at functional/operational scales rather than at large remote sensing footprints.

7.3 Materials and Methods

As mentioned above, the first objective of this study was to develop a drought severity assessment framework using remote sensing soil moisture products for a GCM scenario. We modified a GA-based disaggregation algorithm [*Ines et al*., 2012] to extract land surface information comprising of soil-vegetation combinations within a RS footprint. The modified algorithm, described in detail (section 2.1), was grid-based and enabled incorporation of gridded RS data. This new scheme also determined the uncertainties associated with extracting land surface information and improved the efficiency of the GA. Using the extracted soil-vegetation information with GCM scenarios, long-term daily root zone soil moisture dynamics were predicted at fine-scales with the help of a hydrological model as depicted in Fig. 7.1a. Then, we evaluated drought severities with an appropriate drought severity assessment index under various environmental conditions (e.g., soil textures, vegetation covers, presence of shallow water tables, etc.) at local-scales. The modified grid-based algorithm is described below.

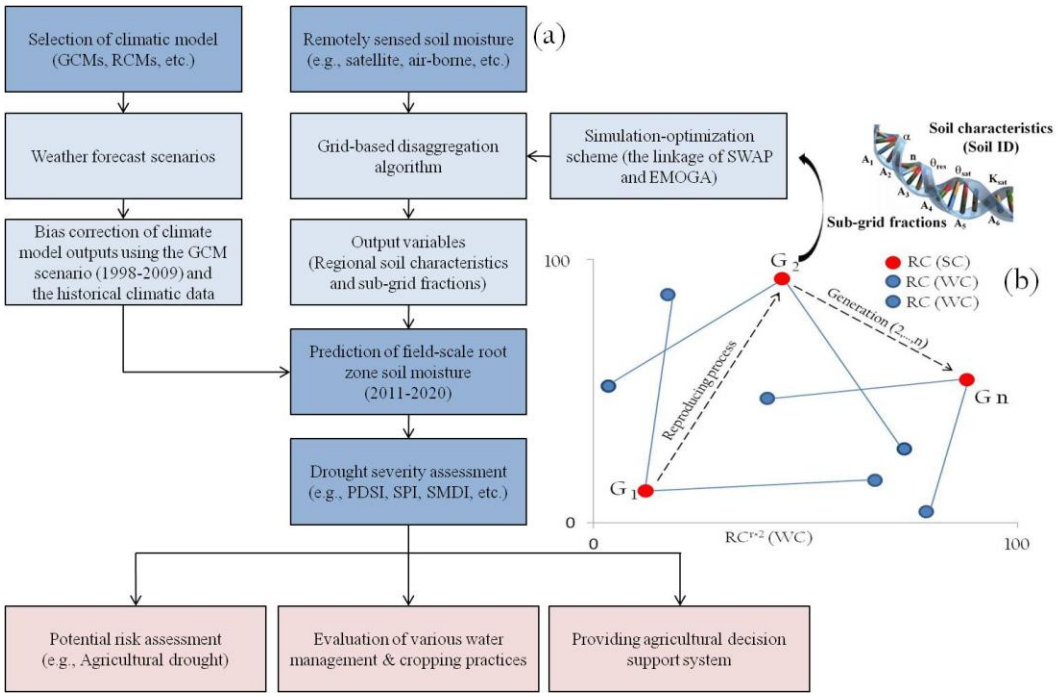

Figure 7.1: (a) Schematic of drought severity assessment framework using climate changes and a modified grid-based disaggregation algorithm; (b) basic concept of Ensemble Multiple Operators Genetic Algorithm (EMOGA)

Note: RC-reproduced chromosomes, SC-strongest chromosome, WC-weak chromosome, and Gn-generations

7.3.1 Modified Grid-based Disaggregation Algorithm

7.3.1.1 Basic Concepts of Linear Mixture/Un-mixing Models for Grid-based Disaggregation Algorithm

A linear relation has been used successfully to characterize the spectral mixture of a RS pixel-based product [*Holben and Shimabukuro*, 1993; *Ferreira et al.*, 2007]. Based on this approach, the response of individual pixel in any spectral wavelength is assumed as a linear combination of the responses of individual component of the mixture. Thus, each image pixel contains land surface conditions (s_i: soil identification-

soil ID and $a_{i,j}$: sub-grid fraction) with respect to the fraction and spectral response of each component within the ground resolution unit. Hence, it is possible to model each sub-pixel spectrum (s_i and $a_{i,j}$ components **k** in Eq. (7.1)) of the image as linear mixture (Eq. (7.2)) model within a multispectral image. An un-mixing (Eq. (7.3)) can be described by the linear mixture with error term (e_t).

$$\mathbf{k} = \{ s_i = s_1, \ldots, s_M,\ a_{i,j} = a_{1,1}, \ldots,\ a_{M,N-1} \} \tag{7.1}$$

$$\hat{\theta}_t^{sum}(\mathbf{k}) = \sum_{j=1}^{N}\sum_{i=1}^{M} a_{i,j}\hat{\theta}_{i,j,t}^{sub} \quad \forall t \tag{7.2}$$

$$\hat{\theta}_t^{RS}(\mathbf{k}) = \sum_{j=1}^{N}\sum_{i=1}^{M} a_{i,j}\hat{\theta}_{i,j,t}^{sub} + e_t \quad \forall t \tag{7.3}$$

$$\sum_{j=1}^{N}\sum_{i=1}^{M} a_{i,j} = 100 \tag{7.4}$$

$$0 \le a_{i,j} \le 100 \tag{7.5}$$

$$Min s_i \le s_i \le Max s_i \tag{7.6}$$

where **k** is the set of variables whose s_i component is the soil hydraulic properties contained within the pixel and $a_{i,j}$ is the sub-grid cell ($a_{M,N}$ was calculated as 100 - $\sum_{j=1}^{N-1}\sum_{i=1}^{M} a_{i,j}$) of corresponding the soil-vegetation combinations within a RS product, $\hat{\theta}_{i,j,t}^{sub}$ is the simulated soil moisture of each sub-grid, $\hat{\theta}_t^{sum}$ is the sum of simulated soil moisture ($\hat{\theta}_{i,j,t}^{sub}$) of each sub-grid fraction, $\hat{\theta}_t^{RS}$ is the RS soil moisture product, *i* is the number of soil textures, *j* is the number of vegetation covers, M is the soil domain, N is the

vegetation domain, and t is the time index, respectively. The internal constraints are shown in Eqs. (7.4-6).

The objective function used to search for the parameters (**k**) by minimizing the errors between the mixed-scale ($\hat{\theta}_t^{sum}$) and RS ($\hat{\theta}_t^{RS}$) near-surface soil moisture is shown in Eq. (7.7),

$$Z(\mathbf{k}) = \text{Min}\{\frac{1}{T}\sum_{t=1}^{T}|\hat{\theta}_t^{sum} - \hat{\theta}_t^{RS}|\} \quad (7.7)$$

$$\text{Fitness}(\mathbf{k}) = \text{Max}[Z(\mathbf{k})]^{-1} \quad (7.8)$$

where Z is the (minimization) objective function. The suitability of searched parameters (**k**) is evaluated using the fitness function (Eq. 7.8).

7.3.1.2 Ensemble Multiple Operators Genetic Algorithm (EMOGA)

Genetic algorithms (GAs) are powerful search algorithms based on the survival of the fittest mechanism [*Holland,* 1975; *Goldberg,* 1989]. In a GA, genes (called chromosomes) are arranged in a population (an array) to represent the unknown parameters. A GA initializes a population comprised of genes (i.e., binary "0" or "1") for the first generation (g=1) as the starting point in the search space and evaluates the suitability of chromosomes using a fitness function. Then, a new population is produced through a GA operator composed of selection, crossover, and mutation for the next generation (g+1). In the GA operator, elite (best fitness) chromosomes were selected from the mating pool, reproduced via the crossover process, and mutated for diversity of genes. The searching ability of GAs is dependent on initial random seed numbers (e.g., 1000, 2000, and 3000) along with the number of generations. Also, as the number of

parameters (**k**) increases, the searching ability of GA decreases. In this study, we developed a new genetic algorithm for searching solutions more efficiently in the unknown search space. A GA was integrated with multiple operators (MOGA) to address the weakness of GA in searching solutions more efficiently from the unknown space. The unique ability of the MOGA is to reproduce the strongest (elite) chromosome in each population as many times as the number of parameters to be searched (i.e., the number of sub-grid fractions: $M \times N$) as shown in Fig. 7.1b. The reproduced chromosomes ($RC^{r=1,\ldots,M\times N}$) have new genetic information through the GA operators, which can search more spaces. Then the MOGA restarts under the assumption that the chromosomes are converged to one region when the maximum fitness of populations is not changed for fifty generations (note that the number of generations is subjective) before the maximum number of generations are completed. Along with restarting the population, creep (at base 10) and intermittent jump mutation operators are introduced for new genetic materials and a time-saving mechanism allows the MOGA to remember both the elite chromosome of the previous generation (g-1) and the remaining chromosomes [*Ines and Honda,* 2005]. In the MOGA, the elite chromosomes are always reproduced in the next generation. A random re-sampling (ensemble e) algorithm [IBM Programmers' Guide; Efron, 1982] is integrated with the MOGA, to create an Ensemble Multiple Operators Genetic Algorithm (EMOGA).

We used multiple populations (seeds: 1000, 2000, and 3000) and the re-sampling technique (ensemble e) with the EMOGA for uncertainty analysis. At the end of the GA generations, the average of individual maximum fitness values generated by the

combinations of ensemble processes and multi-populations (ensemble e × seeds) was calculated and classified as above- and below-average. Then, we determined the filtered (above-average) and un-filtered (both above- and below-average) solutions and calculated the statistics (average and ±95 percent confidence interval-PCI) of searched solutions in the numerical experiments.

A near-surface soil moisture assimilation scheme based on inverse modeling (IM) [*Ines and Droogers,* 2002; *Ines and Mohanty,* 2008a,b] was integrated with the newly developed EMOGA. It was developed under the assumption that the land surface information (soil ID and sub-grid fractions) can be obtained using only near-surface remotely sensed (0-1/0-5 cm) soil moisture by tuning the hydrological model using a genetic algorithm. The search spaces for soil ID values and sub-grid fractions used in the EMOGA are shown in Table 7.1. We used the physically-based Soil Water Atmosphere Plant (SWAP) model for simulating soil water flow between the soil, water, atmosphere, and plant systems in the soil profile [*Kroes et al.,* 1999*; van Dam et al.,* 1997]. The SWAP model is well known in the hydro-climate model literature. The readers are referred to *Kroes et al.,* [1999] and *van Dam et al.,* [1997].

7.3.1.3 Soil Database

In this study, we used soil database shown in Table 7.2 based on synthesis of selected textures from available national and global soil databases (HYPRES [*Wösten et al*., 1999], Rosetta [*Schaap et al.,* 1999], Staring soil database [*Wösten et al.,* 1994], UNSODA [*Leij et al.,* 1999], and USDA-SCS (Soil Survey) [*Carsel and Parrish,* 1988]. These databases have been established with point scale measurements using laboratory-

based techniques. By using the soil database containing soil textural information at a local-scale (homogeneous) soil unit, we can save the computational time for quantifying soil hydraulic parameters at the sub-grid or fine scales. This is because as opposed to searching for multiple hydraulic parameters, the GA searches for one soil ID value. The soil moisture dynamics are determined based on the soil hydraulic parameters associated with the soil ID value. The soil database has various soil textural ranges for the Mualem-van Genuchten soil hydraulic parameters.

Table 7.1: Search spaces of linear mixture/un-mixing parameters (soil identification-soil ID and sub-grid fractions) used in the EMOGA

	Linear mixture/ un-mixing parameters	Searching space of parameters		No. of bit (L)	2^L
		Minimum values	Maximum values		
Simple land surface Conditions (1v-3s)	s_1	1	64	6	64
	s_2	1	64	6	64
	s_3	1	64	6	64
	$a_{1,1}(s_1v_1)$	0	100	8	256
	$a_{2,1}(s_2v_1)$	0	100	8	256
	$^{*}a_{3,1}(s_3v_1)$	-	-	-	-
Compolex land surface conditions (2v-3s)	s_1	1	64	6	64
	s_2	1	64	6	64
	s_3	1	64	6	64
	$a_{1,1}(s_1v_1)$	0	100	8	256
	$a_{1,2}(s_1v_2)$	0	100	8	256
	$a_{2,1}(s_2v_1)$	0	100	8	256
	$a_{2,2}(s_2v_2)$	0	100	8	256
	$a_{3,1}(s_3v_1)$	0	100	8	256
	$^{*}a_{3,2}(s_3v_2)$	-	-	-	-

Note: 1v-3s (one vegetation cover and three different soil textures); 2v-3s (two vegetation covers and three different soil textures); $a_{i,j}$-sub-grid fractions. $^{*}a_{M,N} = 100 - (\sum_{j=1}^{N-1}\sum_{i=1}^{M} a_{i,j})$

Table 7.2: Soil database used in the modified grid-based disaggregation algorithm

Soil ID	Soil texture	Shape parameters		Scaling parameters				Soil ID	Soil texture	Shape parameters		Scaling parameters			
		α	n	θ_{res}	θ_{sat}	K_{sat}	λ			α	n	θ_{res}	θ_{sat}	K_{sat}	λ
1[a]	Sandy Loam	0.021	1.61	0.067	0.37	41.6	0.5	33[a]	Silt Loam	0.012	1.39	0.061	0.43	30.5	0.5
2[b]	Sandy Loam	0.075	1.89	0.065	0.41	106.1	0.5	34[b]	Silt Loam	0.02	1.41	0.067	0.45	10.8	0.5
3[c]	Sandy Loam	0.027	1.45	0.039	0.39	38.3	-0.861	35[c]	Silt Loam	0.005	1.66	0.065	0.44	18.2	0.365
4[d]	Tsand B1	0.025	1.51	0.01	0.43	17.5	-0.14	36[a]	SandyClayLoam	0.033	1.49	0.086	0.4	9.7	0.5
5[d]	Tsand B2	0.023	1.55	0.02	0.43	9.7	-0.983	37[b]	SandyClayLoam	0.059	1.48	0.1	0.39	31.4	0.5
6[d]	Tsand B3	0.015	1.41	0.01	0.45	17.8	-0.213	38[c]	SandyClayLoam	0.021	1.33	0.063	0.38	13.2	-1.28
7[d]	Tsand B4	0.016	1.56	0.01	0.42	54.8	0.177	39[e]	Tfine	0.037	1.1	0.01	0.52	24.8	-1.9772
8[e]	Scoarse	0.043	1.52	0.025	0.37	70	1.25	40[a]	Clay Loam	0.03	1.37	0.129	0.47	1.8	0.5
9[d]	SsandO3	0.021	1.56	0.01	0.34	18.3	-0.522	41[b]	Clay Loam	0.019	1.31	0.095	0.41	6.2	0.5
10[d]	SsandO4	0.022	1.54	0	0.36	53.1	-0.52	42[c]	Clay Loam	0.016	1.42	0.079	0.44	8.2	-0.763
11[d]	SsandO5	0.06	2.06	0.01	0.32	43.6	0.343	43[a]	Silty Clay Loam	0.027	1.41	0.098	0.55	7.4	0.5
12[d]	SsandO6	0.029	1.15	0	0.41	5.5	-6.864	44[b]	Silty Clay Loam	0.01	1.23	0.089	0.43	1.7	0.5
13[e]	Tcoarse	0.038	1.38	0.025	0.4	60	1.25	45[c]	Silty Clay Loam	0.008	1.52	0.09	0.48	11.1	-0.156
14[a]	Loam	0.025	1.31	0.083	0.46	38.3	0.5	46[b]	Sandy Clay	0.027	1.23	0.1	0.38	2.9	0.5
15[b]	Loam	0.036	1.56	0.078	0.43	25	0.5	47[c]	Sandy Clay	0.033	1.21	0.117	0.39	11.4	-3.665
16[c]	Loam	0.011	1.47	0.061	0.4	12.1	-0.371	48[a]	Silty Clay	0.023	1.39	0.163	0.47	8.4	0.5
17[d]	TloamB7	0.019	1.25	0	0.4	14.1	-0.802	49[b]	Silty Clay	0.005	1.09	0.07	0.36	0.5	0.5
18[d]	TloamB8	0.01	1.28	0	0.43	2.3	-2.733	50[c]	Silty Clay	0.016	1.32	0.111	0.48	9.6	-1.287
19[d]	TloamB9	0.007	1.33	0	0.43	1.5	-2.161	51[a]	Clay	0.021	1.2	0.102	0.51	26	0.5
20[d]	SloamO8	0.014	1.34	0	0.47	9.1	-0.803	52[b]	Clay	0.008	1.09	0.068	0.38	4.8	0.5
21[d]	SloamO9	0.009	1.4	0	0.46	2.2	-1.382	53[c]	Clay	0.015	1.25	0.098	0.46	14.8	-1.561
22[d]	SloamO10	0.011	1.28	0	0.49	2.2	-2.123	54[d]	TclayB10	0.012	1.22	0.01	0.42	1.2	-4.795
23[e]	Tmedium	0.031	1.18	0.01	0.44	12.1	-2.3421	55[d]	TclayB11	0.024	1.11	0	0.6	5.3	-5.395
24[a]	Silt	0.006	1.53	0.123	0.48	55.7	0.5	56[d]	TclayB12	0.053	1.08	0	0.55	15.5	-8.823
25[b]	Silt	0.016	1.37	0.034	0.46	60	0.5	57[d]	SclayO11	0.019	1.15	0	0.42	13.8	-1.384
26[c]	Silt	0.007	1.68	0.05	0.49	43.8	0.624	58[d]	SclayO12	0.01	1.16	0	0.56	1.1	-4.171
27[d]	TsiltB14	0.005	1.31	0.01	0.42	0.8	0	59[d]	SclayO13	0.017	1.11	0	0.57	3.3	-4.645
28[d]	SsiltO14	0.003	1.69	0	0.38	0.4	0.057	60[e]	Sfine	0.02	1.09	0.01	0.48	8.5	-3.7124
29[d]	SsiltO15	0.007	1.3	0.01	0.41	3.7	0.912	61[e]	Tveryfine	0.027	1.1	0.01	0.61	15	2.5
30[e]	Smedium	0.025	1.17	0.01	0.39	10.8	-0.7437	62[e]	Sveryfine	0.017	1.07	0.01	0.54	8.2	0.0001
31[e]	Tmediumfine	0.008	1.25	0.01	0.43	2.3	-0.5884	63[d]	TpeatB16	0.013	1.32	0	0.73	13.4	0.534
32[e]	Smediumfine	0.008	1.22	0.01	0.41	4	0.5	64[d]	TpeatB17	0.018	1.14	0	0.72	4.5	-0.35

[a]UNSODA; [b]SoilSurvey; [c]Rosetta; [d]Staring; [e]Hypres (T and S prefixes mean Top- and Sub-surface soils) [*Ines et al.*, 2012]

7.3.2 Numerical Experiments

We conducted synthetic and field validation experiments to evaluate the proposed approach for the current decade (2010-2020, with growing season from March 01 to July 31). The numerical experiments were composed of various steps under rain-fed conditions including; Step 1: testing of the modified grid-based disaggregation algorithm for simple/complex synthetically generated land surface conditions based on the coupled IM-EMOGA, Step 2: incorporation of future climatic scenarios from global circulation models (GCMs) into the algorithm, Step 3: prediction of daily root zone soil moisture at sub-grid scales, Step 4: drought severity assessment for the study domain, and Step 5: field validation experiment and drought severity assessment in the future.

