SOIL MOISTURE DATA ASSIMILATION AT MULTIPLE SCALES AND ESTIMATION OF REPRESENTATIVE FIELD SCALE SOIL MOISTURE CHARACTERISTICS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Eunjin Han

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2011

Purdue University

West Lafayette, Indiana

To my parents and family

ACKNOWLEDGEMENTS

I would like to thank my advisors Dr. Venkatesh Merwade and Dr. Gary Heathman for their immeasurable guidance and encouragement during my graduate studies. I extend many thanks to my thesis committee members, Dr. Rao Govindaraju and Dr. Laura Bowling, who have provided insights and helpful comments for improving the quality of my research. I also wish to thank the Agricultural Research Service-National Soil Erosion Research Laboratory for their financial support for me through my graduate studies. I express special thanks to Scott McAfee and Jim Frankenberger at the National Soil Erosion Research Laboratory for their technical support and data provision for my research. I would like to thank all my colleagues and friends, especially in Hydraulics and Hydrology group of Civil Engineering, the National Soil Erosion Research Laboratory and Korean Presbyterian Church at Purdue for their encouragement. They have made my graduate life richer and more meaningful. I also thank Mrs. Judith Han (area secretary) for her support. Finally, I would not have been able to complete my graduate studies without the support and encouragement of my family.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

ABBREVIATIONS

SYMBOL	UNIT	DESCRIPTION
ALPHA_BF	days	Alpha factor for groundwater recession curve
RCHRG_DP		Deep aquifer percolation fraction
REVAPMN	mm	Threshold water depth in the shallow aquifer for "revap"
CANMX		Maximum canopy storage
CN2		Initial SCS CN II value
SOL_K	mm hr^{-1}	Saturated hydraulic conductivity
SOL_Z	mm	Depth to bottom of soil layer
SOL_AWC	mm H_2O $mmsoil^{-1}$	Available water capacity of soil layer
SOL_ALB		Albedo when soil is moist
SLOPE	m m^{-1}	Average slope steepness
ESCO		Soil evaporation compensation factor
EPCO		Plant water uptake compensation factor
SURLAG	day	Surface runoff lag time
SMTMP	◦C	Snow melt base temperature
TIMP		Snow pack temperature lag factor
CH_N2		Manning's "n" value for the main channel
CH_K2	mm hr^{-1}	Effective hydraulic conductivity in main channel alluvium
BLAI		Maximum (potential) leaf area index
SOL_SW	mm H_2O	Amount of water stored in the soil profile on any given day
SHALLST	mm H_2O	Depth of water in shallow aquifer
DEEPST	mm H_2O	Depth of water in deep aquifer
INFLPCP	mm H_2O	Amount of precipitation that infiltrates into soil
QDAY	mm H_2O	Surface runoff loading to main channel for day in HRU
LATQ	mm H_2O	Amount of water in lateral flow in HRU for the day

SYMBOL	UNIT	DESCRIPTION
RCHRG	mm H_2O	Amount of water recharging both aquifers on current day in HRU
GW_Q	mm H_2O	Groundwater contribution to streamflow from HRU on current day
GWSEEP	mm H_2O	Amount of water recharging deep aquifer on current day
CNDAY		Curve number for current day, HRU and at current soil moisture
ET	mm H_2O	Actual amount of evapotranspiration that occurs on day in HRU

ABSTRACT

Han, Eunjin. Ph.D., Purdue University, December 2011. Soil Moisture Data Assimilation at Multiple Scales and Estimation of Representative Field Scale Soil Moisture Characteristics. Major Professors: Venkatesh Merwade and Gary Heathman.

Soil moisture is a key variable in understanding the hydrologic processes and energy fluxes at the land surface. Therefore, accurate prediction of soil moisture in the vadose zone benefits irrigation planning and crop management, flooding and drought prediction, water quality management, climate change forecasts, and weather prediction. The three objectives of this study are to: (1) investigate the effects of surface soil moisture data assimilation on hydrological responses at the field scale using *in situ* soil moisture measurements, (2) explore the effect of surface soil moisture data assimilation on each hydrologic process in a simulation model, including the effect of spatially varying inputs on the potential capability of surface soil moisture assimilation at the watershed scale, and (3) link two different scales of soil moisture estimates by upscaling single point measurements to field averages.

First, a well-proven data assimilation technique, the Ensemble Kalman Filter (EnKF), is applied to a field scale water quality model, the Root Zone Water Quality Model, with

in situ soil moisture data from two agricultural fields in Indiana. Through daily update, the EnKF improves all statistical results compared to the direct insertion method and model results without assimilation for the 5 cm and 20 cm depths while less improvement is achieved for deeper layers. Optimal update interval and ensemble size are also tested for the operational potential of data assimilation. This study demonstrates the potential of surface soil moisture assimilation to improve water quality and crop yield simulation, as well as, soil moisture estimation at the agricultural field scale.

Second, the EnKF is coupled with a watershed scale, semi-distributed hydrologic model, the Soil and Water Assessment Tool. Results show that daily assimilation of surface soil moisture with the EnKF improves model predictions of almost all hydrological processes. However, the EnKF does not produce as much of a significant improvement in streamflow predictions as compared to soil moisture estimates in the presence of large precipitation errors and due to the limitations of the infiltration-runoff model mechanism. Distributed errors of the soil water content show effects of spatially varying inputs such as soil and landuse types on the assimilation results. Results from this study suggest that soil moisture update through data assimilation can be a supplementary way to overcome the errors created by limited or inaccurate rainfall data.

Proper linkage of soil moisture estimates across different scales of observations and model predictions is essential for the validation of remotely-sensed soil moisture products, as well as the successful application of data assimilation techniques. Thus, this study also examines different upscaling methods to transform point measurements to field averages in representing small agricultural watersheds (~ 2 ha). The cumulative distribution

function (CDF) matching approach is found to provide best estimates of field average soil moisture out of several statistical methods. Tests for temporal and spatial (horizontal and vertical) transferability of the upscaling equations indicate that they are transferable in space, but not in time. Rainfall characteristics and crop types are most likely major factors affecting the success of the transferability. In addition, the CDF matching approach is found to be an effective method to estimate spatial soil moisture variance from single point measurements.

Overall, the results presented in this work can be utilized to improve applications of soil moisture data assimilation at field and watershed scales and better evaluate the scaling behavior of soil moisture.

CHAPTER 1 INTRODUCTION

1.1 Background and Motivation

Soil moisture is an important hydrological state variable that controls the evolution of various hydrological and energy balance processes at the land surface (Pielke et al., 2010). Therefore, there have been considerable efforts to estimate soil moisture in the vadose zone in different research areas for: 1) weather forecast, 2) climate change studies, 3) flood or drought prediction, 4) water quality management, 5) irrigation operation, and 6) soil erosion studies.