Step 1 is to test the approach under various synthetic land surface conditions (representing the areal heterogeneity) during the simulation period (2002) at the Lubbock site in Texas with (1) simple land surface condition (1 vegetation and 3 soil textures, 1v-3s) and (2) complex land surface condition (2 vegetations and 3 soil textures, 2v-3s), respectively. Figure 7.2a shows the complex land surface condition designed for this experiment. Additionally, six different scenarios comprised of various sub-grid fractions of soil textures and vegetation covers were tested in the simple land surface case (Table 7.3).

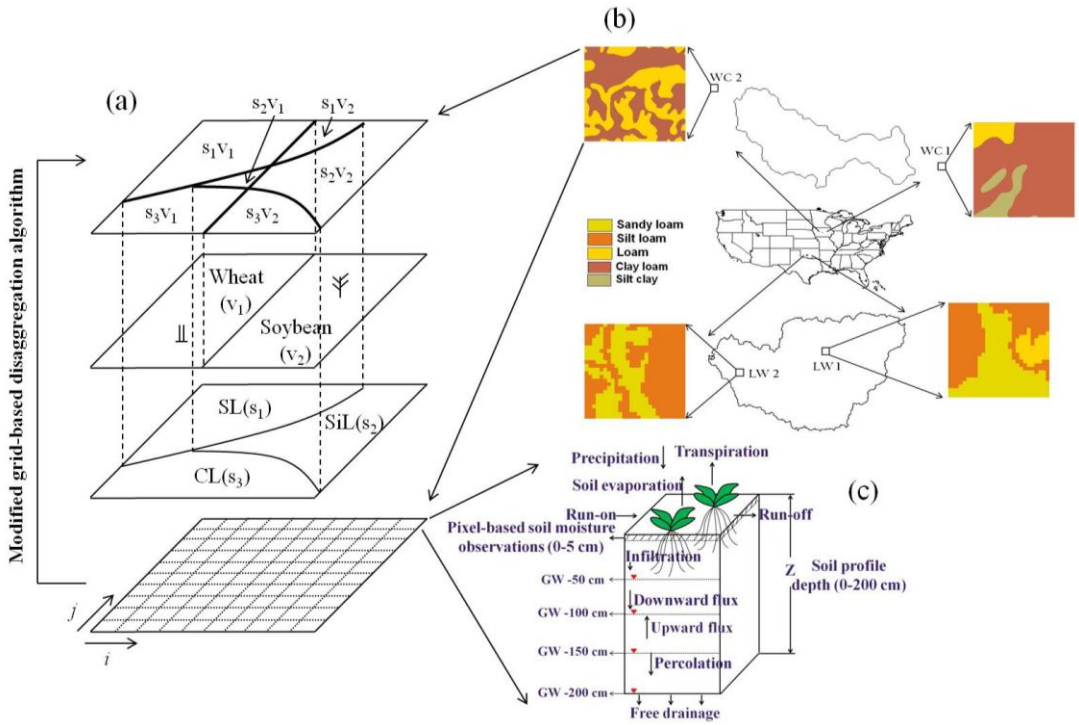

Figure 7.2: (a) Grid domain (10 × 10) for the numerical experiments (e.g., complex land surface condition for Step 1); (b) the Walnut Creek (WC 1 and 2: 800 m × 800 m) and Little Washita (LW 1 and 2: 800 m × 800 m) sites in Iowa and Oklahoma for the field validation experiments; (c) homogeneous soil column under free-drainage and various ground water table depths (-200, -150, -100, and -50 cm)

Note: v_1 - vegetation 1 (wheat), v_2 – vegetation 2 (soybean), s_1 – soil 1 (sandy loam-SL), s_2 – soil 2 (silt loam-SiL), s_3 – soil 3 (clay loam-CL)

Table 7.3: Filtered/un-filtered solutions derived by the modified grid-based disaggregation algorithm using the EMOGA with the simple land surface condition for Step 1

Categories	Filtered/ un-filtered	Soil types	Target values		Solutions		
			Soil ID	Sub-grid fractions ($a_{i,j}$)	Soil ID	Average of sub-grid fractions($a_{i,j}$)	SD of sub-grid fractions
Scenario 1 (v_1: wheat)	Filtered	s_1	1	20	1	20	0
		s_2	33	30	33	30	0
		s_3	40	50	40	50	0
	Un-filtered	s_1	1	20	1	19.5	1.5
		s_2	33	30	16, 33	30.6	1.2
		s_3	40	50	40	49.9	0.6
Scenario 2 (v_1: wheat)	Filtered	s_1	1	30	1	30	0
		s_2	33	30	33	30	0
		s_3	40	40	40	40	0
	Un-filtered	s_1	1	30	1	29.5	1.8
		s_2	33	30	16, 33	30.2	0.9
		s_3	40	40	40	40.2	0.8
Scenario 3 (v_1: wheat)	Filtered	s_1	1	80	1	80	0
		s_2	33	10	33	10	0
		s_3	40	10	40	10	0
	Un-filtered	s_1	1	80	1	80	0
		s_2	33	10	33	10	0
		s_3	40	10	40	10	0
Scenario 4 (v_1: soybean)	Filtered	s_1	1	20	1	20	0
		s_2	33	30	33	30	0
		s_3	40	50	40	50	0
	Un-filtered	s_1	1	20	1	20	0
		s_2	33	30	33	30.2	0.4
		s_3	40	50	40	49.8	0.4
Scenario 5 (v_1: soybean)	Filtered	s_1	1	30	1	30	0
		s_2	33	30	33	30	0
		s_3	40	40	40	40	0
	Un-filtered	s_1	1	30	1	29.4	2.6
		s_2	33	30	33, 38	30.8	3.8
		s_3	40	40	40	39.8	0.9
Scenario 6 (v_1: soybean)	Filtered	s_1	1	80	1	80	0
		s_2	33	10	33	10	0
		s_3	40	10	40	10	0
	Un-filtered	s_1	1	80	1	80	0
		s_2	33	10	33	10	0
		s_3	40	10	40	10	0

Population: 30; Total generations: 5000 (restart 8); seed number: 1000, 2000, 3000; p_{creep}: 0.05; p_{mutate}: 0.05

Note: the target values –soil types (s_1: 1-sandy loam, s_2: 33-silt loam, s_3: 40-clay loam)**;** filtered – above average solutions; unfiltered-all solutions

In Step 2, we selected weather input datasets for the hydrological model from the Coupled Global Climate Model 2 (CGCM2) provided by Canadian Centre for Climate modeling and analysis (CCCma) for predicting root zone soil moisture dynamics. The A_2 (regionally oriented economic growth and continuously increasing global population) story line [see details, *Nakicenovic et al.*, 2000] was selected for the CGCM2 scenarios. The CGCM2 has a very coarse resolution (3.75° by 3.75°, corresponding to more than 400 km by 400 km), which means that it is not representative of climate changes at the local-scale. Therefore, we corrected the monthly CGCM2 products (e.g., precipitation, humidity, wind speed, solar radiation, maximum and minimum temperature) to represent local scales using the historical weather data (1998-2009) based on a simple multiplicative shift method (Eq. 7.9) suggested by *Ines and Honda* [2005],

$$x_t' = x_t \frac{\overline{X}_{obs}}{\overline{X}_{GCM}} \quad \forall t \tag{7.9}$$

where x_t and x_t' are the daily raw and corrected GCM scenarios and $\overline{X}_{GCM}$ and $\overline{X}_{obs}$ are the monthly mean weather predictions of the GCM and historical observed weather data, respectively.

As the bias of raw CGCM2 scenarios were corrected in Eq. (7.9), the bias correction coefficients can be obtained for each month during the previous decade (1998-2009). By using the derived monthly bias correction coefficients of the past decade, future climate changes were compensated for the current decade (2010-2020). In order to capture the variation range of corrected CGCM2 scenarios, we estimated the monthly average and ±95 percent confidence interval (PCI) of CGCM2 scenarios.

With the derived land surface information (Step 1) and bias corrected CGCM2 products (Step 2), we predicted the long-term root zone soil moisture dynamics in the study domain for Step 3 during the current decade (2010-2020). This was computed for both the disaggregated pixels composed of individual soil vegetation combinations (s_1v_1, s_2v_1, and s_3v_1) and the non-disaggregated pixels consisting of mixed soil vegetation combinations ($s_1v_1 + s_2v_1 + s_3v_1$). Three different soil thicknesses (0-1, 0-30, and 0-60 cm) and four different ground water table depths (-200, -150, -100, and -50 cm) were used for the root zone soil moisture predictions. When the bottom boundary condition was governed by the presence of ground water table, it was assumed that the initial pressure condition in the soil profile was in equilibrium with the ground water table. Also, we tested the impacts of various vegetation covers (i.e., wheat, soybean, grass, and maize) for predicting the root zone soil moisture dynamics. Finally we evaluated a local-scale drought severity using the predicted soil moisture dynamics based on the soil moisture deficit index (SMDI) [see details, *Narasimhan and Srinivasan*, 2005] in Step 4. In order to verify our proposed approach for evaluating grid-scale drought severity, the Little Washita sites in Oklahoma (LW1 and 2, June 22 – July 16, 1997) and Walnut Creek sites in Iowa (WC1 and 2, June 25 – July 12, 2002) were selected for the field validation experiments under different hydro-climates in Step 5, as shown in Figure 7.2b. Based on fine-scale soil moisture predictions estimated from the LW and WC sites, we evaluated drought severity for the current decade (2010-2020).

7.3.3 Model Condition and Data Description

In this study, the soil column (total depth 200 cm) had 33 discretized computational layers. The depth of 0-10 cm from the soil surface was finely discretized at intervals of 1 cm. The depths of 10-60 cm and 60-200 cm were discretized at intervals of 5 cm and 10 cm (the 33rd layer has a soil depth of 20 cm), respectively. In the numerical study, the bottom boundary condition was set as free-drainage representing semi-/arid-regions. We assumed the initial condition of h(z,t=0) = -200 cm below soil surface. In the field validation experiments, the bottom boundary conditions were chosen based on the observed range of water table in the given field, i.e., 100-200 cm below soil surface. Ground water interacts with the vadose zone through upward capillary flow of water. In other words, a higher ground water table can lead to more uncertainties in soil hydraulic parameter estimation for the soil profile. Therefore, the bottom boundary condition was prescribed with initial soil water pressure head distribution in hydrostatic equilibrium with an initial water table depth of -200cm from the soil surface.

For the synthetic experiment (Step 1), the (target) soil moisture dynamics were generated by the SWAP model using the soil ID values from the soil database (Table 7.2) with combinations of various sub-grid fractions (Table 7.3) in a forward mode. Synthetic experiments included less complexity compared to RS pixels or *in-situ* soil moisture measurements, which have variations due to heterogeneity of land surface (distributed soil textures and vegetation covers), land management practices, weather conditions, etc. Thus, the synthetic experiments were considered suitable for assessing the application of the approach in a strict sense. In the field validation experiments (Step

5), the RS soil moisture footprints (ESTAR [*Jackson et al.*, 1999] for 17 days at the LW1 and LW2 sites and PSR [*Bindlish et al*, 2006] for 10 days at the WC1 and WC2 sites, 800 m × 800 m) were downscaled for extracting the soil ID values and sub-grid fractions.

Then we validated the results derived by our approach with the observed soil type and vegetation cover. The vegetation cover and soil information were obtained from Geographic Information Systems Resources-GISR (http://www.webgis.com/index.html) and Natural Resources Conservation Service-NRCS (http://www.nrcs.usda.gov/), respectively. Remotely sensed soil moisture products usually have a penetration depth of 0-1 cm (C-band) and 0-5 cm (L-band) from the soil surface. Thus, land surface information (soil ID values and sub-grid fractions) were determined using only the near-surface (0-1 cm for the synthetic conditions and 0-5 cm for the field validations) soil moisture with the modified grid-based disaggregation algorithm.

Field observed hydraulic parameters were used to verify our approach. Soil parameters were derived from the soil core samples at the soil depth of 3-9 cm collected near the LW1 and 2 sites. However, no field-observed data were available for the WC1 and 2 sites. Thus, we only compared the observed and estimated (from soil ID values) water retention curves ($\theta(h)$) for validation at the LW1 and 2 sites. The LW (1 and 2) and WC (1 and 2) regions have a grass cover. Daily climatic data (e.g., precipitation, wind speed, maximum and minimum temperature, and solar radiation) for the model inputs were collected from the USDA-ARS Micronet weather station

(http://ars.mesonet.org/) in Oklahoma and the Soil Climate Analysis Network (SCAN: http://www.wcc.nrcs.usda.gov/scan/) in Iowa.

7.4 Results and Discussion

7.4.1 Step 1: Various Land Surface Conditions

Figures 7.3a,b show the performance (maximum fitness and average errors) of newly developed coupled IM-EMOGA linkage with the combinations of multi-populations (1000, 2000, and 3000) and ensembles (e=10) for the simple land surface condition. The performance of maximum fitness and average errors for the six scenarios were similar under the simple land surface conditions. Thus, we present the results of only one of the scenarios (scenario 6, see section 2.2) in this discussion (Figs, 7.3a,b). The improvement of the solutions as the generations progress indicates that the parameters derived by the EMOGA converge to the global/local solutions within the unknown search space. The maximum fitness of elite chromosomes with the combinations of multi-populations and ensembles showed different trends, which meant that their genetic traits were attributable to their initial points in the search space. As shown in Figure 7.3b, the average errors of soil moisture corresponded to the maximum fitness of solution combinations. Overall, the EMOGA usually found well the optimized solutions during the initial generations.

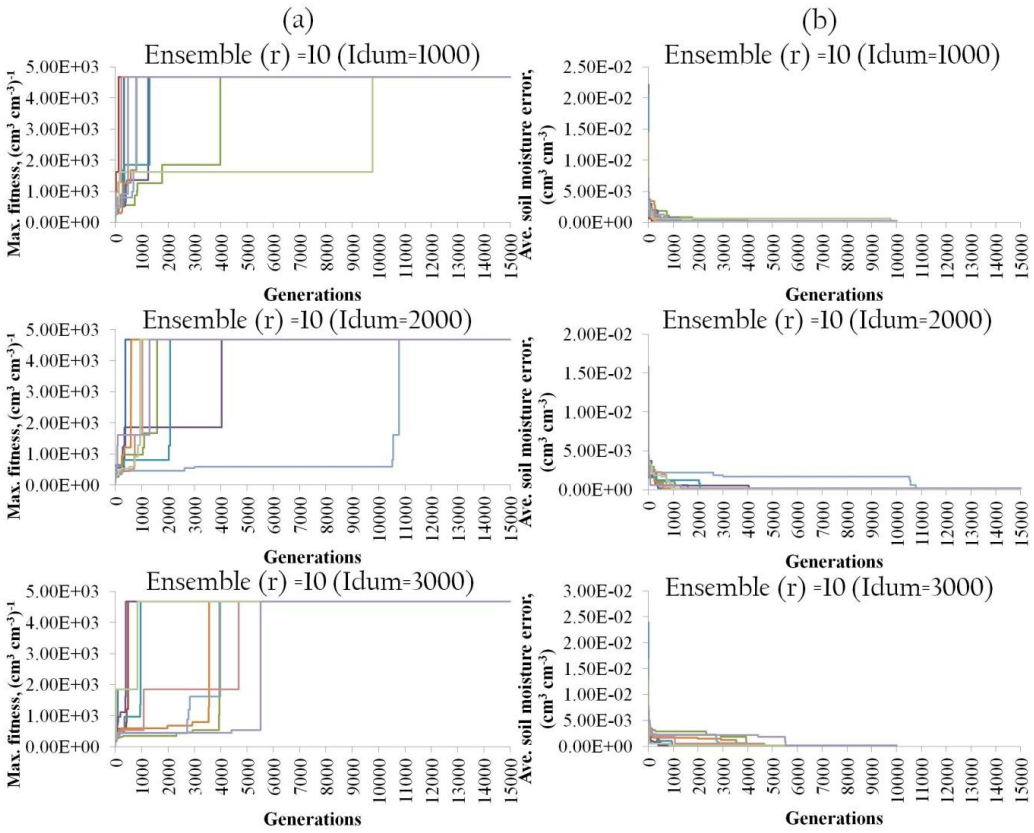

Figure 7.3: Example of the EMOGA solutions for Step 1 (scenario 6 - soybean cover; a_1: 80, a_2: 10, a_3: 10) with combinations of different random number generator seeds and ensemble e; (a) maximum fitness and (b) average soil moisture error

Seeds: 1000, 2000, and 3000
Ensemble e: 10

Table 7.3 shows the filtered/un-filtered solutions (described in section 2.1.2 above) for the combinations of various sub-grid fractions derived by the EOMGA. Overall, the filtered/un-filtered solutions for the simple land surface condition corresponded with the (synthetic) target values. The filtered solutions of soil ID (s_1=1, s_2=33, s_3=40) values and sub-grid fractions ($a_{1,1}$=20, $a_{2,1}$=30, $a_{3,1}$=50 for the scenario 1,

and $a_{1,1}$=30, $a_{2,1}$=30, $a_{3,1}$=40 for scenario 2) with the wheat crop matched well with the target values. But the un-filtered solutions have small variations compared to the filtered ones. The filtered/un-filtered solutions for scenario 3 representing the relatively irregular land surface ($a_{1,1}$=80, $a_{2,1}$=10, $a_{3,1}$=10) were perfectly identifiable with the observations compared to the results of scenario 1 and 2. The results of scenarios 4 to 6 with the soybean cover and three soil textures showed similar trends as shown in those of scenarios 1 to 3. The filtered solutions of scenarios 4 and 5 matched exactly with the target values, although the un-filtered ones had small errors in modeling. The filtered/un-filtered solutions for scenario 6 also appeared to be identifiable with the target values. In most cases, the solutions were effectively found under the synthetic experiment conditions, but the un-filtered soil ID (s_2: 16, **33**, 38, **bold** means the exact solution) values with silt loam soil had more variations in modeling for scenarios 1 to 6. It showed that a silt loam soil caused more uncertainties than other soils in the model performance. Table 7.4 shows the derived (filtered/un-filtered) solutions for the complex land surface conditions (2 vegetation covers and 3 different soil textures), which had six sub-grid fractions ($a_{1,1}$: 17, $a_{1,2}$: 17, $a_{2,1}$: 17, $a_{2,2}$: 17, $a_{3,1}$: 17, $a_{3,2}$: 15). In the filtered solutions, the soil ID values (s_1: 1, s_2: 33, s_3: 40) corresponded well to the target values. Although the searched sub-grid fractions ($a_{1,1}$: 16.7, $a_{1,2}$: 16.3, $a_{2,1}$: 17.0, $a_{2,2}$: 16.7, $a_{3,1}$: 16.5, $a_{3,2}$: 16.8) have small errors in estimating parameters, the solutions were close to the target values. The un-filtered solutions of soil ID (s_1: **1**, 15, 16, 38; s_2: **33**, 36, 38; s_3: **40**, 46) and sub-grid fractions ($a_{1,1}$: 15.3, $a_{1,2}$: 17.3, $a_{2,1}$: 18.5, $a_{2,2}$: 16.7, $a_{3,1}$: 16.4, $a_{3,2}$: 15.8) have more uncertainties compared to the filtered ones. The standard deviations of

sub-grid fractions were also higher than those of the filtered solutions. Thus, these findings showed that the modified grid-based disaggregation algorithm using the new simulation-optimization (SWAP-EMOGA) with both filtered (above-average) and un-filtered (both above- and below-average) solutions performed well in extracting valuable information within a RS footprint using only the near-surface (0-1/0-5 cm) soil moisture, although small variations exist.

7.4.2 Step 2: Future Climate Changes Using Global Climate Models (GCMs)

Figures 7.4a,b show the comparison of bias uncorrected and corrected CGCM2 scenarios at the Lubbock site, Texas, during the historical period (1998 to 2009). In this study, the precipitation data at the Lubbock site was only shown for the sake of brevity. The bias uncorrected CGCM2 scenario was considerably overestimated compared to the historical weather data (Fig. 7.4a), but the CGCM2 data corrected by the multiplicative approach showed the similar frequency and intensity with the observation (Fig. 7.4b) with some uncertainties. Using the coefficients (not shown in this text) obtained through the bias correction for the historical period (1998-2009), we compensated the CGCM2 scenarios for the current decade (2010-2020) as shown in Fig. 7.4c. The monthly average CGCM2-based precipitation forecasts show seasonal periodicity during the simulation period.

Table 7.4: Filtered/un-filtered solutions derived by the modified grid-based disaggregation algorithm using the EMOGA with complex land surface conditions from Step 1

Categories	Soil types	Target values			Solutions				
		Soil ID	Sub-grid fraction ($a_{i,j}$)		Soil ID	Average of sub-grid fractions ($a_{i,j}$)		SD of sub-grid fractions	
			v_1	v_2		v_1	v_2	v_1	v_2
Filtered solutions	s_1	1	1/7	17	1	16.7	16.3	1.1	1.1
	s_2	33	1/7	17	33	17.0	16.7	0.0	0.5
	s_3	40	1/7	15	40	16.5	16.8	1.0	0.6
Un-filtered solutions	s_1	1	1/7	17	1, 15, 16, 38	15.3	17.3	6.6	4.7
	s_2	33	1/7	17	33, 36, 38	18.5	16.7	5.0	3.1
	s_3	40	1/7	15	40, 46	16.4	15.8	3.3	2.2

Population: 30; Total generations: 5000 (restart 8); seed number: 1000, 2000, 3000; p_{creep}: 0.05; p_{mutate}: 0.05

Note: the target values – vegetations (v_1 - wheat cover, v_2 - soybean cover); soil types (s_1: 1-sandy loam, s_2: 33-silt loam, s_3: 40-clay loam)

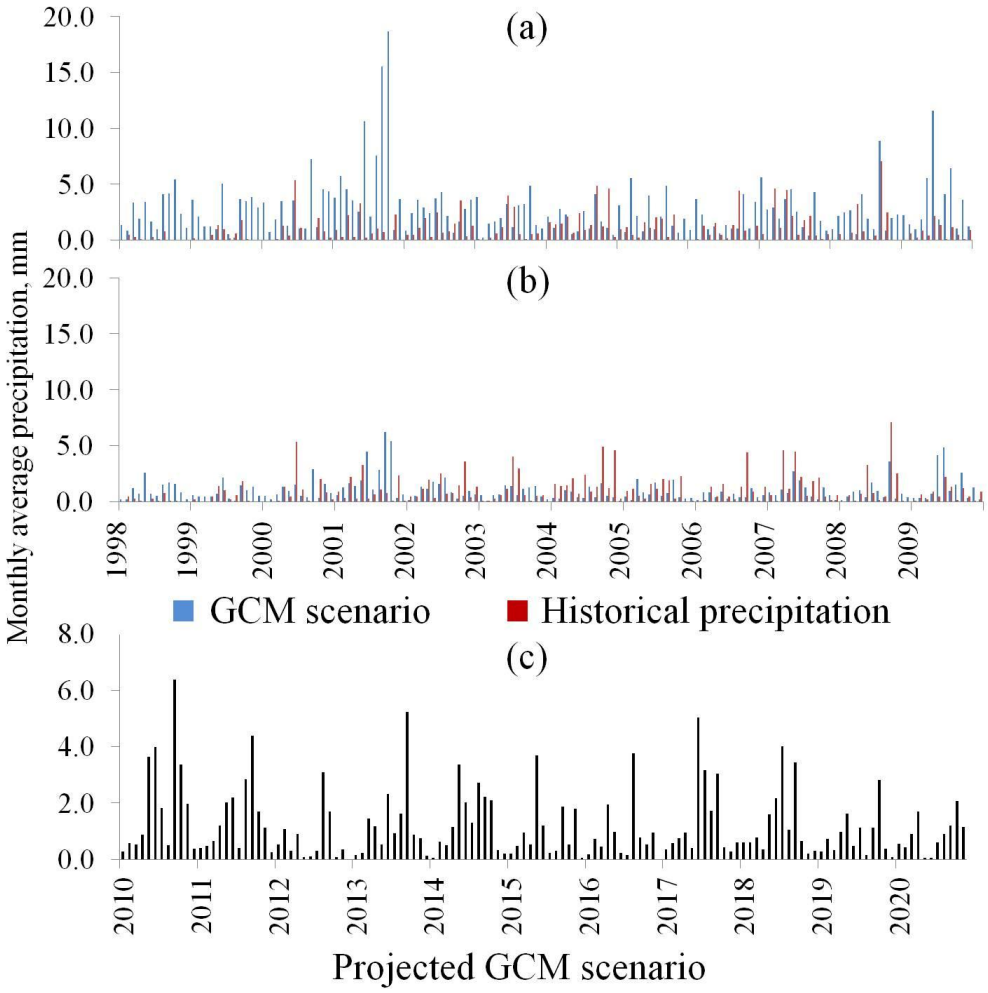

Figure 7.4: (a) Bias uncorrected monthly average GCM and historical precipitation at the Lubbock site, Texas during 1998 – 2009; (b) bias corrected monthly average GCM and historical precipitation based on the multiplicative shift method at the Lubbock site, Texas during 1998 – 2009; (c) bias corrected GCM precipitation using the bias correction coefficients (obtained from 1998 to 2009) at the Lubbock site, Texas during 2010 - 2020

7.4.3 Step 3: Prediction of Daily Root Zone Soil Moisture

In Step 3, we predicted the daily root zone soil moisture dynamics at the Lubbock site using only the filtered solutions for various soil-vegetation conditions (Step 1) with the corrected CGCM2 scenarios (Step 2) under the synthetic conditions during 2010 - 2020. The predicted soil moisture dynamics at the soil depth 0-1 cm were more variable than those of the soil depths of 0-30 and 0-60 cm as shown in Figures 7.5a-c. It indicated that soil water storage at the soil depth of 0-1 cm is minimal and moisture actively evaporates near the soil surface. The predicted soil moisture estimates at the soil depths of 0-30 and 0-60 cm were relatively stable compared to that at the soil depth of 0-1 cm, but they have no significant differences between each other. This result showed that the soil water contents near the soil surface (only 0-1/0-5 cm) are not appropriate for assessing a drought severity, because the bulk of the root zone exists in the deeper (below top 5cm) soil profile. Also, the near-surface soil moisture is significantly affected by the weather conditions directly. Thus, we used the root zone soil moisture from a soil thickness of 0-30 cm for evaluating drought conditions.

Figures 7.5d-i show the statistics (average and ±95 PCI) of disaggregated-(s_1v_1, s_2v_1, s_3v_1) and mixed-($s_1v_1+s_2v_1+s_3v_1$) scale root zone soil moisture predictions for scenarios 1 to 3 (from Table 7.3) with wheat cover (v_1). At the disaggregated-scale, soil texture greatly influences the predictions of soil moisture dynamics as the disaggregated soil moisture for sandy loam (s_1) soil were considerably lower than those of silt loam (s_2) and clay loam (s_3) soils. The prediction of mixed-scale soil moisture for scenario 3 (Fig. 7.5i), which had a (dominant) sandy loam soil (80%) within a pixel, was slightly lower

than those of scenarios 1 and 2. Usually, the mixed-scale soil moistures of scenarios 1 to 3 had less variations compared to the disaggregated (s_1v_1, s_2v_1, s_3v_1) results.

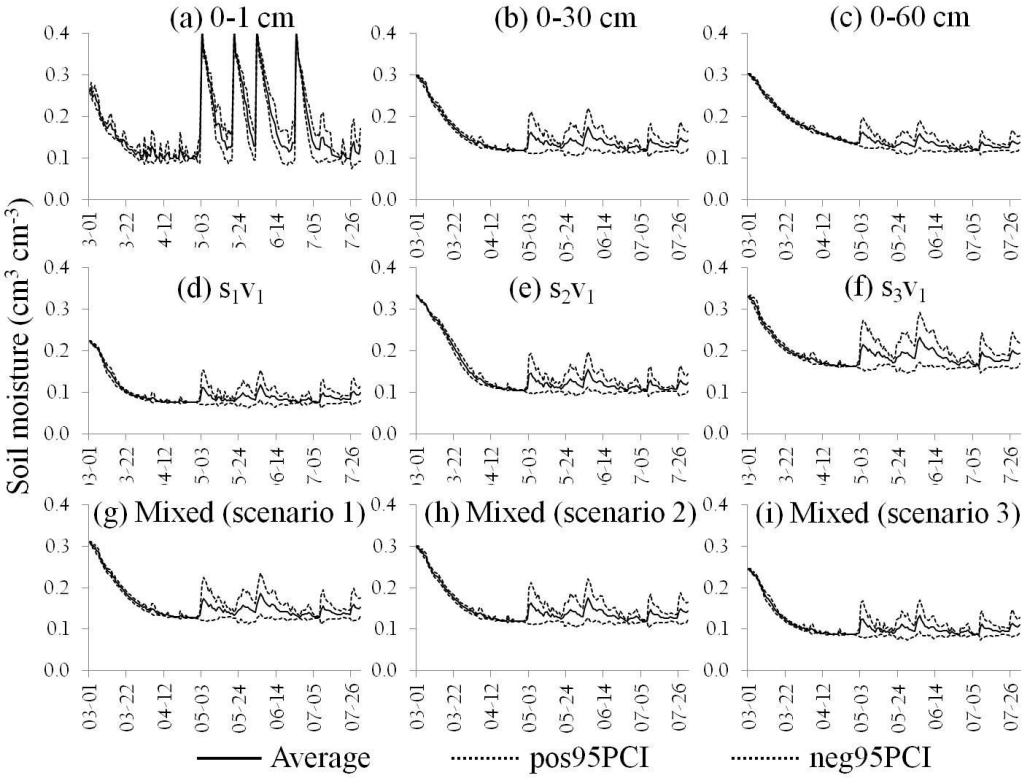

Figure 7.5: Predicted daily long-term root zone soil moisture; (a-c) different soil thicknesses, (d-f) disaggregated-scale (s_1v_1, s_2v_1, s_3v_1) with soil thickness of 0-30 cm, (g-i) mixed-scale (s_1v_1+s_2v_1+s_3v_1) with soil thickness of 0-30 cm using the CGCM2 scenario during 2010-2020 (s_1: sandy loam, s_2: silt loam, s_3: clay loam, v_1: wheat cover, scenario 1 - $a_{1,1}$: 20, $a_{2,1}$: 30, $a_{3,1}$: 50, scenario 2 - $a_{1,1}$: 30, $a_{2,1}$: 30, $a_{3,1}$: 40, scenario 3 - $a_{1,1}$: 80, $a_{2,1}$: 10, $a_{3,1}$: 10)

In Fig. 7.6, we present the effects of various vegetation covers (v_1: wheat, v_2: soybean, v_3: grass, and v_4: maize) on forecasting the root zone soil moisture (only scenario 2 with regular sub-grid fractions was shown). Vegetation covers have significant impact on the prediction of root zone soil moisture during the simulation period (2010-2020). The predicted root zone soil moisture estimates with soybean, grass, and maize covers were considerably higher than that of the wheat cover, especially for the grass cover. This may indicate that the grass cover, which has less bare soil, can endure a drought better than the other vegetation covers. As shown in the findings of simple land surface condition (Fig. 7.6), the predicted soil moisture of clay loam under the complex land surface condition (figure for the complex land surface condition was excluded, because the trends of soil moisture dynamics were similar to those of the simple land surface condition) was higher than those of sandy loam and silt loam soils. Also, for the two vegetation covers studied under the complex land surface conditions, soybean cover was shown to hold more water content than a wheat cover in the soil depth (0-30 cm).

We considered the impacts of various shallow water table depths (-200, -150, -100, and -50 cm) on the prediction of mixed-scale root zone soil moisture with the wheat cover (scenario 2) as shown in Figures 7.7a-d. The predicted root zone soil moisture with the shallow water table depth of -50 cm was considerably higher compared to that of the free-drainage condition (Fig. 7.5h), but the deeper water table depths of -200, -150, and -100 cm have less direct impacts on the root zone soil moisture quantities.