Data assimilation is a technique that optimally integrates observational information with model predictions for improved forecast of state variables. Applications of soil moisture data assimilation in hydrology have a relatively short history; these studies were initiated in the early 1980s and have made rapid progress in the last decade with the development of remote sensing technologies and increasingly available remotely-sensed soil moisture data (Ni-Meister, 2008). Shallow soil moisture estimates from remotely-sensed data have led researchers to merge measured surface soil moisture data and hydrological models in an effort to obtain more accurate soil moisture estimates in the vadose zone through data assimilation (Reichle et al., 2002b; Ni-Meister et al., 2006a; Reichle et al., 2007; Sabater et al., 2007; Das et al., 2008; Draper et al., 2009).

Since these soil moisture assimilation studies were motivated by how to best use remotely-sensed soil moisture data obtained from satellites to improve climate forecasts (McLaughlin, 1995), they have been conducted at regional or global scales using land surface models. Thus, there is still a lack of research on utilizing soil moisture data assimilation techniques with hydrologic models at field or watershed scales which are the fundamental operative unit for water resources management (Troch et al., 2003). Most operational hydrologic models for agricultural applications, flood forecast and water quality management are applied at field or watershed scales. Therefore, successful application of soil moisture data assimilation at the field or watershed scale would be a significant contribution for improving the prediction capability of hydrologic models with increased operational benefits.

One of the obstacles in the application of soil moisture data assimilation at the field or watershed scale is an intermediate gap in the scale of soil moisture measurements (Robinson et al., 2008). Two major soil moisture observational methods have advanced at significantly different scales; *in situ* sensors typically provide point soil moisture measurements in an area less than 1 m^2, while remotely sensing data, especially satellite-based, may be on the order of a few kilometers. Since successful improvement in soil moisture estimation through data assimilation is possible only with good quality and unbiased soil moisture observations (Ni-Meister, 2008), it is important to resolve the scale issues between different soil moisture observations and areal estimates of field or watershed scale hydrologic models.

Ideally the scales between measurement, modeling and processing should be interchangeable or identical (Verstraeten et al., 2008). However, in reality there can be significant differences between scales which pose problems for full utilization of available soil moisture measurements and modeling capabilities (Western et al., 2002; Robinson et al., 2008). In addition, proper linkage of soil moisture estimates across different scales of observations and model predictions is essential for the validation of current and upcoming space-borne soil moisture retrievals (Grayson and Western, 1998; Crow et al., 2005; Cosh et al., 2006; Miralles et al., 2010).

1.2 Research Objectives

This study aims to extend the applicability of soil moisture data assimilation for local or watershed scale hydrologic simulations based on the progress made using land surface models that have been conducted primarily at the regional or global scale. This study is also expected to contribute to other important agricultural, environmental and hydrologic research issues at local or watershed scales. It is hypothesized in this study that assimilating surface soil moisture observations can improve predictions of total soil water content in the root zone and thus, enhance overall model performance of the hydrological system at the field or watershed scale. This hypothesis is tested at both field and watershed scales with detailed objectives described below.

Another goal of this study is to tackle a scaling issue in soil moisture estimation by upscaling single point soil moisture measurements to field averages that are representative of a small agricultural watershed (~ 2 ha). The hypothesis of this approach

is that representative soil moisture characteristics such as the field average and spatial variance can be predicted using point soil moisture measurements from a single location. Each objective is accomplished as a separate study resulting in a journal article. A brief description of each paper is given below.

(1) **Application of soil moisture data assimilation with a field scale hydrologic model:** A well-proven data assimilation technique, the Ensemble Kalman Filter (EnKF), is applied to a field scale water quality model, the Root Zone Water Quality Model (RZWQM), with *in situ* soil moisture data from two agricultural fields in Indiana. Specific tasks of this objective include: (i) to compare the capability of the EnKF with simple direct insertion method, and (ii) to explore the effect of ensemble size and update interval for practical application of the EnKF.

(2) **Application of soil moisture data assimilation with a watershed scale hydrologic model:** The EnKF is coupled with a watershed scale, semi-distributed hydrologic model, the Soil and Water Assessment Tool (SWAT), in order to investigate the effect of surface soil moisture assimilation on hydrological responses at the watershed scale through a synthetic experiment. This objective is achieved with specific tasks of: (i) investigating how soil moisture update through assimilation may compensate for errors in the hydrologic prediction due to inaccurate rainfall input, and (ii) exploring how spatially varying inputs such as land use and soil types affect the potential capability of surface soil moisture assimilation.

(3) **Estimation of representative field scale soil moisture characteristics**: The main objective for this phase in the study is to find the most appropriate means to

transform single point measurements to field averages by applying different upscaling methods. Additional tasks of this objective involve: (i) exploring whether the best upscaling method is transferable in time and space, (ii) finding which factors most influence the temporal and spatial transferability of the upscaling method, and (iii) estimating spatial variability of soil moisture within the fields using single point measurements and the cumulative distribution function matching method.

1.3 Dissertation outline

The remainder of this dissertation is organized as follows. Chapters 2 to 4 address the three major research objectives mentioned earlier. Each chapter is considered an independent study having its own introduction to conclusions, but they are all connected under the umbrella of soil moisture data assimilation and scaling study at local scales. Final conclusions, future works and remarks synthesizing the all three studies are included in Chapter 5.

CHAPTER 2 SOIL MOISTURE DATA ASSIMILATION AT FIELD SCALE

2.1 Introduction

Estimation of soil moisture has received considerable attention in the areas of hydrology, agriculture, meteorology and environmental studies because of its role in the partitioning of water and energy at the land surface, specifically, precipitation into runoff and infiltration, and energy into latent and sensible heat fluxes. Better understanding and prediction of soil moisture in the root zone is beneficial in irrigation planning and crop management, flooding and drought prediction, water quality management, climate change and weather prediction. However, despite its significance, the expenses associated with field measurements and the high spatiotemporal variability of soil properties have placed limitations in obtaining *in situ* measurements of profile soil moisture. Recent developments in remote sensing technology have provided possibilities to overcome these limitations through soil moisture retrieval from remote sensing images (Schmugge and Jackson, 1994; Jackson et al., 1995; Jackson and Vine, 1996; Jackson et al., 2002; Njoku et al., 2003; Verstraeten et al., 2006; Baup et al., 2007). However, in most cases, the actual sensing depth is limited to a few centimeters allowing soil moisture estimates only for the top 0-5 cm and 1 cm surface layer using the L-band (~1-2 GHz)and C-band (~4-8 GHz), respectively (Jackson and Schmugge, 1989; Njoku et al., 2003).