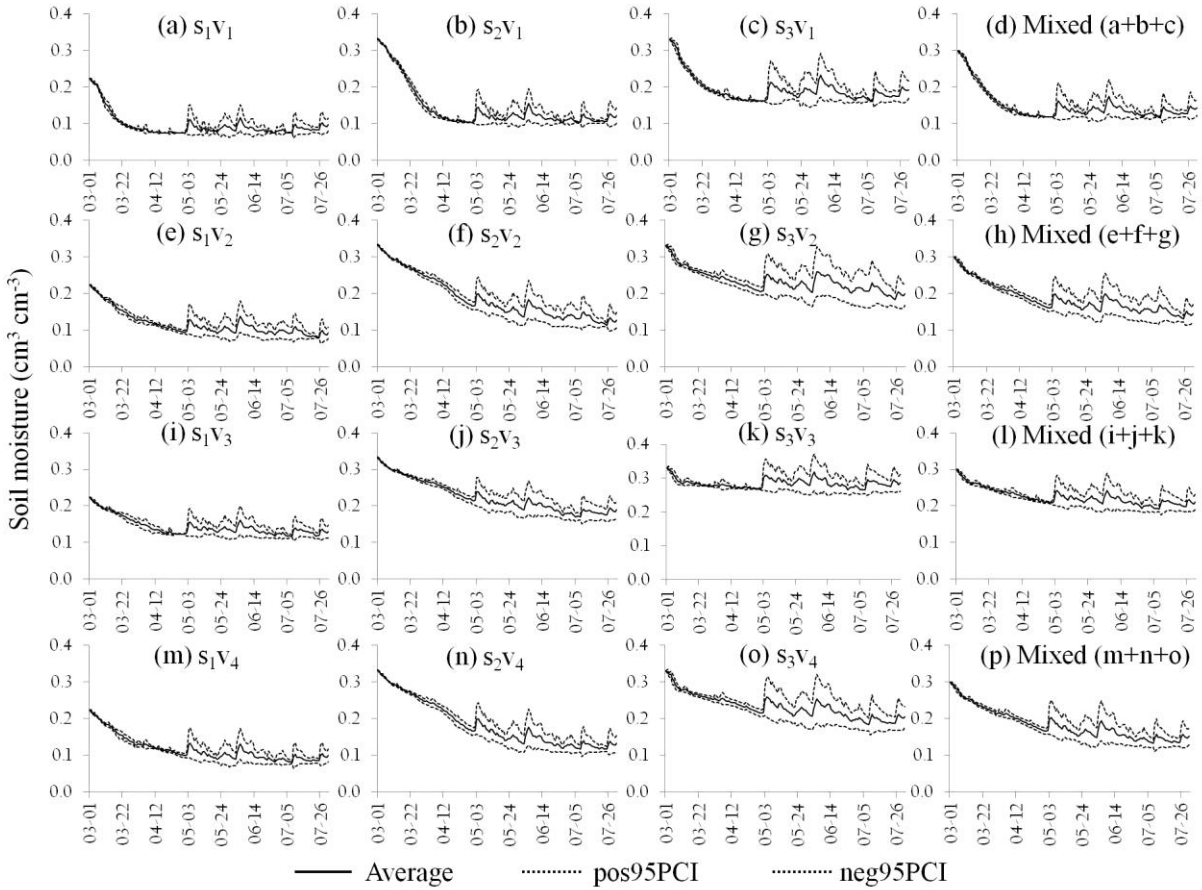

Figure 7.6: Predicted daily root zone soil moisture of disaggregated- (s_1v_1, s_2v_1, s_3v_1) and mixed- (s_1v_1 + s_2v_1 + s_3v_1) scale soil-vegetation conditions in Step 1 during 2010-2020 (s_1: sandy loam, s_2: silt loam, s_3: clay loam, v_1: wheat, v_2: soybean, v_3: grass, v_4: maize covers)

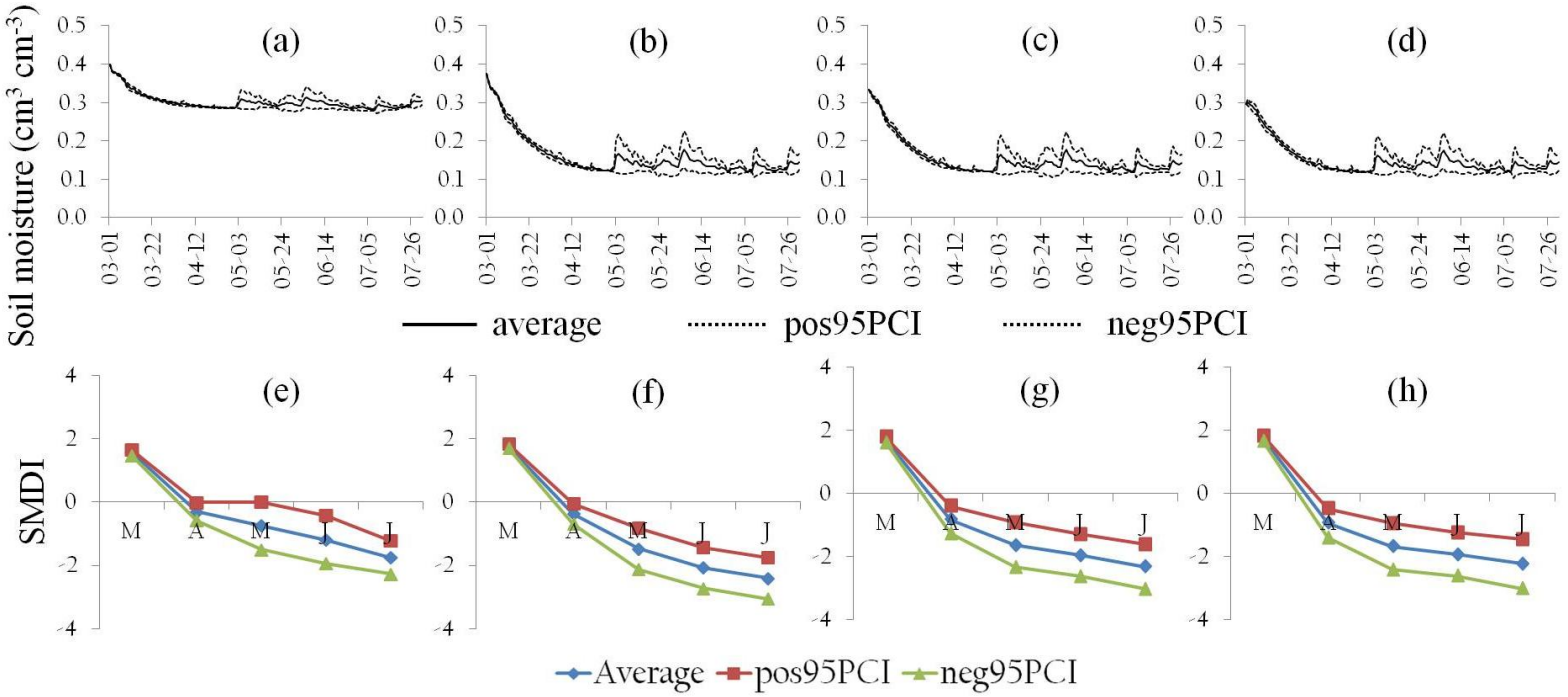

Figure 7.7: Predicted daily root zone soil moisture and monthly soil moisture deficit index (SMDI) values of mixed-scale ($s_1v_1+s_2v_1+s_3v_1$) soil-vegetation conditions with the presence of water table depths of -50 (a,e), -100 (b,f), -150 (c,g), and -200 (d,h) cm for scenario 2 during 2010-2020 (s_1: sandy loam, s_2: silt loam, s_3: clay loam, v_1: wheat cover)

7.4.4 Step 4: Drought Severity Assessment

Based on the findings of Step 3, we estimated the monthly statistics (average and ±95PCI) of SMDI values for evaluating the drought severity. The SMDI values of silt loam and clay loam soils with the wheat cover were slightly higher than that of sandy loam soil (Table 7.5) at the disaggregated-scale. But the SMDI values at the mixed-scale (scenarios 1, 2, and 3 were only shown) have no significant differences.

Table 7.6 presents the monthly average SMDI values for the disaggregated- and mixed-scale root zone soil moisture with the combinations of different soil textures (s_1, s_2, s_3) and vegetation covers (v_1, v_2, v_3, v_4) under the simple land surface condition (one vegetation and three soil textures). The mixed-scale SMDI values for soybean (v_2), grass (v_3), and maize cover (v_4) were similar indicating that these crops were affected by the drought condition from May to July, but the SMDI of wheat cover decreased steeply from April. Overall, it is inferred that the Lubbock site may be affected by a drought condition from April or May to July during the current decade and different vegetation covers considerably influence the SMDI values. The field sites with soybean, grass, and maize covers had more positive SMDI values, as compared to the sites with a wheat cover, indicating that the former vegetation covers would handle drought better. In the complex land surface condition (2 vegetations and 3 soil textures, Table 7.7), the SMDI values had a trend similar to the simple land surface condition. The SMDI values with the soybean cover and silt loam/clay loam soils were higher than those, which had sandy loam soil and wheat cover. But they were still influenced by the drought from May to July under the complex land surface conditions. Figures 7.7e-h show the monthly SMDI

values for various water table depths of -200, -150, -100, and -50 cm with the wheat cover (scenario 2) at the Lubbock site. As shown in the predicted root zone soil moisture (Figures 7.8a-d), the shallow water table depth of -50 cm only increased the SMDI. This result indicated that the water table depths of below -100 cm may have less direct influence on reducing drought severity at the field site.

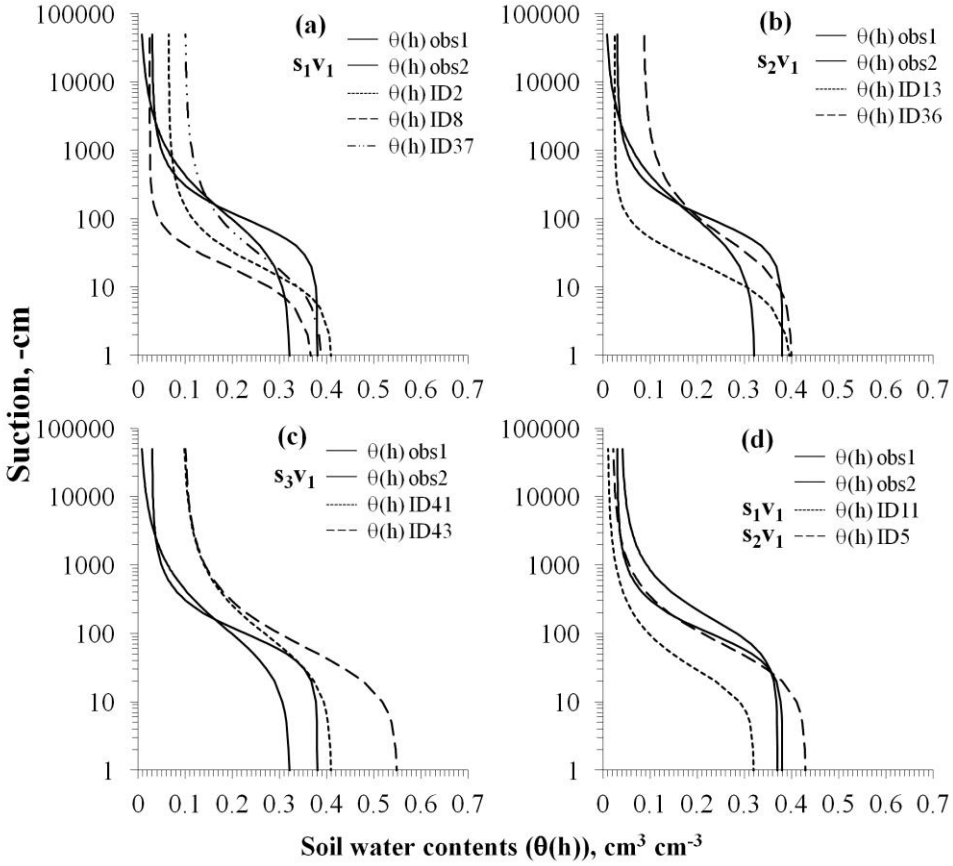

Figure 7.8: Field-observed and estimated (from soil ID values) water retention curves (θ(h)); (a-c) sub regions (s_1v_1, s_2v_1, s_3v_1) at the LW 1 site, (d) sub-regions (s_1v_1 and s_2v_1) at the LW 2 site

Table 7.5: Monthly average and ±95 PCI of soil moisture deficit index (SMDI) values of disaggregated- (s_1v_1, s_2v_1, s_3v_1) and mixed- ($s_1v_1 + s_2v_1 + s_3v_1$) scale soil-vegetation conditions for Step 4; a) s_1: sandy loam, b) s_2: silt loam, c) s_3: clay loam , d) scenario 1 - a_1: 20, a_2: 30, a_3: 50, e) scenario 2 - a_1: 30, a_2: 30, a_3: 40, f) scenario 3 - a_1: 80, a_2: 10, a_3: 10 with a wheat cover (v_1)

Month	Sandy loam (s_1v_1)			Silt loam (s_2v_1)			Clay loam (s_3v_1)			Scenario 1 Mixed (20-30-50)			Scenario 2 Mixed (30-30-40)			Scenario 3 Mixed (80-10-10)		
	Avg.	Pos 95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI
March	1.5	1.6	1.4	2.1	2.1	2.0	1.4	1.6	1.2	1.7	1.8	1.6	1.7	1.8	1.6	1.6	1.7	1.5
April	-1.4	-1.0	-1.8	-0.6	-0.2	-0.9	-1.3	-0.9	-1.8	-1.0	-0.5	-1.5	-1.1	-0.6	-1.5	-1.0	-0.5	-1.5
May	-2.1	-1.5	-2.7	-1.6	-1.0	-2.3	-1.7	-1.0	-2.5	-1.5	-0.7	-2.2	-1.7	-1.0	-2.4	-1.5	-0.8	-2.2
June	-2.1	-1.5	-2.7	-1.8	-1.4	-2.3	-1.8	-0.9	-2.8	-1.4	-0.8	-2.1	-2.0	-1.2	-2.7	-1.4	-0.9	-1.8
July	-2.2	-1.3	-3.0	-2.2	-1.7	-2.8	-2.1	-1.2	-3.1	-2.0	-1.3	-2.7	-2.2	-1.4	-3.0	-1.9	-1.2	-2.7

Table 7.6: Monthly average and ±95 PCI of disaggregated (s_1v_1, s_2v_1, and s_3v_1) and mixed ($s_1v_1+s_2v_1+s_3v_1$) scale SMDI values with various vegetation covers for Scenario 2 during the current decade (2010-2020)

Index	March			April			May			June			July		
	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI
s_1v_1	1.5	1.6	1.4	-1.4	-1.0	-1.8	-2.1	-1.5	-2.7	-2.1	-1.5	-2.7	-2.2	-1.3	-3.0
s_2v_1	2.1	2.1	2.0	-0.6	-0.2	-0.9	-1.6	-1.0	-2.3	-1.8	-1.4	-2.3	-2.2	-1.7	-2.8
s_3v_1	1.4	1.6	1.2	-1.3	-0.9	-1.8	-1.7	-1.0	-2.5	-1.8	-0.9	-2.8	-2.1	-1.2	-3.1
Mixed	1.7	1.8	1.6	-1.1	-0.6	-1.5	-1.7	-1.0	-2.4	-2.0	-1.2	-2.7	-2.2	-1.4	-3.0
s_1v_2	1.9	2.1	1.7	0.1	0.4	-0.2	-1.2	-0.6	-1.9	-1.7	-0.9	-2.4	-2.4	-1.6	-3.1
s_2v_2	2.4	2.5	2.4	1.2	1.4	0.9	-0.7	-0.1	-1.3	-2.0	-1.4	-2.7	-2.8	-2.2	-3.4
s_3v_2	1.4	1.8	1.0	0.2	0.7	-0.2	-0.6	0.1	-1.3	-1.2	-0.1	-2.3	-2.1	-0.9	-3.2
Mixed	1.9	2.2	1.7	0.6	0.9	0.2	-0.7	0.0	-1.4	-1.6	-0.6	-2.5	-2.3	-1.4	-3.3
s_1v_3	1.9	2.2	1.6	-0.3	0.2	-0.7	-0.9	-0.2	-1.6	-1.3	-0.5	-2.1	-1.7	-1.0	-2.5
s_2v_3	2.3	2.5	2.2	1.1	1.4	0.8	-0.6	0.1	-1.3	-1.9	-1.1	-2.7	-2.7	-1.9	-3.5
s_3v_3	0.5	1.0	0.0	-0.6	0.1	-1.3	-0.3	0.5	-1.1	-0.9	0.1	-2.0	-1.2	0.0	-2.3
Mixed	1.6	2.0	1.3	0.3	0.8	-0.2	-0.6	0.2	-1.3	-1.5	-0.5	-2.4	-2.0	-1.1	-3.0
s_1v_4	1.9	2.1	1.7	0.3	0.6	0.0	-1.2	-0.5	-1.9	-2.0	-1.2	-2.8	-2.5	-1.8	-3.1
s_2v_4	2.5	2.6	2.4	1.3	1.5	1.1	-0.7	-0.1	-1.2	-2.3	-1.7	-2.9	-2.9	-2.4	-3.5
s_3v_4	1.5	1.9	1.1	0.5	1.0	-0.1	-0.5	0.2	-1.2	-1.5	-0.4	-2.7	-2.4	-1.3	-3.4
Mixed	2.1	2.3	1.9	0.8	1.2	0.5	-0.6	0.0	-1.3	-1.9	-1.0	-2.8	-2.6	-1.8	-3.3

s_1: sandy loam, s_2: clay loam, s_3: clay loam, v_1: wheat, v_2: soybean, v_3: grass, v_4: maize

Table 7.7: Monthly average and ±95 PCI of disaggregated (s_1v_1, s_1v_2, s_2v_1, s_2v_2, s_3v_1, s_3v_2) and mixed (s_1v_1+s_1v_2+s_2v_1+ s_2v_2+s_3v_1+s_3v_2) scale SMDI values with complex land surface conditions during the current decade (2010-2020)

Index	March			April			May			June			July		
	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI
s_1v_1	1.5	1.6	1.4	-1.4	-1.0	-1.8	-2.1	-1.5	-2.7	-2.1	-1.5	-2.7	-2.2	-1.3	-3.0
s_1v_2	1.9	2.1	1.7	0.1	0.4	-0.2	-1.2	-0.6	-1.9	-1.7	-0.9	-2.4	-2.4	-1.6	-3.1
s_2v_1	2.1	2.1	2.0	-0.6	-0.2	-0.9	-1.6	-1.0	-2.3	-1.8	-1.4	-2.3	-2.2	-1.7	-2.8
s_2v_2	2.4	2.5	2.4	1.2	1.4	0.9	-0.7	-0.1	-1.3	-2.0	-1.4	-2.7	-2.8	-2.2	-3.4
s_3v_1	1.4	1.6	1.2	-1.3	-0.9	-1.8	-1.7	-1.0	-2.5	-1.8	-0.9	-2.8	-2.1	-1.2	-3.1
s_3v_2	1.4	1.8	1.0	0.2	0.7	-0.2	-0.6	0.1	-1.3	-1.2	-0.1	-2.3	-2.1	-0.9	-3.2
Mixed	1.9	2.0	1.8	0.1	0.4	-0.2	-1.0	-0.3	-1.6	-1.8	-0.9	-2.6	-2.3	-1.4	-3.1

s_1: sandy loam, s_2: clay loam, s_3: clay loam, v_1: wheat, v_2: soybean

7.4.5 Step 5: Field Validation Experiments and Drought Severity Assessments in the Future

In Step 5, this approach was applied to several field sites. Table 7.8 and 7.9 present the filtered/un-filtered (soil ID and sub-grid fractions) solutions at the WC (1 and 2) and LW (1 and 2) sites, respectively. The searched soil ID values and sub-grid fractions at the WC2 and LW2 sites, which had relatively simple land surface (2 soils and 1 vegetation cover-grass), appeared to be more identifiable than those of the others WC1 and LW1 (3 soils and 1 vegetation-grass) at the field scales. Overall, uncertainties of the unfiltered and filtered sub-grid fractions were similarly shown for both the WC (1 and 2) and LW (1 and 2) sites, but the unfiltered soil ID values had more variations in modeling compared to those of the filtered solutions. Generally, land surface is heterogeneous in terms of soil textures in a real world scenario, but the soil characteristics are assumed to be homogeneous for modeling purposes. It showed that ignoring the heterogeneity of soil textures within a RS pixel may cause errors in the model predictions at the airborne-scale (800 m × 800 m).

The water retention curves ($\theta(h)$) derived from the estimated soil ID values for all sub-regions were compared with the field-observations obtained near the LW1 and 2 sites as shown in Fig. 7.8. Although limited observed $\theta(h)$ curves were used for validation, the results for sub-regions at the LW1 and 2 sites matched the observed $\theta(h)$ functions with small uncertainties demonstrating the robustness of our approach, especially for the LW2 site. Overall, the estimated soil ID and sub-grid fractions at the field-scales had more uncertainties than the results under the synthetic conditions, because the field-scale experiments were affected by various environmental factors such

as the dynamic local weather conditions, heterogeneity of land surface conditions, unknown initial and bottom boundary conditions in soil profile, measurement errors, etc. Furthermore, in searching the solutions, the simulation-optimization approach also had its own weakness of model structures (e.g., improper description of rooting depth, root distribution, etc.). It is usual that the field-scale experiments have more variations than the results of synthetic conditions (step 1).

Based on the filtered solutions at the LW1 and WC1 sites, we predicted the disaggregated- and mixed-scale root zone soil moisture dynamics for the current decade (2010-2020) in Fig. 7.9. The LW1 and WC1 sites comprised of relatively complicated land surface conditions were only shown here. The predicted root zone soil moisture dynamics (quantities) showed different trends at the disaggregated-scale (s_1v_1, s_2v_1, s_3v_1 at LW1 and WC1 sites) with the grass cover. The sub-regions, which had lower hydraulic conductivities (s_3v_1: 6.2 and 7.4 mm/d at LW1; s_3v_1: 0.4 mm/d at WC1), can hold more soil water than the other sites with high K_{sat} values (LW1: 9.7~106.1 mm/d; WC1: 43.6~106.1 mm/d). It indicated that the physical soil characteristics were significantly influencing soil water contents at the disaggregated-scales. Overall, the mixed-scale root zone soil moistures for the LW1 and WC1 sites were similar with those of the disaggregated sub-grid fractions (s_2v_1 for the LW1 and s_3v_1 for the WC1), because these sub-regions had the sub-grid fractions of 61.9 % and 58.1 % within the RS pixels, respectively.

The SMDI values were similar to the trends of predicted soil moisture dynamics as shown in Table 7.10. The monthly average SMDI values at the mixed-scales were

positive for the LW 1 site, which meant that the land surface may have a wet condition. But the sub-region (s_1v_1) had negative (average) SMDI values in March and April. This shows that crops at the disaggregated-region can be damaged by the drought severity partially owing to the land surface conditions (i.e., different soil textures, vegetation cover, etc.). The average SMDI value at the sub-region (s_1v_1) from April to July was positive for the WC 1 site, except in March. The other regions (s_2v_1 and s_3v_1) also showed positive SMDI values, but the SMDI values in June and July were negative. There was a decreasing trend of SMDI starting in April, except of the sub-region (s_1v_1). This trend follows the soil moisture predictions in the sub-regions (s_2v_1 and s_3v_1), because these sub-regions (s_2v_1 and s_3v_1) had total sub-grid fractions of 77.8 % within the RS pixel. For the LW1 and WC1 sites, overall the -95PCI of SMDI values were negative (-2 to 0) during the simulation period. This meant that these sites were at a potential risk of agricultural drought in the worst scenario. Also, the average and ± 95PCI of SMDI values in July were negative probably because the simulation period for crops was set during March 01-July 31 indicating that crop growths decreased in July causing a reduction in soil moisture. This could cause the negative SMDI values in July. In this study, we evaluated drought severity based on fine-scale soil moisture predictions for individual soil-crop combinations using GCM scenarios in the current decade. These findings showed that the drought severity of each sub-region was variable based on the local land surface conditions (i.e., various soil textures, vegetations, ground water tables, etc.) in the study sites. Thus, our proposed methodology can assess drought severity at each sub-region within a RS pixel for efficient water resources management.

Table 7.8: Filtered/un-filtered solutions derived by the modified grid-based disaggregation algorithm using the EMOGA at the Walnut Creek (WC 1 and 2) sites

Field sites	Categories	Observations		Derived solutions					
				Filtered			Un-filtered		
		Soil ID	Sub-grid fractions	Soil ID (Soil texture)	Average of sub-grid fractions ($a_{i,j}$)	SD of sub-grid fractions	Soil ID (Soil texture)	Average of sub-grid fractions ($a_{i,j}$)	SD of sub-grid fractions
WC 1 (v_1: grass)	s_1	Loam	12	2 (Sandy loam)	22.2	4.0	2 (Sandy loam)	22.6	4.9
	s_2	Silty clay	12	10 (SsandO4), 11 (SsandO5)	19.8	9.0	1 (Sandy loam), 10 (SsandO4), 11 (SsandO5), 28 (SsiltO14), 52 (Clay)	20.8	8.7
	s_3	Clay loam	76	28 (SsiltO14)	58.0	5.9	28 (SsiltO14), 52 (Clay)	56.6	7.4
WC 2 (v_1: grass)	s_1	Clay loam	57	2 (Sandy loam)	52.0	0.0	2 (Sandy loam)	51.7	0.9
	s_2	Loam	43	49 (Silty clay)	48.0	0.0	46 (Sandy clay), 49 (Silty clay)	48.3	0.9

Population: 30; Total generations: 5000 (restart 8); seed number: 1000, 2000, 3000; p_{creep}: 0.05; p_{mutate}: 0.05

Table 7.9: Filtered/un-filtered solutions derived by the modified grid-based disaggregation algorithm using the EMOGA at the Little Washita (LW 1 and 2) sites

Field sites	Categories	Observations		Derived solutions					
				Filtered			Un-filtered		
		Soil ID	Sub-grid fractions ($a_{i,j}$)	Soil ID (Soil texture)	Average of sub-grid fractions ($a_{i,j}$)	SD of sub-grid fractions	Soil ID (Soil texture)	Average of sub-grid fractions ($a_{i,j}$)	SD of sub-grid fractions
LW 1 (v_1: grass)	s_1	f sandy loam	44	2 (Sandy loam), 8 (Scoarse), 37 (Sandy Clay Loam)	28.2	9.0	1 (Sandy loam), 2 (Sandy loam), 8 (Scoarse), 9 (SsandO3), 37 (Sandy clay loam)	24.9	9.1
	s_2	Silt loam	47	13 (Tcoarse), 36 (Sandy clay loam)	61.9	13.9	13 (Tcoars), 15 (Loam), 36 (Sandy Clay Loam)	63.5	12.6
	s_3	Loam	9	41 (Clay loam), 43 (Silty clay loam)	9.9	8.8	6 (TsandB3), 14 (Loam), 28 (SsiltO14), 34 (Silt loam), 39 (Tfine), 40 (Clay loam), 41 (Clay loam), 43 (Silty clay loam), 48 (Silty clay)	11.6	7.0
LW 2 (v_1: grass)	s_1	Sandy loam	38	11 (SsandO5)	33.0	0.0	2 (Sandy loam), 3 (Sandy loam), 11 (Sandy loam)	25.7	10.5
	s_2	Silt loam	62	5 (Tsand B2)	67.0	0.0	5, 9 (SsandO3)	74.3	10.5

Population: 30; Total generations: 5000 (restart 8); seed number: 1000, 2000, 3000; p_{creep}: 0.05; p_{mutate}: 0.05

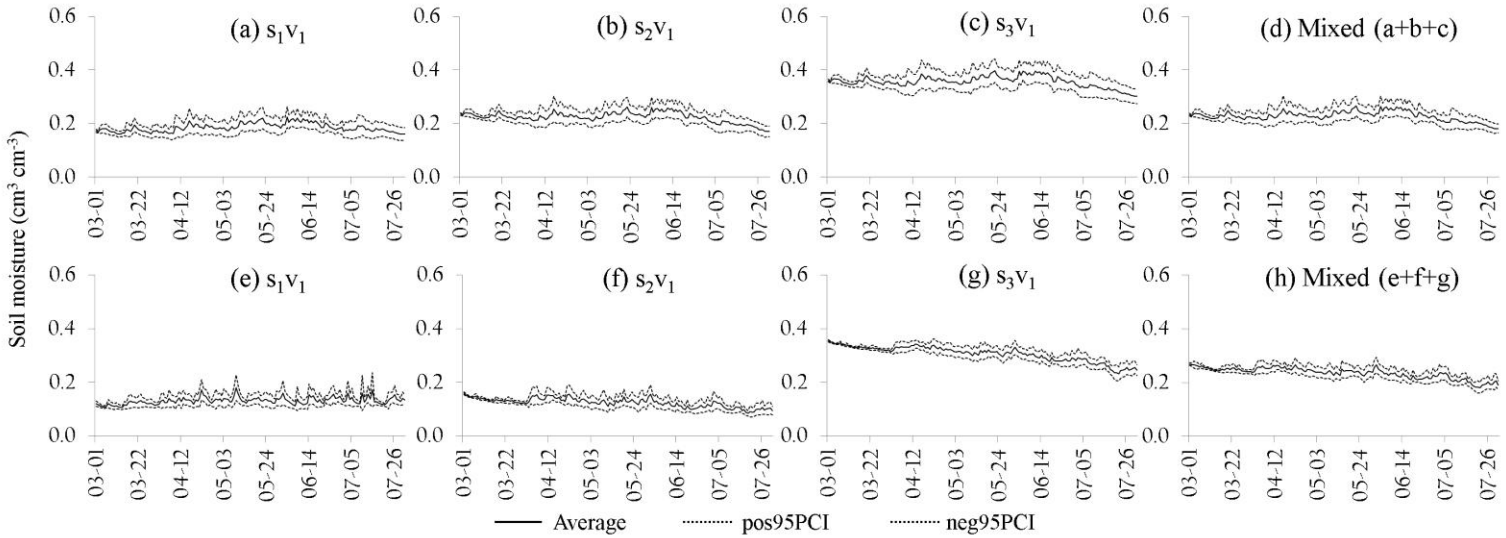

Figure 7.9: Disaggregated- (s_1v_1, s_2v_1, s_3v_1) and mixed- ($s_1v_1 + s_2v_1 + s_3v_1$) scale root zone soil moisture predictions with the grass cover; (a-d) Little Washita (LW1) site and (e-h) Walnut Creek (WC1) site

Table 7.10: Monthly average and ±95 PCI of disaggregated (s_1v_1, s_2v_1, and s_3v_1) and mixed ($s_1v_1+s_2v_1+s_3v_1$) scale SMDI values with the grass cover at the Little Washita (LW1) and Walnut Creek (WC1) sites during the current decade (2010-2020)

Sites	Month	Disaggregated-scale									Mixed-scale		
		s_1v_1			s_2v_1			s_3v_1			$s_1v_1+s_2v_1+s_3v_1$		
		Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI	Avg.	pos95PCI	neg95PCI
LW1	March	-0.7	-0.2	-1.3	0.2	0.5	0.0	-0.1	0.3	-0.4	0.0	0.3	-0.3
	April	-0.6	0.4	-1.6	0.0	0.9	-1.0	-0.1	0.8	-1.0	-0.1	0.9	-1.1
	May	0.2	1.1	-0.6	0.4	1.2	-0.4	0.4	1.3	-0.5	0.4	1.3	-0.4
	June	0.5	1.2	-0.1	0.5	1.2	-0.2	0.5	1.2	-0.2	0.6	1.2	-0.1
	July	-0.9	0.1	-1.9	-1.4	-0.5	-2.4	-1.3	-0.3	-2.3	-1.4	-0.5	-2.3
WC1	March	-1.1	-0.4	-1.7	0.6	0.9	0.4	1.3	1.7	1.0	0.9	1.3	0.6
	April	0.0	0.6	-0.6	0.8	1.4	0.3	1.3	1.8	0.9	1.2	1.6	0.7
	May	-0.1	0.5	-0.6	0.0	0.4	-0.5	0.3	0.9	-0.3	0.2	0.8	-0.3
	June	0.1	0.7	-0.4	-0.9	-0.3	-1.4	-0.7	-0.2	-1.1	-0.6	-0.1	-1.1
	July	0.4	0.9	-0.2	-1.7	-1.1	-2.3	-2.0	-1.6	-2.5	-1.8	-1.4	-2.3

7.5 Conclusions

We explored a drought severity assessment framework using remotely sensed soil moisture footprints with global circulation model (GCM) scenarios at fine sub-grid scales for the current decade (2010-2020). The soil hydraulic properties were quantified from remotely sensed soil moisture footprints using the newly developed simulation-optimization scheme by coupling SWAP with EMOGA. The estimated soil parameters were used along with climatic variables provided by GCMs to predict fine-scale soil moisture dynamics. Finally, we evaluated drought severity using SMDI based on the predicted soil moisture at finer-scales. Synthetic and field experiments were conducted for testing the new approach under different hydro-climates (Lubbock, Texas; Little Washita watershed, Oklahoma; and Walnut Creek watershed, Iowa). These experiments were comprised of different combinations of soil textures, vegetation covers, and ground water table depths.

The estimated solutions under the synthetic case matched well with the target values for the simple/complex land surface conditions indicating that this approach was able to extract the available land surface information from a RS footprint. The predictions of disaggregated (s_1v_1, s_2v_1, and s_3v_1) soil moisture derived by the estimated solutions were affected more by the soil textures, whereas the vegetation covers influenced the soil moisture dynamics at the mixed-scale ($s_1v_1+s_2v_1+s_3v_1$). A shallow water table depth of -50 cm influenced the root zone soil moisture quantity by leading to an increase, but there were barely any impacts due to deeper water table depths of -200, -150, and -100 cm. The SMDI values showed a similar trend with the predicted root zone

soil moisture. The SMDI values were increased by conditioning on various soil textures, vegetation covers, and presence of ground table depths at local-scales, but crops were still at a potential risk of drought severity during April to July. This indicated that conditioning of soil texture, vegetation, and groundwater table depth can exert only a limited influence in overcoming a drought. In the field validation experiments, the results at the WC and LW sites had more uncertainties compared to the experiment under the synthetic conditions. This was attributed to the model performance being influenced by not only the soil, vegetation, and ground water, but also by the limitations in model structures and observation errors. Although the results at the field sites had some variations in estimating the soil parameters and sub-grid fractions, the derived solutions matched the GIS-based observations. The comparison of field-observed and estimated water retention curves $\theta(h)$ for the LW1 and 2 sites supports the robustness of our approach. Overall, positive monthly average SMDI values at the LW1 and WC1 sites indicated that these sites had less drought risk for the current decade. However, the average SMDI values of sub-regions were variable indicating that these sub-sites could be partially affected by the drought severity in March and April.

In this study, our approach performed well in extracting land surface information within a RS product. According to the predicted root zone soil moisture and SMDI based on the derived land surface information and bias corrected CGCM2 scenarios, the field sites can be influenced by drought conditions during crop growing season for the current decade. This approach has been demonstrated to be helpful in forecasting root zone soil moisture dynamics and assessing drought severity at the field-scales.

CHAPTER VIII

GENERAL CONCLUSIONS

In this dissertation, we focused on quantifying soil hydraulic properties at multiple scales across the land surface. A new inversion model, various scaling (downscaling and upscaling) algorithms, optimization techniques, and a stochastic evolutionary approach were developed with a goal to contribute to the understanding of water resources/agricultural water management, rainfall-runoff processes, seasonal climate predictions, and ecosystem dynamics and alteration.

A layer-specific soil moisture assimilation scheme in Chapter II was developed for determining the soil hydraulic properties in the layered soil profile. In Chapter III, a new inverse modeling algorithm by combining soil moisture (SM) based Noisy Monte Carlo Genetic Algorithm (NMCGA) and Surface Energy Balance Algorithm for Land (SEBAL) based evapotranspiration (ET) products (using remote sensing data) was tested for quantifying pixel-scale soil hydraulic properties at multiple scales.

In Chapters IV and V, new deterministic downscaling and scaling (joint downscaling and upscaling) algorithms were developed for improving usage of remotely sensed soil moisture footprints and *in-situ* data in complex terrains at various hydro-climatic regions. In order to predict root zone soil moisture in the soil profile using rainfall data and (limited) soil moisture history at the network of multiple weather stations under two different hydro-climatic regions, a new non-parametric evolutionary algorithm (genetic algorithm-based hidden Markov model, HMMGA) was developed in Chapter VI.