The limitations of shallow soil moisture estimates from remotely-sensed data have led researchers to integrate measured surface data and hydrologic models to obtain more accurate estimates of soil moisture content in the root zone through data assimilation techniques (Houser et al., 1998; Li and Islam, 1999; Hoeben and Troch, 2000; Walker et al., 2001b; Heathman et al., 2003a). Various data assimilation techniques have been used in soil moisture studies including the Kalman Filter (Hoeben and Troch, 2000; Walker et al., 2001b; Crosson et al., 2002), direct insertion method (Walker et al., 2001b; Heathman et al., 2003a), extended Kalman Filter (Reichle et al., 2002b; Seuffert et al., 2004; Draper et al., 2009) and Ensemble Kalman Filter (Reichle et al., 2002a, 2002b; Zhang et al., 2006). Among these techniques, the Ensemble Kalman Filter (EnKF) has been widely adopted because of its strength in handling non-linear systems and computational efficiency (Crow and Wood, 2003; Huang et al., 2008). As one of surface measurements from a remote platform, brightness temperature has been assimilated into hydrologic models for better estimation of soil moisture (Margulis et al., 2002; Crow and Wood, 2003; Huang et al., 2008). Recently, satellite-based surface soil moisture observations (0-5 cm) from the Scanning Channel Microwave Radiometer (SMMR) and the Advanced Microwave Scanning Radiometer – Earth Observing System (AMSR-E) have been integrated with land surface models through data assimilation (Reichle and Koster, 2005; Ni-Meister et al., 2006b; Reichle et al., 2007; Draper et al., 2009).

Most previous investigations that explored assimilation of remotely-sensed surface soil moisture involved regional or global scales for more accurate climate forecasts by improving temporal and spatial soil moisture estimations (Walker and Houser, 2001;

Reichle et al., 2002b; Ni-Meister, 2008; Draper et al., 2009). However, Troch et al. (2003) stressed the necessity of extending the applicability of data assimilation, which has so far been studied in land surface models using synthetic datasets, to many other important hydrological issues at catchment scale for a variety of water resources management problems. Therefore, there is a strong need for more application-focused research to determine the operational potentials of data assimilation within hydrologic modeling using actual observations.

There have been a few studies which have applied soil moisture data assimilation techniques for improved hydrologic predictions at field or catchment scales. Walker et al. (2002) demonstrated the feasibility of near-surface soil moisture assimilation to retrieve profile soil moisture in a 6 ha catchment using a three dimensional distributed soil moisture model. *In situ* soil moisture observations in an agricultural field (21 ha) were used for soil moisture assimilation with the EnKF, especially for correcting bias (De Lannoy et al., 2007a; De Lannoy et al., 2007c) and for applying an adaptive EnKF (De Lannoy et al., 2009), with the Community Land Model. Recently, a catchment scale hydrologic model, the Soil and Water Assessment Tool, was used to investigate the performance of the EnKF for hydrologic predictions (Xie and Zhang, 2010; Chen et al., 2011).

Heathman et al. (2003a) used a physically based and field-scale agricultural model, the Root Zone Water Quality Model (RZWQM) to explore the possibility of assimilating observed surface soil moisture for better estimates of root zone soil water content. Based on substantial field measured data, they showed that the direct insertion data assimilation

technique produced dynamics of model simulation results in the top 30 cm layers better than the model simulation without assimilation. Since they focused on using field measured surface soil moisture to estimate root zone soil water content rather than synthetic analyses, their study used the simple direct insertion method. As an extension to Heathman et al. (2003a), this paper examines the benefit of assimilating surface soil moisture through the EnKF to improve model simulated soil moisture in the root zone using the RZWQM. In addition, contrary to many previous studies which assimilated remotely-sensed surface soil moisture into land surface models at the regional or global scale, this study applies a point scale model (RZWQM) to field measured surface soil moisture data for the purpose of data assimilation.

Field scale application of the EnKF in this study provides the fundamental groundwork necessary to advance our knowledge of surface soil moisture data assimilation in hydrologic modeling. First, field scale, deterministic models such as RZWQM can serve as the basis for the watershed scale hydrologic or water quality models (Abrahamson et al., 2006) and has been directed to extend the modeling methodology from one dimension to two or three dimensions (Walker et al., 2001a; Ma et al., 2007a). Second, *in situ* soil moisture observations have fewer uncertainties than the remotely-sensed soil moisture data, which is better for testing different data assimilation schemes and validating the assimilation results more effectively. This is because observation errors for *in situ* soil moisture data can be estimated in a more reasonable way and thus, the *in situ* observations can be considered to represent the true system state with certain errors. Third, a field scale assimilation study can be based on more accurate soil characteristics

data which are essential for water flow simulation in the unsaturated zone possibly reducing systematic model bias before assimilation. Walker et al. (2001a) discussed that incorrect information about soil porosity or residual soil moisture content restrained the soil moisture model from correct retrieval of profile soil moisture, even with an assimilation scheme. Lastly, it is possible with the field scale studies to investigate the impacts of local horizontal or vertical soil heterogeneity on data assimilation results, which is generally ignored in large scale studies.

RZWQM is one of the most representative agricultural system models with components for water movement, plant growth, and chemical transport with management effects (Ma et al., 2007a). Since its release in 1992 (Ahuja et al., 2000), the RZWQM has been widely applied and tested in different areas for various purposes (Ma et al., 2001): to determine soil hydraulic properties (Cameira et al., 2000), to simulate tile drainage and leached nitrate (Abrahamson et al., 2005; 2006), and crop production and nitrogen balances (Ma et al., 2007c). Among major components of the RZWQM, soil hydrological components have performed satisfactorily (Cameira et al., 2000; 2005; Kozak et al., 2007) and soil water content were reasonably simulated (Farahani et al., 1999; Ma et al., 2003; 2007b; 2007c). In addition, the RZWQM has been successfully used to induce saturated hydraulic conductivities for spatially distributed soils by coupling with microwave remotely-sensed soil moisture (Mattikalli et al., 1998).

Thus, this study has the following three objectives: (i) apply EnKF to the RZWQM using point scale soil moisture data; (ii) compare the results from EnKF application with simple direct insertion method; and (iii) explore the effect of ensemble size and update interval

on model output. The above objectives were accomplished by using point scale profile soil moisture data collected for Upper Cedar Creek Watershed (UCCW) in northeast Indiana. Details of the study area including data acquisition are provided in the next section.