A drought severity assessment platform based on a grid-based disaggregation algorithm adapting a combined optimization and hydrological model (Soil-Water-Atmosphere-Plant, SWAP and Ensemble Multiple Operators Genetic Algorithm, EMOGA) was developed in Chapter VII. We demonstrated that soil layers and vertical heterogeneity could impact the uncertainty of quantifying soil hydraulic parameters. Although the sub-surface flows dominated by the upward fluxes were more elusive, this approach successfully matched root zone soil moisture estimates with observations at the individual soil layers suggesting that this approach could be applied in real world conditions.

Considering both soil moisture and evapotranspiration components in the optimization algorithm improved the estimations of soil hydraulic properties and reduced their uncertainties better than those of using SM-only (Noisy Monte Carlo Genetic Algorithm, NMCGA) or ET-only (Surface Energy Balance Algorithm for Land, SEBAL). Also it improved the predictions of soil moisture dynamics in the deep soil depth (180-200 cm) dominated by upward flows with the presence of ground water table depth (-100 cm from the soil surface). These results demonstrated that the ET component plays the key role in estimating the soil hydraulic parameters along the soil column as well as soil moisture.

The deterministic disaggregation algorithm estimated well the soil moisture dynamics of sub-pixels from synthetic and field validation experiments with the observations under heterogeneous land surface condition with uncertainties. It indicated that the ET component can be used to capture the heterogeneity of land surface within a

remotely sensed soil moisture data. The good correspondence of observed water retention ($\theta(h)$) functions (from the soil core samples) and soil parameters searched by the genetic algorithm at the study (LW 13 and 21) sites showed the robustness of this algorithm.

Furthermore, the scaling (joint downscaling and upscaling) algorithm based on inversion model could scale down and up near-surface soil moisture estimates considerably well in the airborne-/satellite-scales compared to the *in-situ* root zone soil moisture measurement. Although the upscaling approach was excluded at the satellite-scale due to no available *in-situ* root zone soil moisture datasets, these findings demonstrated that our algorithm performs well in scaling down and up across complicated land surface at various scales.

The hidden Markov model genetic algorithm (HMMGA) performs quite well in forecasting rainfall occurrence probabilities, dry/wet spell lengths, and daily root zone soil moisture dynamics, although uncertainties were included in predicted estimates. Especially, the measured and predicted root zone soil moisture estimates were considerably affected by both the precipitation and land surface characteristics (e.g., soil texture, vegetation covers, topography, etc.). We demonstrated that our approach could provide reasonable predictions over multiple locations with the historical precipitation and (limited) root zone soil moisture data in the future.

A drought severity assessment platform based on a grid-based disaggregation algorithm adapting a combined optimization and hydrological model (SWAP-EMOGA) executed quite well in disaggregating the RS soil moisture products under the synthetic

and field validation experiments, although the uncertainties due to the RS resolutions and variability of climatic conditions contribute to the modeling performance. Finer-scale root zone soil moisture predictions were considerably influenced by various combinations of environmental factors (soils, crops, groundwater table, etc.) along with GCM scenarios.

However, environmental factors had relatively limited impacts on reducing drought severity. The absolute soil moisture deficit index values do indicate the occurrence of agricultural drought during 2010-2020. Thus, our proposed approach can be used to assess drought severity at finer-scales using a remotely sensed soil moisture product for efficient agricultural/water resources management.

The methodologies developed in this dissertation can contribute significantly to the spatial and temporal analysis of root zone soil moisture using remotely sensed and *in-situ* soil moisture data at the multiple scales in the real world. Also, our proposed approaches may provide better input parameters for large-scale hydrologic and hydro-climatic models, resulting in better understanding of the hydrologic cycle. Furthermore, a better understanding of water cycle would help us to be better prepared for sustainable water resources, agricultural production, and devastating natural disasters in the real world.

REFERENCES

Abbaspour, K. C., M. T. van Genuchten, R. Schulin, and E. Schläppi (1997), A sequential uncertainty domain inverse procedure for estimating subsurface flow and transport parameters, *Water Resources Research, 33*(8), 1879–1892, doi:10.1029/97WR01230.

Adegoke, J. O., and A. M. Carleton (2002), Relations between soil moisture and satellite vegetation indices in the U.S. Corn Belt, *American Meteorological Society, 3*(4), 395–405.

Ahmad, M., W. G. M. Bastiaanssen, and R. A. Feddes (2002), Sustainable use of groundwater for irrigation: A numerical analysis of the subsoil water fluxes, *Irrigation and Drainage, 51*(3), 227–241.

Allen, R. G., M. Tasumi, and R. Trezza (2007), Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC) model, *Journal of Irrigation and Drainage Engineering*, *133*(4), 380–394, doi:10.1061/(ASCE)0733-9437(2007)133:4(380).

Aronica, G., P. D. Bates, and M. S. Horritt (2002), Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE, *Hydrological Processes, 16*, 2001–2016, doi:10.1002/hyp.398.

Baghdadi, N., S. Gaultier, and C. King (2002), Retrieving surface roughness and soil moisture from synthetic aperture radar (SAR) data using neural networks, *Canadian Journal of Remote Sensing*, *28*(5), 707–711.

Barron, J., J. Rockström, F. Gichuki, and N. Hatibu (2003), Dry spell analysis and maize yields for two semi-arid location in east Africa, *Agricultural and Forest Meteorology*, *117*, 23–37, doi:10.1016/S0168-1923(03)00037-6.

Bastiaanssen, W. G. M., M. Menenti, R. A. Feddes, and A. A. M. Holtslag (1998), A remote sensing surface energy balance algorithm for land (SEBAL): I. Formulation, *Journal of Hydrology*, *212*, 198–212.

Bastiaanssen, W. G. M., E. J. M. Noordman, H. Pelgrum, G. Davids, B. P. Thoreson, and R. G. Allen (2005), SEBAL model with remotely sensed data to improve water-resources management under actual field criterion, *Journal of Irrigation and Drainage Engineering*, *131*(1), 85–93, doi:10.1061/(ASCE)0733-9437(2005) 131:1(85).

Bates, B. C., and E. P. Campbell (2001), A markov chain monte carlo scheme for parameter estimation and influence in conceptual rainfall-runoff modeling, *Water Resources Research*, *37*(4), 937–947, doi:10.1029/2000WR900363.

Belmans, C., J. G. Wesseling, and R. A. Feddes (1983), Simulation of the water balance of a cropped soil: SWATRE, *Journal of Hydrology*, *63*, 271–286.

Beven, K., and A. Binley (1992), The future of distributed models: Model calibration and uncertainty predictions, *Hydrological Processes*, *6*, 279–298.

Beven, K., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology, *Journal of Hydrology*, *249*, 11–29.

Bindlish, R., J. A. J. Thomas, M. K. Gasiewski, and E. G. Njoku (2006), Soil moisture mapping and AMSR-E validation using the PSR in SMEX02, *Remote Sensing of Environment*, *103*(2), 127–139, doi:10.1016/j.rse.2005.02.003.

Bonan, G. B. (1996), A land surface model (LSM version 1.0) for ecological, hydrological, and atmospheric studies: Technical Description and User's Guide, Technical Note NCAR/TN-417+STR, National Center for Atmospheric Research, Boulder, CO.

Bosch, D. D. (1991), Error associated with point observation of matric potential in heterogeneous soil profiles, *Transactions of American Society of Agricultural Engineers*, *34*(6), 2427–2436.

Bouttier, F., J. F. Mahfouf, and J. Noilhan (1993a), Sequential assimilation of soil moisture from atmospheric low-level parameters: Part I. Sensitivity and calibration studies, *Journal of Applied Meteorology, 32*(8), 1335–1351.

Bouttier, F., J. F. Mahfouf, and J. Noilhan (1993b), Sequential assimilation of soil moisture from atmospheric low-level parameters: Part II. Implementation in a mesoscale model, *Journal of Applied Meteorology*, *32*(8), 1352–1364.

Brooks, R. H., and A. T. Corey (1964), Hydraulic properties of porous media, Hydrology Paper No. 3, Colorado State University, Fort Collins, CO.

Campbell, G. S. (1974), A simple method for determining unsaturated hydraulic conductivity from moisture retention data, *Soil Science, 117*, 311–314.

Cancelliere, A., G. D. Mauro, B. Bonaccorso, and G. Rossi (2007), Drought forecasting using the standardized precipitation index, *Water Resources Management*, *21*(5), 801–819, doi:10.1007/s11269-006-9602-y.

Carroll, D. L. (1996), Genetic algorithms and optimizing chemical oxygen-Iodine lasers, in *Developments in theoretical and applied mechanics*, vol. 18, edited by H. B. Wilson et al., pp. 411–424, University of Alabama, Tuscaloosa, AL.

Carsel, R. F, and R. S. Parrish (1988), Developing joint probability-distributions of soil-water retention characteristics, *Water Resources Research, 24*(5), 755–769.

Cayan, D. R., E. P. Maurer, M. D. Dettinger, M. Tyree, and K. Hayhoe (2008), Climate change scenarios for the California region, *Climatic Change*, *87*, 21–42, doi:10.1007/s10584-007-9377-6.

Chan-Hilton, A. B., and T. B. Culver (2000), Constraint handling for genetic algorithms in optimal remediation design, *Journal of Water Resources Planning Management*, *126*(3), 128–137, doi:10.1061/(ASCE)0733-9496(2000)126:3 (128).

Charney, J. G., W. J. Quirk, S. H. Chow, and J. Kornfield (1977), A comparative study of the effects of albedo change on drought in semi-arid region, *Journal of Atmospheric Sciences*, *34*, 1366–1385.

Clark, C. A., and R. W. Arritt (1995), Numerical simulations of the effect of soil moisture and vegetation cover on the development of deep convection, *Journal of Applied Meteorology, 34*(9), 2029–2045.

Clapp, R. B., and G. M. Hornberger (1978), Empirical equations for some soil hydraulic properties, *Water Resources Research, 14*(4), 601–604, doi:10.1029/WR014i004 p00601.

Crow, W. T., D. Ryu, and J. S. Famiglietti (2005), Upscaling of field-scale soil moisture measurements using distributed land surface modeling, *Advances in Water Resources, 28*, 1–14, doi:10.1016/j.advwatres.2004.10.004.

Crow, W. T., E. F. Wood, and R. Dubayah (2000), Potential for downscaling soil moisture maps derived from space borne imaging radar data, *Journal of Geophysical Research, 105*, 2203–2212.

Dai, Y., and Q. C. Zeng (1997), A land surface model (IAP94) for climate studies: Part I. Formulation and validation in off-line experiments, *Advances Atmospheric Sciences, 14*(4), 433–460.

Dai, Y., X. Zeng, R. E. Dickinson, I. Baker, G. B. Bonan, M. G. Bosilovich, S. A. Denning, P. A. Dirmeyer, P. R. Houser, G. Niu, K. W. Oleson, A. C. Schlosser, and Z. L. Yang (2003), The common land model, *Bulletin of American Meteorological Society, 84*(8), 1013–1023.

Daroonwan, K., K. Honda, C. Sujittra, and S. A. Khun (2008), Near real-time decision support system for drought monitoring and impact assessment on rice in Ubon Ratchathani, northeastern Thailand, paper presented at 29th Asian Conference on Remote Sensing, World Conference on Agricultural Information and IT, Colombo, Sri Lanka.

Das, N. N., and B. P. Mohanty (2006), Root zone soil moisture assessment using passive microwave remote sensing and vadose zone modeling, *Vadose Zone Journal*, *5*(1), 296–307, doi:10.2136/vzj2005.0033.

Das, N. N., D. Entekhabi, and E. Njoku (2011), An algorithm for merging SMAP radiometer and radar data for high resolution soil moisture retrieval, *IEEE Transactions on Geoscience and Remote Sensing*, *49*(5), 1504–1512, doi:10.1109/TGRS.2010.2089526.

Das, N. N., B. P. Mohanty, and E. G. Njoku (2008a), A markov chain monte carlo algorithm for upscaled soil-vegetation-atmosphere-transfer modeling to evaluate satellite-based soil moisture measurements, *Water Resources Research*, *44*, doi:10.1029/2007WR006472.

Das, N. N., B. P. Mohanty, and E. G. Njoku (2008b), Characterization of backscatter by surface features in L-band active microwave remote sensing of soil moisture, *IEEE International Geoscience and Remote Sensing Symposium*, *2*, 817–820, doi:10.1109/IGARSS.2008.4779119.

De Lannoy, G. J. M., V. R. N. Pauwels, P. R. Houser, T. Gish, and N. E. C. Verhoest (2007), Representativeness of point soil moisture observations, upscaling, and assimilation, in *Quantification and prediction of predictive uncertainty for sustainable water resources management*, vol. 313, edited by IAHS Publications, pp. 249–257, Proceedings of Symposium HS2004 at IUGG2007, Perugia, Italy.

Dickinson, R., A. Henderson-Sellers, P. J. Kennedy, and M. F. Wilson (1993), Biosphere atmosphere transfer scheme (BATS) version 1e as coupled to the NCAR

Community Climate Model, Technical Note NCAR/TN-378+STR, National Center for Atmospheric Research, Boulder, CO.

Droogers, P., and W. G. M. Bastiaanssen (2002), Irrigation performance using hydrological and remote sensing modeling, *Journal of Irrigation and Drainage Engineering*, *128*(1), 11–18, doi:10.1061/(ASCE)0733-9437(2002)128:1(11).

Droogers, P., W. G. M. Bastiaanssen, M. Beyazgül, Y. Kayam, G. W. Kite, and H. Murray-Rust (2000), Distributed agro-hydrological modeling of an irrigation system in western Turkey, *Agricultural Water Management*, *43*(2), 183–202, doi:10.1016/S0378-3774 (99)00055–4.

Duan, Q. Y., S. Sorooshian, and H. V. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resources Research*, *28*(4), 1015–1031.

Efron, B. (1982), *The jackknife, the bootstrap and other resampling plans*, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Engman, T. (1991), Application of microwave remote sensing of soil moisture for water resources and agriculture, *Remote Sensing of Environment*, *35*, 213–226.

Entekhabi, D., G. R. Asrar, A. K. Betts, K. J. Beven, R. L. Bras, and C. J. Duffy (1999), An agenda for land surface hydrology research and a call for the second international hydrological decade, *Bulletin of American Meteorological Society*, *80*(10), 2043–2058.

Farrar, T. J., S. E. Nicholson, and A. R. Lare (1994), The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana: II. NDVI response to soil moisture, *Remote Sensing of Environment, 50*, 121–133.

Feddes, R. A., P. J. Kowalik, and H. Zarandy (1978), *Simulation of field water use and crop yield*, John Wiley & Sons, New York.

FEMA (1995), National mitigation strategy: Partnerships for building safer communities, Federal Emergency Management Agency, Washington, D.C.

Ferreira, M. E., L. G. Ferreira, E. E. Sano, and Y. E. Shimabukuro (2007), Spectral linear mixture modeling approaches for land cover mapping of tropical savanna areas in Brazil, *International Journal of Remote Sensing, 28*(2), 413–429, doi:10.1080/01431160500181507.

Gardner, W. R. (1958), Some steady state solutions of the unsaturated moisture flow equation with application from a water table, *Soil Science, 85*, 228–232.

Goldberg, D. E. (1989), *Genetic algorithms in search and optimization and machine learning*, Addison-Wesley Publishing, New York.

Goldberg, D. E. (2002), *The design of innovation: Lessons from and for competent genetic algorithms*, Kluwer Academic Publishers, Norwell, MA.

Gong X., A. G. Barnston, and M. N. Ward (2003), The effect of spatial aggregation on the skill of seasonal precipitation forecasts, *Journal of Climate, 16*(18), 3059–3071.

Green, R. E., L. R. Ahuja, and S. K. Chong (1986), *Hydraulic conductivity, diffusivity, and sorptivity of unsaturated soils*, American Society of Agronomy, Madison, WI.

Hansen, J. D., K. W. Rojas, and M. J. Schaffer (1999), Calibrating the root zone water quality model, *Agronomy Journal*, *91*, 171–177.

Hartigan, J., and M. Wong (1979), A k-means clustering algorithm, *Applied Statistics*, *28*, 100–108.

Hayhoe, K., D. Cayan, C. B. Field, P. C. Frumhoff, E. P. Maurer, N. L. Miller, S. C. Moser, S. H. Schneider, K. N. Cahill, E. E. Cleland, L. Dale, R. Drapek, R. M. Hanemann, L. S. Kalkstein, J. Lenihan, C. K. Lunch, R. P. Neilson, S. C. Sheridan, and J. H. Verville (2004), Emissions pathways, climate change, and impacts on California, *Proceeding of National Academy of Sciences*, *101*(34), 12422–12427.

Heathman, G. C., P. J. Starks, L. R. Ahuja, and T. J. Jackson (2003), Assimilation of surface soil moisture to estimate soil water content, *Journal of Hydrology*, *279*, 1–17, doi:10.1016/S0022-1694(03)00088-X.

Hewitson, B. C., and R. G. Crane (1996), Climate downscaling: Techniques and application, *Climate Research*, *7*, 85–95.

Holben, B. N., and Y. E. Shimabukuro (1993), Linear mixing applied to coarse spatial resolution data from multispectral satellite sensors, *International Journal of Remote Sensing*, *14*(11), 2231–2240.

Holland, J. H. (1975), *Adaptation in natural and artificial system*, University of Michigan press, Ann Arbor, MI.

Hollinger, S. E., and S. A. Isard (1994), A soil moisture climatology of Illinois, *Journal of Climate*, *7*(5), 822–833.