2.2 Study Area and Data

The Matson Ditch sub-catchment in the upper Cedar Creek Watershed located in northeastern Indiana (Figure 2.1) was selected as the test bed for this study. Since 2004, the National Soil Erosion Research Laboratory (NSERL) of the USDA – Agricultural Research Service (ARS) has established an extensive soil moisture and weather monitoring network in the upper Cedar Creek Watershed as shown in Figure 2.1. The weather stations in this network collect the following meteorological data for model simulations: ten minute rainfall, air temperature, solar radiation, wind speed, and relative humidity.

Among the ten soil monitoring sites within the upper Cedar Creek network, AS1 and AS2 field-size watersheds were selected for RZWQM simulations in this study because of the availability of measured soil properties data (e.g., soil texture, bulk density) and crop management information. The drainage area of the AS1 and AS2 field sites is 2.23 ha and 2.71 ha, respectively. Both sites are fairly flat (slope < 5%), and are in agriculture production (alternative cropping of corn and soybean) with AS1 being in no-till, and AS2 under a rotational tillage system. The major soil types found at these sites are Glynwood

(GnB2) silt loam (Fine, illitic, mesic Aquic Hapludalfs) for AS1 and Blount (BaB2) silt loam (Fine, illitic, mesic, Aeric Epiaqualfs) for AS2.

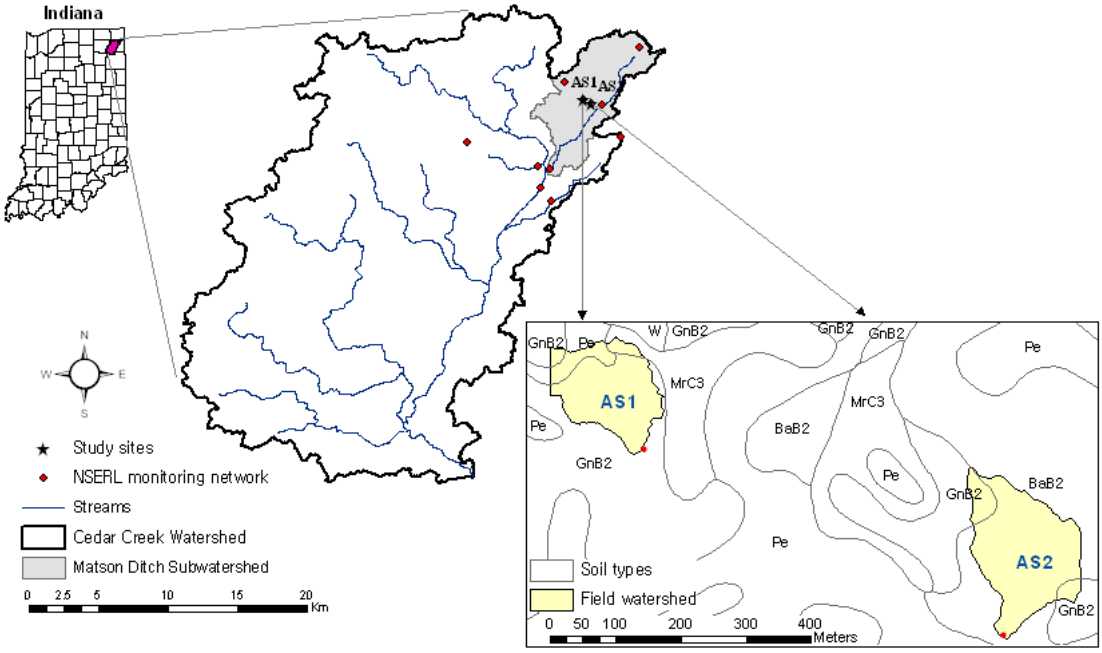


Figure 2.1 Study area in Upper Cedar Creek Watershed and AS1 and AS2 field sites

The soil sensors in both fields are Stevens SDI-12 Hydra Probes that measure soil moisture and temperature every ten minutes by propagating an electromagnetic signal into the soil, which is a Frequency Domain Reflectometry (vs Time Domain Reflectometry). Out of four factory calibration equations for soil moisture measurements, the linear calibration equation was used in this study. The calibration equation is based on the linear relationship between volumetric water content and the real component of the complex dielectric permittivity, which has been proven to be reasonably accurate for many soils (Topp et al., 1980; Topp and Davis, 1985; Heathman et al., 2003b; Seyfried and Murdock, 2004). In addition, Seyfried et al. (2005) concluded that linear expression

outperformed the other factory-supplied calibrations. According to texture analysis at the NSERL and Seyfried *et al* (2005), the linear calibration coefficients for loam soil type were selected. Since the Loam setting is applicable for Loam, Clay Loam, and Silty Clay Loam textures according to the Hydra Probe manual (Stevens Water Monitoring Systems, Inc., Portland, Oregon, USA, 2007), the selection of the calibration coefficients for loam is suitable for the soil types in this study (Table 2.1).

Table 2.1 Soil Physical and Hydraulic Properties

Site	Depth (cm)	Measured in laboratory							Model estimation	
		Soil[a)] Texture	Sand (%)	Silt (%)	Clay (%)	Bulk density (g cm-^{3})	FC[b)] (m^3m^{-3})	WP[c)] (m^3m^{-3})	WP[c)] (m^3m^{-3})	Ks[d)] (cm h^{-1})
AS1	0-5	L	51.3	31.6	17.1	1.42	0.303	0.207	0.146	1.91
	5-15	L	42.3	37.0	20.7					
	15-30	L	37.6	36.3	26.0	1.49	0.318	0.246	0.153	1.32
	30-45	SCL	45.7	23.4	30.9	1.44	0.349	0.282	0.176	0.49
	45-60	CL	36.4	27.1	36.5	1.48	0.351	0.293	0.207	0.30
AS2	0-5	L	30.8	45.4	23.8	1.39	0.318	0.218	0.153	1.76
	5-15	L	27.7	46.0	26.4					
	15-30	CL	25.4	33.4	41.3	1.53	0.305	0.236	0.185	0.66
	30-45	CL	26.3	38.0	35.7	1.40	0.358	0.293	0.210	0.60
	45-60	CL	25.1	39.4	35.5	1.38	0.346	0.298	0.205	1.00

[a)] Acronym for soil texture: SCL=sandy clay loam, CL= clay loam, L= loam, C= clay
[b)] FC: field capacity, water content at -33kPa
[c)] WP: wilting point, water content at -1500 kPa
[d)] Ks: Saturated hydraulic conductivity

Soil moisture measurements were obtained at four depths by individual sensors installed at 5 cm, 20 cm, 40 cm and 60 cm. Accuracy of the soil moisture measurement is reported by the manufacturer as ±0.03 water fraction by volume in typical soil. Seyfried et al. (2005) demonstrated the dielectric loss corrected calibration could reduce soil moisture measurement errors. Considering that measured soil moisture in this study was not

corrected with the dielectric loss and near saturated soil condition has the largest errors (Seyfried et al., 2005), actual observation error of the soil moisture may be slightly higher than 0.03 m^3m^{-3}.

In this study, the measured surface soil moisture data (5 cm) were used for data assimilation, and the data at 5, 20, 40 and 60 cm were used for validation purposes. In order to separate the observations that are assimilated into the model from the observations that are used to verify subsequent data assimilation results, especially for the 5 cm observations, the observations (Y_k in Equation 2.4) are compared to the model predicted soil moisture before assimilation step (X_k^{i-} in Equation 2.6) for results analysis. That is, the two sets of observations are considered mutually exclusive. The measured soil moisture and temperature data are also used to set up the initial conditions in the RZWQM.

Measured soil moisture data during the simulation period (April – October 2007) at the two study sites are displayed in Figure 2.2. Observed meteorological data was collected at AS1, and used as model input for both sites since the fields are less than 500 m apart. The top soil layers (0-5 cm), especially for AS1, show very dynamic variations in soil water content with rainfall; whereas deeper layers show more stable soil water content conditions.

2.3 Methodology

In this experiment, I first ran the RZWQM with the input date described in the previous section without any assimilation. This run is called the open loop simulation. Then the model was run for the same period with two different data assimilation techniques: the direct insertion method (DIR) and the Ensemble Kalman Filter (EnKF). The ability of the EnKF in improving the simulated soil moisture profile is evaluated by comparing the results of the open loop and the DIR. Time series graphs and statistical indices such as correlation coefficient, root mean square error and mean bias error are used for the evaluation.

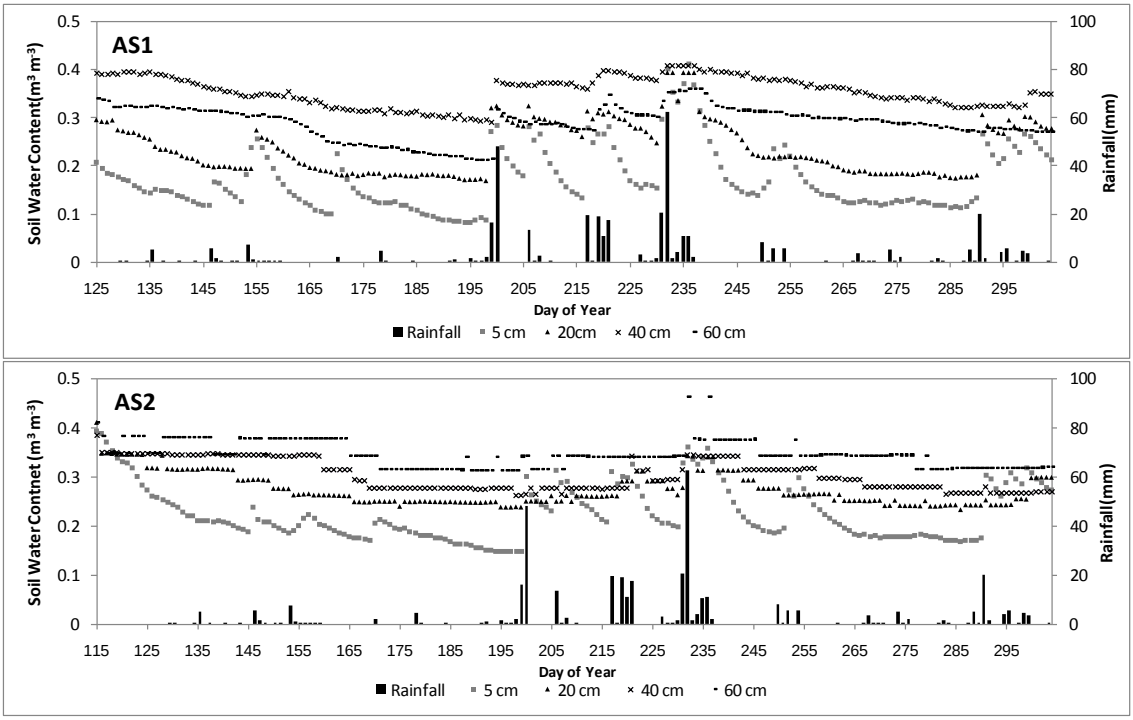


Figure 2.2 Observed Rainfall and Soil Water Content at 5, 20, 40 and 60cm depth (Observation Periods: 5 May to 31 October, 2007 for AS1 and 25 April to 31 October, 2007for AS2)

In this section, I begin with a brief description of the RZWQM focusing on hydrologic processes, which is the model operator in the data assimilation system. A brief review of the EnKF is given in section 2.3.2. Section 2.3.3 describes how the data assimilation algorithm is implemented in the RZWQM to integrate observed surface soil moisture with the model prediction.

2.3.1. RZWQM

The RZWQM is a physically-based, one-dimensional deterministic model that uses fundamental flow equations to simulate infiltration and redistribution in the subsurface region. During a rainfall event, water infiltration is simulated by a modified form of the Green-Ampt equation (Equation 2.1) given below after (Ahuja et al., 2000).

$$V = K_s \frac{\tau_c + H_0 + Z_{wf}}{Z_{wf}} \qquad (2.1)$$

where V is the infiltration rate at any given time (cm h^{-1}), K_s is the effective average saturated hydraulic conductivity of the wetting zone (cm h^{-1}), τ_c is the capillary drive or suction head at the wetting front (cm), H_0 is the depth of surface ponding (cm), and Z_{wf} is the depth of the wetting front (cm).

Redistribution between storm events is simulated by a mixed form of the Richards' equation as given by Equation 2.2 after (Ahuja et al., 2000).

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

Where θ is the volumetric soil water content ($cm^3\ cm^{-3}$), t is time (h), z is the soil depth (cm), h is the soil-water pressure head (cm), K is the unsaturated hydraulic conductivity ($cm\ h^{-1}$), and $S(z,t)$ is the sink term (h^{-1}).

RZWQM adopts the general mass-conservative numerical solution of Celia et al. (1990) to solve the Richards' equation. The sink term takes into account plant water uptake and tile drainage, if present. All other biological and chemical processes for plant growth and movement of nutrients and pesticides in RZWQM are simulated following the physical water flow process. More detailed information for biological and chemical processes in RZWQM can be found in Ahuja et al. (2000).