Holtan, H. N. (1961), A concept for infiltration estimates in watershed engineering, Bulletin 41–51, U.S. Department of Agriculture, Academic press, New York.

Hughes, J. P., P. Guttorp, and S. P. Charles (1999), A non-homogeneous hidden Markov model for precipitation occurrence, *Applied Statistics*, *48*, 15–30.

Illston, B. G., J. B. Basara, D. K. Fisher, R. Elliott, C. A. Fiebrich, K. C. Crawford, K. Humes, and E. Hunt (2008), Mesoscale monitoring of soil moisture across a statewide network, *Journal of Atmospheric and Oceanic Technology*, *25*, 167–182, doi:10.1175/2007JTECHA993.1.

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

Ines, A. V. M., and J. W. Hansen (2006), Bias correction of daily GCM rainfall for crop simulation studies, *Agricultural and Forest Meteorology*, *138*, 44–53, doi:10.1016/j.agrformet.2006.03.009.

Ines, A. V. M., and K. Honda (2005), On quantifying agricultural and water management practices from low spatial resolution RS data using genetic algorithms: A numerical study for mixed pixel environment, *Advances in Water Resources*, *28*(8), 856–870, doi:10.1016/j.advwatres.2004.11.015.

Ines, A. V. M., and B. P. Mohanty (2008a), Near-surface soil moisture assimilation for quantifying effective soil hydraulic properties using genetic algorithm: 1. Conceptual modeling, *Water Resources Research*, *44*, W06422, doi:10.1029/2007WR005990.

Ines, A. V. M., and B. P. Mohanty (2008b), Near-surface soil moisture assimilation for quantifying effective soil hydraulic properties under different hydro-climatic conditions, *Vadose Zone Journal*, *7*, 39–52, doi:10.2136/vzj2007.0048.

Ines, A. V. M., and B. P. Mohanty (2009), Near-surface soil moisture assimilation for quantifying effective soil hydraulic properties using genetic algorithm: 2. Using airborne remote sensing during SGP97 and SMEX02, *Water Resources Research*, *45*, W01408, doi:10.1029/2008WR007022.

Ines, A. V. M., B. P. Mohanty, and Y. Shin (2012), A disaggregation algorithm for remotely sensed soil moisture, *Water Resources Research*, Submitted.

IPCC (2007), The physical science basis, contribution of working group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge university press, New York.

Jackson, T. J, D. M. Le Vine, A. Y. Hsu, A. Oldak, and P. J. Starks (1999), Soil moisture mapping at regional scales using microwave radiometry: The Southern Great Plains hydrology experiment, *IEEE Transactions on Geoscience and Remote Sensing*, *37*(5), 2136–2151.

Jackson, T. J., R. Bindlish, A. J. Gasiewski, B. Stankov, M. Klein, E. G. Njoku, D. Bosch, T. L. Coleman, C. Laymon, and P. J. Starks (2005), Polarimetric scanning

radiometer C and X band microwave observations during SMEX03, *IEEE Transactions on Geoscience and Remote Sensing*, *43*(11), 2418–2430, doi:10.1109/TGRS.2005.857625.

Jain, A. K., M. N. Murty, and P. J. Flynn (1999), Data clustering: A review, *ACM Computing Surveys*, *31*(3), 264–323.

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

Jana, R. B., and B. P. Mohanty (2012b), A topography-based scaling algorithm for soil hydraulic parameters at hill-slope scales: Field testing, *Water Resources Research, 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 hill-slope scale, *Water Resources Research, 48*, W02520, doi:10.1029/2010WR010185.

Jana, R. B., B. P. Mohanty, and E. P. Springer (2007), Multiscale pedotransfer functions for soil water retention, *Vadose Zone Journal*, *6*(4), 868–878, doi:10.2136/vzj2007.0055.

Jana R. B., B. P. Mohanty, and E. P. Springer (2008), Multiscale bayesian neural networks for soil water content estimation, *Water Resources Research, 44*, W08408, doi:10.1029/2008WR006879.

Jobbágy, E. G., and R. G. Jackson (2004), Groundwater use and salinization with grassland afforestation, *Global Change Biol*ogy, *10*, 1299–1312, doi:10.1111/ j.1365-2486.2004.00806.x.

Kerr, Y. H., P. Waldteufel, J. P. Wigneron, J. M. Martinuzzi, J. Font, and M. Berger (2001), Soil moisture retrieval from space: The soil moisture and ocean salinity (SMOS) mission, *IEEE Transactions on Geoscience and Remote Sensing*, *39*(8), 1729–1735.

Kim, G., and A. P. Barros (2002), Downscaling of remotely sensed soil moisture with a modified fractal interpolation method using contraction mapping and ancillary data, *Remote Sensing of Environment*, *83*, 400–413.

Kirshner, S. (2005), Modeling of multivariate time series using hidden markov models, Ph.D dissertation, Department of Information and Computer Science, University of California, Irvine, California.

Kool, J. B., and J. C. Parker (1988), Analysis of the inverse problem for transient unsaturated flow, *Water Resources Research*, *24*, 817–830.

Krishnakumar, K. (1989), Microgenetic algorithms for stationary and nonstationary function optimization, *SPIE Intelligent Control and Adaptive Systems*, *1196*, 289–296.

Krishnamurti, T. N., S. O. Han, and V. Misra (1995), Prediction of the dry and wet spell of the Austrian monsoon, *International Journal of Climatology*, *15*, 753–771.

Kroes, J. G., J. C. van Dam, J. Huygen, and R. W. Vervoort (1999), User's guide of SWAP version 2.0: Simulation of water, solute transport, and plant growth in the

soil-atmosphere-plant environment, Report 81, DLO Winand Staring Centre, Wageningen, Netherlands.

Kuczera, G., and E. Parent (1998), Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The metropolis algorithm, *Journal of Hydrology*, *211*, 69–85.

Lall, U., B. Rafagopalan, and D. G. Tarboton (1996), A nonparametric wet/dry spell model for resampling daily precipitation, *Water Resources Research*, *32*(9), 2803–2823.

LBB (2011), Fiscal impact of drought on state agencies and public institutions of higher education, fiscal year 2011, Legislative budget board, Austin, TX.

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 and F. J. Leij, pp.1269–1281, Proceeding of International Workshop, University of California, Riverside, CA.

Leung, L. R., Y. Qian, X. Bian, W. M. Washington, J. Han, and J. O. Roads (2004), Mid-century ensemble regional climate change scenarios for the western United States, *Climate change*, *62*, 75–113.

Liang, X., E. F. Wood, and D. P. Lettenmaier (1996), Surface soil moisture parameterization of the VIC-2L model: Evaluation and modifications, *Global and Planetary Change*, *13*, 195–206.

Liou, Y. A., S. F. Liu, and W. J. Wang (2001), Retrieving soil moisture from simulated brightness temperatures by a neural network, *IEEE Transactions on Geoscience and Remote Sensing*, *39*(8), 1662–1672.

Liu, W. T., and F. N. Kogan (1996), Monitoring regional drought using the vegetation condition index, *International Journal of Remote Sensing*, *17*(14), 2761–2782.

Mahfouf, J. F. (1990), A variational assimilation of soil moisture in meteorological models, paper presented at International Symposium on Assimilation of Observations in Meteorology and Oceanography, World Meteorological Organization, Clermont Ferrand, France.

Mallants, D., B. P. Mohanty, D. Jacques, and J. Feyen (1996), Spatial variability of hydraulic properties in a multi-layered soil, *Soil Science*, *161*(3), 167–181.

Maurer, E. P. (2007), Uncertainty in hydrologic impacts of climate change in the Sierra Nevada, California under two emissions scenarios, *Climatic Change*, *82*, 309–325, doi:310.1007/s10584-10006-19180-10589.

Maurer, E. P., and P. B. Duffy (2005), Uncertainty in projections of streamflow changes due to climate change in California, *Geophysical Research Letters*, *32*, doi:10.1029/2004GL021462.

McKee, T. B., N. J. Doesken, and J. Kleist (1993), The relationship of drought frequency and duration to time scales, paper presented at 8th Conference on Applied Climatology, American Meteorological Society, Anaheim, CA.

Merlin, O., G. Chehbouni, Y. Kerr, E. G. Njoku, and D. Entekhabi (2005), A combined modeling and multi-spectral/multi-resolution remote sensing approach for

disaggregation of surface soil moisture: Application to SMOS configuration, *IEEE Transactions on Geoscience and Remote Sensing*, *43*(9), 2036–2050, doi:10.1109/tgrs.2005.853192.

Miller, B. L. (1997), Noise sampling and efficient genetic algorithms, Report 97001, Illinois Genetic Algorithms Laboratory, University of Illinois, Urbana, IL.

Miller, B. L., and D. E. Goldberg (1996), Optimal sampling for genetic algorithms, in *Intelligent Engineering Systems through Artificial Neural Networks (ANNIE '96)*, vol. 6, edited by C. H. Dagli et al., pp. 291–298, American Society of Mechanical Engineering Press, New York.

Mohanty, B. P., and J. Zhu (2007), Effective hydraulic parameters in horizontally and vertically heterogeneous soils for steady-state land-atmosphere interaction, *Journal of Hydrometeorology, 8*, 715–729, doi:10.1175/JHM606.1.

Mohanty, B. P., R. S. Kanwar, and C. J. Everts (1994), Comparison of saturated hydraulic conductivity measurement methods for a glacial till soil, *Soil Science Society of America Journal*, *58*(3), 672–677.

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 Resources Research*, *38*(5), doi:10.1029/2000WR000076.

Moran, M. S., S. J. Maas, and P. J. Pinter Jr. (1995), Combining remote sensing and modeling for estimating surface evaporation and biomass production, *Remote Sensing Reviews*, *12*, 335–353.

Moulin, S., A. Bondeau, and R. Delécolle (1998), Combining agricultural crop models and satellite observations: From field to regional scales, *International Journal of Remote Sensing*, *19*(6), 1021–1036.

Mualem, Y. (1976), A new model for predicting the hydraulic conductivity of unsaturated porous media, *Water Resources Research, 12*(3), 513–522.

Nakicenovic, N., J. Alcamo, G. Davis, B. D. Vries, J. Fenhann, S. Gaffin, K. Gregory, A. Grübler, T. Y. Jung, T. Kram, E. L. La Rovere, L. Michaelis, S. Mori, T. Morita, W. Pepper, H. Pitcher, L. Price, K. Raihi, A. Roehrl, H. H. Rogner, A. Sankovski, M. Schlesinger, P. Shukla, S. Smith, R. Swart, S. van Rooijen, N. Victor, and Z. Dadi (2000), Emissions Scenarios, contribution of working group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, New York.

Narasimhan, B., and R. Srinivasan (2005), Development and evaluation of soil moisture deficit index (SMDI) and evapotranspiration deficit index (ETDI) for agricultural drought monitoring, *Agricultural and Forest Meteorology*, *133*, 69–88, doi:10.1016/j.agrformet.2005.07.012.

NCEP (2005), The community noah land-surface model (LSM): User's Guide Public Release Version 2.7.1, ftp://emc.ncep.noaa.gov/mmb/gcp/ldas/noahlsm/ver2.7.1.

Nielsen, D. R., J. W. Biggar, and K. T. Her (1973), Spatial variability of field measured soil-water properties, *Hilgardia*, *42*, 215–259.

Njoku, E. (2004), SMEX03 AMSR-E daily gridded soil moisture and brightness temperatures, National Snow and Ice Data Center, Boulder, CO.

Njoku, E. (2008), AMSR-E/Aqua daily L3 surface soil moisture, interpretive parameters & QC EASE-Grids V002, National Snow and Ice Data Center, Boulder, CO.

Njoku, E., and D. Entekkabi (1996), Passive microwave remote sensing of soil moisture, *Journal of Hydrology*, *184*, 101–129.

Nosetto, M. D., E. G. Jobbágy, R. B. Jackson, and G. A. Sznaider (2009), Reciprocal influence of crops and shallow ground water in sandy landscapes of the Inland Pampas, *Field Crops Research*, *113*, 138–148.

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), Report NCAR/TN-478+STR, National Center for Atmospheric Research, Boulder, CO.

Oleson, K. W., Y. Dai, G. Bonan, M. Bosilovich, R. Dickinson, P. Dirmeyer, F. Hoffman, P. Houser, S. Levis, G. Y. Niu, P. Thornton, M. Vertenstein, Z. L. Yang, and X. Zeng (2004), Technical description of the community land model (CLM), Report NCAR/TN-461+STR, National Center for Atmospheric Research, Boulder, CO.

Ottlé, C., and D. Vidal-Madjar (1994), Assimilation of soil moisture inferred from infrared remote sensing in a hydrological model over the HAPEX-MOBILHY region, *Journal of Hydrology*, *158*, 241–264.

Pachepsky, Y. A., D. Timlin, and G. Várallyay (1996), Artificial neural networks to estimate soil water retention from easily measurable data, *Soil Science Society of America Journal*, *60*, 727–773.

Palmer, W. C. (1965), Meteorological drought, Research Paper No. 45, Weather Bureau, Washington, D.C.

Palmer, W. C. (1968), Keeping track of crop moisture conditions, nationwide: The new crop moisture index, *Weatherwise, 21*(4), 156–161.

Panofsky, H. A., and G. W. Brier (1968), Some applications of statistics to meteorology, Pennsylvania State University Press, University Park, PA.

Peck, E. L. (1976), Catchment modeling and initial parameter estimation for the national weather service river forecast system (NOAA Technology Memorandrom; NWS HYDRO-31), National Weather Service, Silver Spring, MD.

Pollacco, J. A., B. P. Mohanty, and A. Efstratiadis (2012), The weighted objective function selector algorithm and its application to invert surface soil moisture and evapotranspiration in SVAT models, *Advances in Water Resources*, Accepted.

Rabiner, L. R. (1989), A tutorial on hidden Markov models and selected applications in speech recognition, *Proceeding of IEEE*, *77*(2), 257–286.

Robertson, A., S. Kirshner, and P. Smyth (2004), Downscaling of daily rainfall occurrence over northeast brazil using a hidden markov model, *Journal of Climate, 17*, 4407–4424.

Roeckner, E., K. Arpe, L. Bengtsson, C. M. Claussen, L. Dümenil, M. Esch, M. Giorgetta, U. Schiese, and U. Schulzweida (1996), The atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate, Report 218, Max-Planck-Institute for Meteorology, Hamburg, Germany.

Roerink, G. J., Z. Su, and M. Menenti (2000), S-SEBI: A simple remote sensing algorithm to estimate the surface energy balance, *Physics and Chemistry of the Earth*, *25*(2), 147–157.

Rose, C. W., W. R. Stern, and J. E. Drummond (1965), Determination of hydraulic conductivity as a function of depth and water content for soil in situ, *Australian Journal of Soil Research*, *3*, 1–9.

Sarwar, A., W. G. M. Bastiaanssen, M. T. Boers, and J. C. van Dam (2000), Evaluating drainage design parameters for the fourth drainage project, Pakistan by using SWAP model: I. Calibration, *Irrigation and Drainage Systems*, *14*(4), 257–280, doi:10.1023/A:1006468905194.

Schaake, J. C., V. I. Koren, Q. Y. Duan, K. Mitchell, and F. Chen (1996), Simple water balance model for estimating runoff at different spatial and temporal scales, *Journal of Geophysical Research*, *101*(D3), 7461–7476.

Schaap, M. G., and W. Bouten (1996), Modeling water retention curves of sandy soils using neural networks, *Water Resources Research*, *32*(10), 3033–3040, doi:10.1029/96WR02278.

Schaap, M. G., and F. J. Leij (1998), Database-related accuracy and uncertainty of pedotransfer functions, *Soil Science, 163*(10), 765–779, doi:10.1097/00010694-199810000-00001.

Schaap, M. G., F. J. Leij, and M. T. van Genuchten (1998), Neural network analysis for hierarchical prediction of soil hydraulic properties, *Soil Science Society of America Journal, 62*, 847–855.

Schaap, M. G., F. J. Leij, and M. T. van Genuchten (1999), A bootstrap-neural network approach to predict soil hydraulic parameters, in *Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media*, edited by M. T. van Genuchten and F. J. Leij, pp. 1237–1250, Proceeding of International Workshop, University of California, Riverside, CA.

Shafer, B. A., and L. E. Dezman (1982), Development of surface water supply index (SWSI) to assess the severity of drought conditions in snow pack runoff areas, paper presented at 50th Western Snow Conference, Colorado State University Press, Fort Collins, CO.

Sharma, S., B. P. Mohanty, and J. Zhu (2006), Including topography and vegetation attributes for developing pedotransfer functions, *Soil Science Society of America Journal, 70*(5), 1430–1440, doi:10.2136/sssaj2005.0087.

Shen, X., M. Kimoto, and A. Sumi (1998), Role of land surface processes associated with interannual variability of broad-scale Asian summer monsoon, *Journal of the Meteorological Society of Japan, 76*, 217–236.

Shin, Y., and B.P. Mohanty (2012), Development of drought severity assessment framework using remotely sensed soil moisture products with climate changes, *Water Resources Research,* Submitted.

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 Resources Research, 48*, W06529, doi:10.1029/2010WR009581.

Singh, R., R. K. Jhorar, J. C. van Dam, and R. A. Feddes (2006a), Distributed ecohydrological modeling to evaluate irrigation system performance in Sirsa district, India: II. Impact of viable water management scenarios, *Journal of Hydrology*, *329*, 714–723, doi:10.1016/j.jhydrol.2006.03.016.

Singh, R., J. G. Kroes, J. C. van Dam, and R. A. Feddes (2006b), Distributed ecohydrological modeling to evaluate the performance of irrigation system in Sirsa district, India: I. Current water management and its productivity, *Journal of Hydrology*, *329*, 692–713, doi:10.1016/j.jhydrol.2006.03.037.