Measured soil physical and hydraulic properties at the two sites (AS1 and AS2) were obtained from field samples analyzed at the NSERL lab or determined *in situ* and used as model input (Table 2.1). RZWQM provides various options for the estimation of soil hydraulic properties depending on the availability of the data (Ahuja et al., 2000). In this study, minimum necessary input (soil texture, -33 kPa water content and bulk density for each soil layer) are provided and other properties are estimated by the model (Table 2.1). For the constitutive relationship between h-θ-K which is key for the numerical solution of the Richards' equation, RZWQM uses functional forms of those relationships based on the modified Brooks-Corey equation. The parameters for the Brooks-Corey equation were compiled by Rawls et al. (1982) for major USDA soil textures. Validation of this estimation technique can be found in Starks et al. (2003) who showed that the limited input data option, based on textural class name, predicted soil water content as well as

the input from more detailed laboratory measurements. Thus, the parameter values for the Brooks-Corey equation from Rawls et al. (1982) are used in this study.

2.3.2 EnKF Assimilation

The main concept of EnKF, which includes forecasting the error statistics using Monte Carlo methods, was first introduced by Evensen (1994), and has been applied to various fields, especially in oceanography and meteorology (Evensen and Leeuwen, 1996; Burgers et al., 1998; Keppenne, 2000; Houtekamer and Mitchell, 2001). There are slight differences between various ensemble-based filter approaches. For instance, Houtekamer and Mitchell (1998) used two ensembles of model states for forecasting, and updated each ensemble with error covariance from the other ensemble. Whitaker and Hamill (2002) suggested a variant of the EnKF by avoiding perturbation of observations. This study follows the algorithm proposed by Evensen (2003; Evensen, 2004) and notation presented by Reichle et al. (2002b). In this approach, a non-linear system model can be expressed with the following generic form:

$$X_{k+1} = f_k \; X_k \; + w_k \qquad (2.3)$$

Where is the system state of interest (in my case soil water content), indicates a non-linear model operator at time step *k*, and is the system error that accounts for all uncertainties in the model physics or forcing data. The w_k term is treated as a normally distributed random variable with zero mean and covariance Q_k.

Observed soil moisture values (Y_k) are related to the true state ( ) through a measurement operator (H_k) as shown in Equation 2.4:

$$Y_k = H_k X_k + v_k \qquad (2.4)$$

Because measured soil moisture data are available in this study, H_k is an identity matrix. The term represents errors in the measurement instrument and procedures, and conforms to a Gaussian distribution with zero mean and covariance R.

The Kalman filter mainly consists of forecasting and updating steps (Gelb, 1974). Before starting the forecasting or updating of EnKF, an initial ensemble of size N is created by adding pseudorandom noise with zero mean and covariance *P* to get the first "guess" state as shown in Equation 2.5 below.

$$X_0^{i+} = X_0 + e_i \qquad e_i \sim N(0, P) \qquad i=1, \ldots..,N. \qquad (2.5)$$

As a next step, forecasting is performed to estimate the state at a future time by integrating the initially created ensemble through the non-linear model operator, which in this study is redistribution or the infiltration process.

$$X_k^{i-} = f_{k-1}(X_{k-1}^{i+}) + w_{k-1}^i \qquad i = 1, \ldots., N. \qquad (2.6)$$

In Equation 2.5 and 2.6, the superscripts '-'and '+' refer to state variables that are forecasted and updated, respectively. In the standard Kalman filter algorithm, uncertainties in the system are expressed with state error covariance. Contrary to the standard Kalman filter or extended Kalman filter, the EnKF does not require the propagation of state error covariance P_k^- explicitly. Propagation of state error covariance is computationally expensive for a large system, and its exclusion is one of the

advantages offered by the EnKF. Instead of forecasting the state error covariance P_k^- from the covariance of the previous time step P_{k-1}^+, the EnKF estimates error covariance from the forecasted ensemble of the state and their mean as shown in Equation 2.7 below.

$$\mathrm{P}_k^- = \frac{1}{N-1} D_k \mathrm{D}_k^{\mathrm{T}} \qquad (2.7)$$

where and –

Whenever measured data are available, forecasted system variables are updated through the weighted sum of forecasted variables and observed data. The Kalman gain, K_k works as a weight for the updating step as shown below.

$$\mathrm{K}_k = \mathrm{P}_k^- \mathrm{H}_k^{\mathrm{T}} \left[\mathrm{H}_k \mathrm{P}_k^- \mathrm{H}_k^{\mathrm{T}} + \mathrm{R}_k \right]^{-1} \qquad (2.8)$$

$$X_k^{i+} = X_k^{i-} + K_k [Y_k - H_k X_k^{i-} + v_k^i] \qquad i = 1, .., N \qquad (2.9)$$

As Equation 2.9 shows, in the update step, the ensemble of observation size *N* is created by adding random perturbations (v_k^i) with the mean equal to zero and observation variance R_k. After each member of ensemble variables are updated, a best estimate of the system variable at time *k* is found by averaging the updated ensemble members.

2.3.3 Application of EnKF to RZWQM

The basic framework of the RZWQM, and how the EnKF is incorporated into RZWQM processes is presented in Figure 2.3. Model initial conditions are set up using the observation data from the previous day of the first simulation date. Initial ensemble members are generated by adding random errors to the initial condition (Equation 2.5).

The initial covariance P_i is determined by averaging measurement variances at four measurement depths in the soil profile ($V_{5\,cm} + V_{20\,cm} + V_{40\,cm} + V_{60\,cm}$ /4) from the actual ten minute data on the same day as the initial condition. 100 ensemble members are used for all base simulations, but the effect of different ensemble size is also investigated by changing the number of ensemble members from 50 to 500.

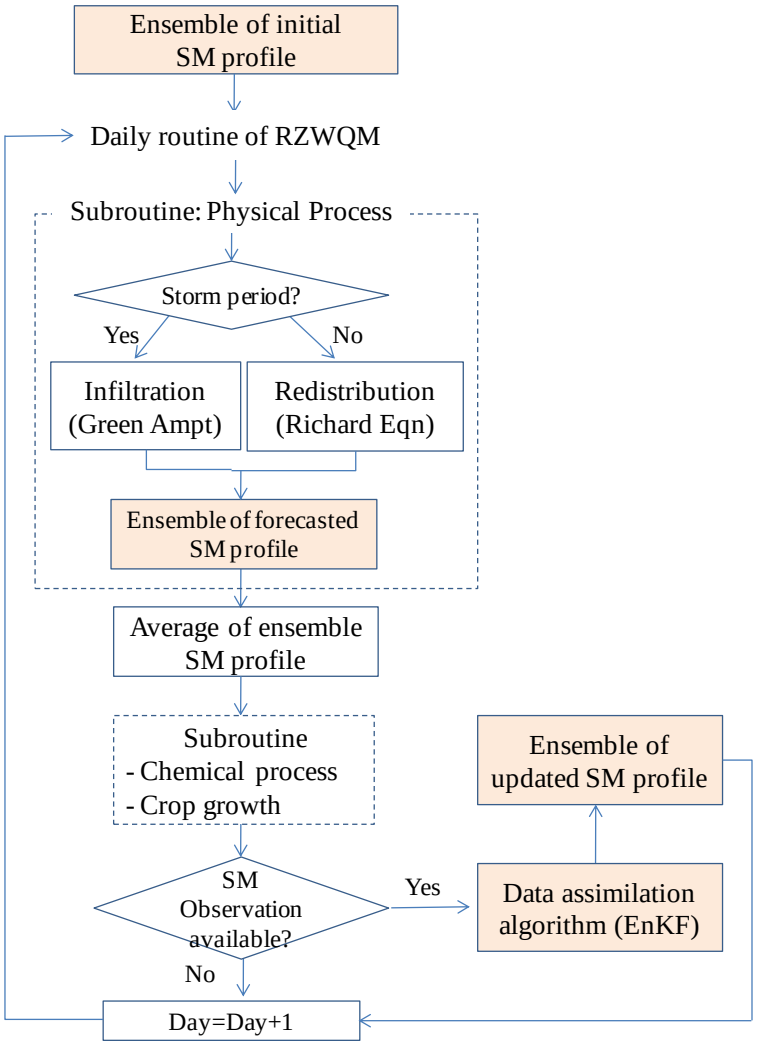


Figure 2.3 Framework of the RZWQM and EnKF Scheme

Most of the subroutines in the RZWQM are based on a daily time step except the physical processes, which are based on a sub-hourly time step (Figure 2.3). In addition, considering that remotely-sensed data from satellites are available on a daily time step or longer, the minimum update interval for this study was taken as one day, even though application of more frequent update intervals is possible from the *in situ* ten minute observations. In order to find the optimum frequency of surface soil moisture assimilation, different update intervals ranging from one day to two weeks were examined in this study with 100 ensemble size.

To evaluate the performance of EnKF compared to other data assimilation techniques, a simple data assimilation scheme, direct insertion method, was also applied in this study. The direct insertion method involves simply the substitution of available observed data for the forecasted system state variable. In this study, it was assumed that the top three numerical layers included in the top 5 cm have the same observed soil moisture as the one measured at the 5 cm depth by Hydra Probe.

Among the measured soil profile moisture data from four different depths (5 cm, 20 cm, 40 cm and 60 cm), data for the top 5 cm depth was used for assimilation. This arrangement mimics the process of assimilating remotely-sensed soil moisture information which is also available only for the top few centimeters of the soil profile. In assimilating data at the daily time step by using measurements collected at ten minute time intervals, the data collected at 11:50 PM is used to update and reinitialize the model simulation at the end of the each day.

In the EnKF application, it is important to have a good knowledge of model and observation errors to define the error covariance (in Equation 2.3 and in Equation 2.4). In studies that use synthetic data, these errors are pre-defined, but in experiments involving actual measurements, these errors are difficult to estimate. In this study, observation errors are mainly from instrument error and error from using point measured soil moisture and ignoring the field-variability of soil water characteristics. Model errors result from the uncertainties in model physics, soil characteristics, and forcing variables (precipitation and other meteorological data). Because these errors were difficult to estimate in this study, appropriate standard deviation of observation and model errors which minimize errors were determined by trial and error. Standard deviations of model errors of 9.58E-3 and 7.20E-3 m^3m^{-3} were used for AS1 and AS2, respectively, and an equal standard deviation of observation errors of 7.07E-3 m^3m^{-3} was used for both sites.

Forecasted soil moisture at each of the four measurement depths was compared with the measured data for evaluation before the updating step. This evaluation was conducted by using time series graphs and standard statistical measures. The correlation coefficient (R) was used to indicate the strength of linear association between the measured and predicted soil moisture values. In addition, root mean square error (RMSE; Equation 2.10) and mean bias error (MBE; Equation 2.11) were used to represent the prediction error and bias respectively.

$$RMSE = \sqrt{\frac{\sum P - O^{2}}{n}} \qquad (2.10)$$

$$MBE = \frac{\sum P - O}{n} \qquad (2.11)$$

Where *P* and *O* are predicted and observed soil moisture, respectively, and *n* is the number of measurements.

2.4 Results and Discussion

2.4.1 Assimilation Results

All modeling results presented in this study are for non-calibrated conditions using a combination of measured, default and model estimated soil parameters or properties. I chose to perform non calibrated simulations to better determine the fundamental affects of data assimilation on the physically-based model performance of soil moisture dynamics in the 0-60 cm root zone. Non calibrated RZWQM simulations were conducted at field sites AS1 and AS2 under the following three conditions: (i) simulation without data assimilation (open loop); (ii) simulation with direction insertion data assimilation method (DIR); and (iii) simulation with EnKF data assimilation (EnKF). As mentioned in section 2.3.3, the soil moisture at 5 cm was used for assimilation, with model predictions compared to measured data at 5, 20, 40 and 60 cm depths. Simulation of RZWQM was conducted from April 1 (Julian day 91) to October 31 (Julian day 304) 2007 to include the effect of plant water uptake and to avoid periods of frozen soil water. Results at AS1 (beginning Julian day 125) and AS2 (beginning Julian day 115) are presented in Figure 2.4 and Table 2.2. For the EnKF simulation, 100 ensemble members were created and the results shown in Figure 2.4 and Table 2.2 are the average of ten simulations.