Srinivasan, G., M. Hulme, and C. G. Jones (1995), An evaluation of the spatial and interannual variability of tropical precipitation as simulated by GCMs, *Geophysical Research Letters*, *22*(13), 1697–1700.

Stamm, J. F., and E. F. Wood (1994), Sensitivity of a GCM simulation of global climate to the representation of land-surface hydrology, *Journal of Climate*, *7*(8), 1218–1239.

Stern, R. D., and R. Coe (1984), A model fitting analysis of daily rainfall data, *Journal of Royal Statistical Society*, *147*(1), 1–34.

Stockton, J. G., and A. W. Warrick (1971), Spatial variability of unsaturated hydraulic conductivity, *Soil Science Society of America Journal*, *35*(5), 847–848.

Su, Z. (2002), The surface energy balance system (SEBS) for estimation of turbulent heat fluxes, *Hydrology and Earth System Sciences*, *6*, 85–100.

Supit, I., A. A. Hooyer, and C. A. van Diepen (1994), System description of the WOFOST 6.0 crop simulation model implemented in CGMS, in *Theory and*

Algorithms, vol. 1, edited by I. Supit et al., EUR 15956, pp. 146, Commission of the European Communities, Luxembourg.

Ulaby, F., P. Dubois, and J. von Zyl (1996), Radar mapping of surface soil moisture, *Journal of Hydrology*, *184*, 57–84.

van Bavel, C. H. M., G. B. Stirk, and K. J. Brust (1968), Hydraulic properties of a clay loam soil and the field measurement of water uptake by roots: I. Interpretation of water content and pressure profiles, *Soil Science Society of America Journal*, *32*, 310–317.

van Dam, J. C. (2000), Field-scale water flow and solute transport: SWAP model concepts, parameter estimation, and case studies, Ph.D. dissertation, Department of Soil Physics, Agricultural Hydrology and Groundwater, Wageningen University, Wageningen, Netherlands.

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 soil-water-atmosphere-plant environment, Technical Document 45, DLO Winand Staring Centre, Wageningen, Netherlands.

van Genuchten, M. T. (1980), A closed-form equation foe predicting the hydraulic conductivity of unsaturated soils, *Soil Science Society of America Journal*, *44*(5), 892–898.

Viterbi, A. J. (1967), Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, *IEEE Transactions on Information Theory*, *13*(2), 260–269.

Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resources Research, 39*(8), 1201, doi:10.1029/2002WR001642.

Vrugt, J. A., G. Schoups, J. W. Hopmans, C. Young, W. W. Wallender, and W. Bouten (2004), Inverse modeling of large-scale spatially distributed vadose zone properties using global optimization, *Water Resources Research, 40*, W06503, doi:10.1029/2003WR00276.

Walker, J., and P. R. Rowntree (1977), The effect of soil moisture on circulation and rainfall in a tropical model, *Quarterly Journal of Royal Meteorological Society*, *103*(435), 29–46.

Wang, Q. J. (1991), The Genetic Algorithm and its application to calibrating conceptual rainfall-runoff models, *Water Resources Research, 27*(9), 2467–2471.

Wang, X., H. Xie, H. Guan, and X. Zhou (2007), Different responses of MODIS-derived NDVI to root-zone soil moisture in semi-arid and humid regions, *Journal of Hydrology*, *340*, 12–24, doi:10.1016/j.jhydrol.2007.03.022.

Wesseling, J. G., and J. G. Kroes (1998), A global sensitivity analysis of the model SWAP, Report 160, DLO Winand Staring Centre, Wageningen, Netherlands.

Wilby, R. L., and T. M. L. Wigley (1997), Downscaling general circulation model output: A review of methods and limitations, *Progress in Physical Geography*, *21*(4), 530–548.

Wilby, R. L., T. M. L. Wigley, D. Conway, P. D. Jones, B. C. Hewistson, J. Main, and D. S. Wilks (1998), Statistical downscaling of general circulation model output: A comparison of methods, *Water Resources Research*, *34*(11), 2995–3008.

Wilks, D. S. (1998a), Conditioning stochastic daily precipitation models on total monthly precipitation, *Water Resources Research*, *25*(6), 1429–1439.

Wilks, D. S. (1998b), Multisite generalization of a daily stochastic precipitation generation model, *Journal of Hydrology*, *210*, 178–191.

Wilks, D. S., and R. L. Wilby (1999), The weather generation game: A review of stochastic weather models, *Progress in Physical Geography, 23*(3), 329–357.

Wood, E. F. (1994), Scaling soil moisture and evapotranspiration in runoff models, *Advances in Water Resources*, *17*, 25–34.

Wood, A. W., E. P. Maurer, A. Kumar, and D. P. Lettenmaier (2002), Long-range experimental hydrologic forecasting for the eastern United States, *Journal of Geophysical Research, 107*(D20), 4429, doi:10.1029/2001JD000659.

Woolhiser, d. A., and J. Roldan (1982), Stochastic daily precipitation models: 2. A comparison of distribution of amounts, *Water Resources Research*, *18*(5), 1461–1468.

Wösten, J. H. M., G. H. Veerman, and J. Stolte (1994), Water retention and hydraulic conductivity functions of top- and sub-soils in the Netherlands: The Staring

series, Technical Document 18, DLO Winand Staring Centre, Wageningen, Netherlands.

Wösten, J. H. M., A. Lilly, A. Nemes, and C. Le Bas (1999), Development and use of a database of hydraulic properties of European soils, *Geoderma, 90,* 169–185.

Wu, J., C. Zheng, C. Cien, and L. Zheng (2006), A comparative study of monte carlo simple genetic algorithm and noisy genetic algorithm for cost-effective sampling network design under uncertainty, *Advances in Water Resources*, *29*, 899–911, doi:10.1016/j.advwatres.2005.08.005.

Xevi, E., J. Gilley, and J. Jeyen (1996), Comparative study of two crop yield simulation models, *Agricultural Water Management*, *30*(2), 155–173.

Yang, S., and K. M. Lau (1998), Influences of sea surface temperature and ground wetness on Asian summer monsoon, *Journal of climate*, *11,* 3230–3246.

Yeh, T. C. J. (1989), One-dimensional steady state infiltration in heterogeneous soils, *Water Resources Research, 25*, 2149–2158.

Young, K. C. (1994), A multivariate chain model for simulating climatic parameters from daily data, *Journal of Applied Meteorology and Climatology*, *33*, 661–671.

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, *Journal of Hydrometeorology*, *10*, 308–319, doi:10.1175/2008JHM1011.1.

Zhu, J., and B. P. Mohanty (2002), Upscaling of soil hydraulic properties for steady state evaporation and infiltration, *Water Resources Research*, *38*(9), 1178. doi:10.1029 /2001WR000704.

Zhu, J., and B. P. Mohanty (2003), Upscaling of hydraulic properties of heterogeneous soils, in *Scaling Methods in Soil Physics*, edited by Y. A. Selim et al., pp. 97–117, CRC Press, Boca Raton, FL.

Zhu, J., and B.P. Mohanty (2004), Soil hydraulic parameter upscaling for steady-state flow with root water uptake, *Vadose Zone Journal*, *3*(4), 1464–1470.

Zhu, J., and B. P. Mohanty (2006), Effective scaling factors for transient infiltration in heterogeneous soils, *Journal of Hydrology*, *319*, 96–108, doi:10.1016/j.jhydrol. 2005.07.004.

Zhu, J., B. P. Mohanty, and N. N. Das (2006), On the effective averaging schemes of hydraulic properties at the landscape scale, *Vadose zone Journal*, *5*(1), 308–316.

Zhu, J., B. P. Mohanty, A. W. Warrick, and M. T. van Genuchten (2004), Correspondence and upscaling of hydraulic functions for steady-state flow in heterogeneous soils, *Vadose Zone Journal*, *3*, 527–533.

APPENDIX A

1. Noah Land Surface Model (Noah LSM)

Noah Land Surface Model (Noah LSM) is widely used to simulate water and energy fluxes based on coupled and uncoupled mode at various scales [*Ek et al*., 2003]. In this study, we used the uncoupled mode for estimating soil moisture dynamics in the vertical direction.

This model simulates soil moisture estimates using the diffusive form of the Richards equation (Eq. (A.1)) based the Campbell [*Campbell*, 1974] soil hydraulic properties (SHP) model in Eqs. (A.2 and A.3), the simple water balance (SWB) model [*Schaake et al*., 1996] for surface runoff, and internal parameterization scheme [*Clapp and Hornberger*, 1978] for parameterization.

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left(D\frac{\partial \theta}{\partial z}\right) + \frac{\partial K(\theta)}{\partial z} + S(\theta) \tag{A.1}$$

$$\psi = \psi_{sat}\left[\frac{\theta}{\theta_{sat}}\right]^{-b} \tag{A.2}$$

$$k = k_{sat}\left[\frac{\theta}{\theta_{sat}}\right]^{2b+3} \tag{A.3}$$

where ψ is the soil matric potential, ψ_{sat} is the saturated soil matric potential, b is the curve fitting parameter, θ is the soil moisture, θ_{sat} is the saturated soil moisture, k is the hydraulic conductivity, and k_{sat} is the saturated hydraulic conductivity in the soil profile, respectively.

2. Community Land Model (CLM)

Community Land Model (CLM) combined with Land surface model (LSM) [*Bonan*, 1996], biosphere-atmosphere transfer scheme (BATS) [*Dickinson et al*., 1993], and the Chinese Academy of Sciences Institute of Atmospheric Physics's LSM, 1994 version [*Dai and Zeng*, 1997] provides land surface forcing as the physical boundary for atmospheric model in Community Climate System Model (CCSM). The CLM model requires input datasets such as land surface type, soil and vegetation parameters, model initialization, and atmospheric boundary conditions [*Dai et al*., 2003].

This model requires the forcing data (i.e., solar radiation, downward wave radiation, air temperature, wind speed, air pressure, humidity, and precipitation). The CLM also adapted simple parameterizations [*Clapp and Hornberger*, 1978] for identifying soil and vegetation parameters.

The CLM simulates soil moisture dynamics in the vertical direction based on a modified Richards' equation (Eq. (A.4)). Hydrostatic equilibrium soil moisture distribution was subtracted to improve the mass-conservative numerical scheme when the water table is within the soil column [*Zeng and Decker*, 2009].

$$\frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{t}} = \frac{\partial}{\partial \boldsymbol{z}}\left[\boldsymbol{k}\left(\frac{\partial(\boldsymbol{\Psi} - \boldsymbol{\Psi}_{\boldsymbol{E}})}{\partial \boldsymbol{z}}\right)\right] - \boldsymbol{Q} \qquad \text{(A.4)}$$

where θ is the volumetric soil water content ($mm^3\ mm^{-3}$), k is the hydraulic conductivity ($mm\ s^{-1}$), ψ is the soil matric potential (mm), ψ_E is the equilibrium soil matric potential (mm), and Q is the soil moisture sink term.

The hydraulic conductivity and equilibrium soil matric potential are given as equations (Eqs. (A.5-A.7)), respectively.

$$\boldsymbol{k}[\boldsymbol{z}_{\boldsymbol{h},\boldsymbol{i}}]=\begin{cases}(1-\dfrac{\boldsymbol{f}_{\boldsymbol{frz},\boldsymbol{i}}+\boldsymbol{f}_{\boldsymbol{frz},\boldsymbol{i}+1}}{2})\boldsymbol{k}_{\boldsymbol{sat}}[\boldsymbol{z}_{\boldsymbol{h},\boldsymbol{i}}][\dfrac{0.5(\boldsymbol{\theta}_{\boldsymbol{i}}+\boldsymbol{\theta}_{\boldsymbol{i}+1})}{0.5(\boldsymbol{\theta}_{\boldsymbol{sat},\boldsymbol{i}}+\boldsymbol{\theta}_{\boldsymbol{sat},\boldsymbol{i}+1})}]^{2\boldsymbol{B}_{\boldsymbol{i}}+3} & 1\le \boldsymbol{i}\le \boldsymbol{N}_{\boldsymbol{levsoi}}-1 \\ (1-\boldsymbol{f}_{\boldsymbol{frz},\boldsymbol{i}})\boldsymbol{k}_{\boldsymbol{sat}}[\boldsymbol{z}_{\boldsymbol{h},\boldsymbol{i}}]\left(\dfrac{\boldsymbol{\theta}_{\boldsymbol{i}}}{\boldsymbol{\theta}_{\boldsymbol{sat},\boldsymbol{j}}}\right)^{2\boldsymbol{B}_{\boldsymbol{i}+3}} & \boldsymbol{i}=\boldsymbol{N}_{\boldsymbol{levsoi}}\end{cases} \tag{A.5}$$

$$\Psi_{\boldsymbol{E}}=\Psi_{\boldsymbol{sat}}\left[\frac{\boldsymbol{\theta}_{\boldsymbol{E}}(\boldsymbol{z})}{\boldsymbol{\theta}_{\boldsymbol{sat}}}\right]^{-\boldsymbol{B}} \tag{A.6}$$

$$\theta_E(z)=\theta_{sat}\left(\frac{\psi_{sat}+z_{\nabla}-z}{\psi_{sat}}\right)^{-\frac{1}{B}} \tag{A.7}$$

where $f_{frz,i}$ is the impermeable fraction, $z_{h,i}$ is the depth of the interface of two adjacent layers, N_{levsoi} is the number of soil layer, B is a function of soil texture, θ_E is the equilibrium volumetric water content (mm^3 mm^{-3}) at depth z, θ_{sat} is the saturated volumetric water content (mm^3 mm^{-3}), and z_{∇} is the water table depth, respectively.

APPENDIX B

1. Expectation-Maximization Algorithm for Hidden Markov Model (HMM)

In this study, HMM trained the parameter set ($\Theta=\{A, B, \pi_i\}$) through the optimized initial parameters by a genetic algorithm (GA). HMM trains parameters through a three-step process as follows,

i) State transition probabilities $A=|a_{i,j}|$

The transition probabilities ($a_{i,j}$) at time t determine states at time t+1. a_j is the probabilities of transitioning state S_i at time t to state S_j at time t+1. K is the orders of hidden states.

$$\left.\begin{array}{c} a_{i,j} = P(S_{t+1} = s_j \mid S_t = s_i), 1 \le i, j \le K \\ \sum_{j=1}^{n} a_{i,j} = 1 \end{array}\right\} \quad \text{(B.1)}$$

ii) Observation probabilities $B=|b_j(o_j)|$ in state j, where

$$\left.\begin{array}{c} b_j(o_k) = P(O_t = o_k \mid S_t = s_j), 1 \le j \le K, 1 \le k \le 2 \\ \sum_{k=1}^{2} b_j(o_k) = 1 \end{array}\right\} \quad \text{(B.2)}$$

iii) Initial transition probabilities $\pi = |\pi_i|$, where

$$\left.\begin{array}{c} \pi_i = P(S_1 = s_i), 1 \le i \le K \\ \sum_{i=1}^{n} \pi_i \end{array}\right\} \quad \text{(B.3)}$$

We used the *Forward-Backward* algorithm (FBA) to compute $P(O|\Theta)$ with the parameter sets (Θ) of HMM. The FBA is to estimate the posterior marginals of hidden state variables with a sequential observations ($O_{1:T}=O_1,\ldots,O_T$). The forward variable $\alpha_t(i)$ can be solved as follows,

$$\alpha_1(i) = \pi_i b_i(O_1), 1 \le i \le K \quad \text{(B.4)}$$

$$\alpha_t(i) = P(O_1,\ldots,O_t, S_t = s_i \mid \Theta) = [\sum_{j=1}^{K} \alpha_{t-1}(j) a_{i,j}] b_i(O_t) \tag{B.5}$$

$$P(O \mid \Theta) = \sum_{j=1}^{K} \alpha_T(i) \tag{B.6}$$

Then, the algorithm computes a set of backward variable $\beta_t(i)$ defined as below,

$$\beta_T(i) = 1, 1 \le i \le K \tag{B.7}$$

$$\beta_t(i) = P(O_{t+1},\ldots,O_T \mid Q_t = S_i \mid \Theta) = [\sum_{j=1}^{K} a_{i,j} b_j(O_{t+1}) \beta_{t+1}(j)] \tag{B.8}$$

Then, P(O|Θ) is maximized using the given parameter set (Θ={*A*, *B*, π_i}) through an iterative procedure using the Baum-Welch method (called as the expectation-maximization algorithm, *Dempster et al*., [1977]). For the procedure of re-estimation of HMM parameter set (Θ), we first define $\kappa_t(i,j)$ indicating the probability of transitioning state S_i at time t to state S_j at time t+1 as follows,

$$\kappa_t(i,j) = \frac{\alpha_t(i) a_{i,j} b_j(O_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^{K}\sum_{j=1}^{K} \alpha_t(i) a_{i,j} b_j(O_{t+1}) \beta_{t+1}(j)}, 1 \le t \le T, 1 \le i \le K, 1 \le j \le K \tag{B.9}$$

We defined $\gamma_t(i)$ as the probability of being in state S_i at time t with the observation sequence and the model. Thus, $\gamma_t(i)$ is related to $\kappa_t(i,j)$ by summing over *j*,

$$\gamma_t(i) = P(S_t = s_i \mid O, \Theta) = \sum_{j=1}^{K} \kappa_t(i,j), 1 \le t \le T, 1 \le i \le K \tag{B.10}$$

$$\alpha_{i,j}^{new} = \frac{\sum_{t=1}^{T-1} \kappa_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, 1 \le i, j \le K \tag{B.11}$$

$$b_i(V_k)^{new} = \frac{\sum_{t=1}^{T} \gamma_t(i)}{\sum_{t=1}^{T} \gamma_t(i)}, 1 \le i \le K, 1 \le k \le 2 \tag{B.12}$$