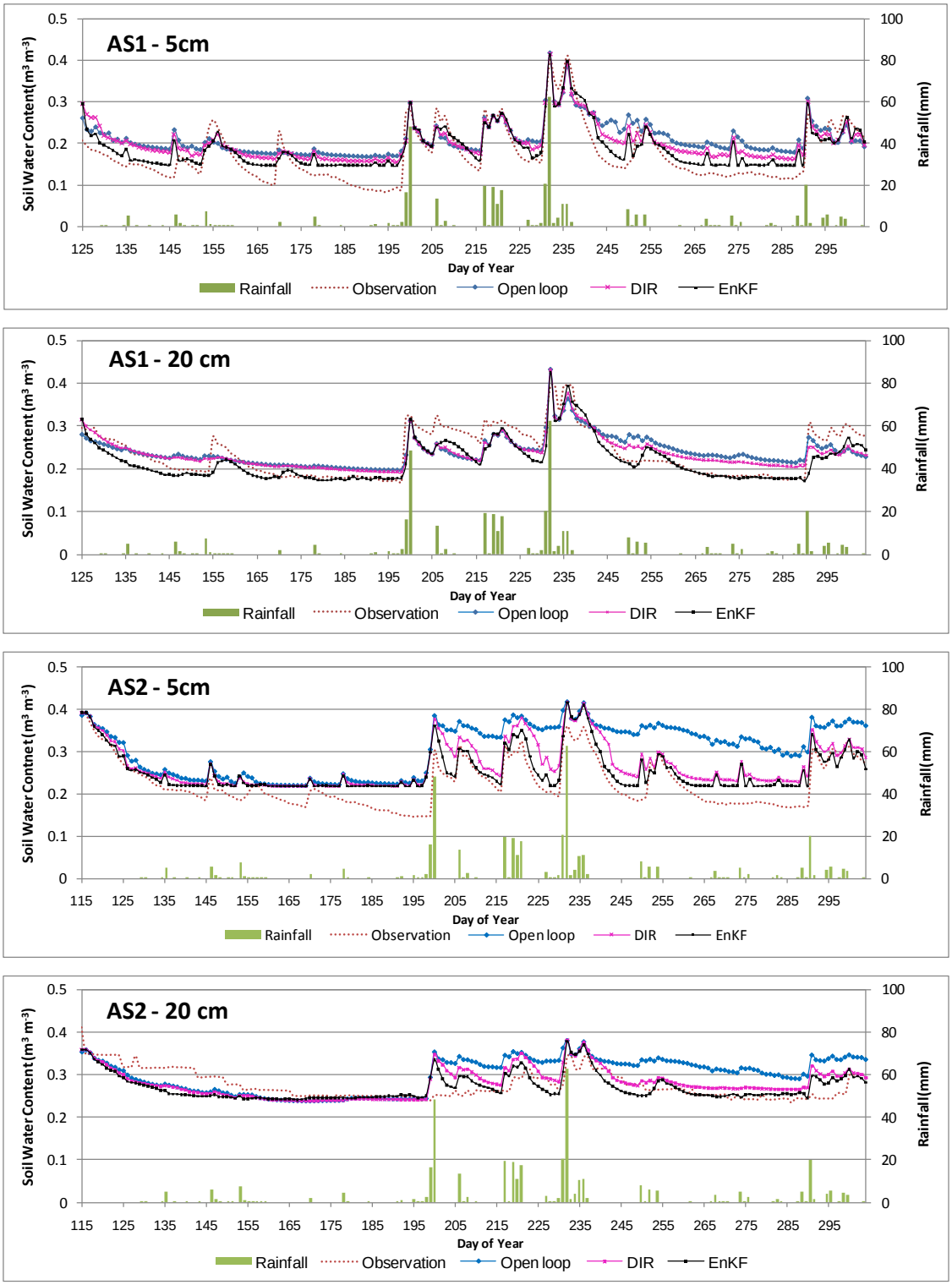


Figure 2.4 Comparison of Observed and Assimilated Soil Moisture

Application of the EnKF improved the open loop results more than the DIR, especially by increasing correlation coefficients to more than 0.90 in the top soil layer (5cm) for both AS1 and AS2, and reducing prediction errors by 38 and 59 % for AS1 and AS2, respectively, compared to open loop (Table 2.2). The AS2 simulation had relatively poor results from the open loop, and data assimilation using both direct insertion and EnKF produced better results (increased correlation coefficients and reduced errors) for all the layers. However, results from the AS1 simulation were more complicated and will be discussed in greater detail below. The open loop simulation did not capture the dynamic variation of soil moisture conditions in the top soil layer especially during the dry-down period. In most cases, the open loop simulation overestimated soil moisture in the top layer while surface soil moisture data assimilation improved simulated soil moisture dynamics and soil moisture estimates.

Table 2.2 Statistical Data Analysis

Site	Depth (cm)	Open Loop			Direct Insertion			EnKF		
		R	RMSE	MBE	R	RMSE	MBE	R	RMSE	MBE
AS1	5	0.816	0.055	0.038	0.873	0.046	0.030	0.912	0.034	0.018
	20	0.785	0.035	0.003	0.843	0.031	-0.001	0.892	0.031	-0.018
	40	0.836	0.086	-0.084	0.861	0.088	-0.087	0.757	0.103	-0.099
	60	0.834	0.027	-0.018	0.831	0.028	-0.020	0.683	0.043	-0.031
AS2	5	0.690	0.091	0.078	0.891	0.049	0.042	0.904	0.037	0.028
	20	0.320	0.050	0.024	0.573	0.031	0.005	0.612	0.029	-0.005
	40	0.109	0.073	0.056	0.336	0.054	0.041	0.383	0.046	0.034
	60	0.412	0.033	0.022	0.460	0.031	0.019	0.485	0.028	0.015

*Simulation period: April – October 2007

The consistent discrepancies between model forecasting and observations shown in Figure 2.4 have been found in other previous studies (Walker et al., 2001a; Reichle and Koster, 2004b; Reichle et al., 2004). Walker et al. (2001a) concluded that removal of the systematic bias in the model forecasts or observation is essential for correct retrieval of profile soil moisture using the Kalman filter assimilation scheme. In regard to remotely-sensed soil moisture data, cumulative distribution function (CDF) matching has been used to reduce the systematic differences between satellite-based and model-based soil moisture estimates effectively (Reichle and Koster, 2004b; Drusch et al., 2005b). The need for the correction or rescaling of the discrepancies (bias) before assimilating soil moisture observations into the model are confirmed in this study as well. Removal of the systematic forecast bias is expected to improve assimilation results especially during the dry-down period.

The observed soil moisture data at the AS1 site (Figure 2.2) shows consistently higher values of soil moisture being measured at the 40 cm depth compared to those at 60 cm, which is not typical of most profile soil moisture distributions. This is because the AS1 sensor site has a dense till soil layer derived from unsorted glacial material at approximately 40 cm which makes lateral flow more dominant than the vertical flow. As a result, the measured soil moisture at the 40 cm depth at AS1 is always higher than the soil moisture at 60 cm depth. Based on soil sampling data, this appears to be a local phenomenon near the sensors and not representative of the entire field. Thus, the model is unable to capture the unique phenomenon at AS1 (predicting higher water content at the 40 cm depth than at 60 cm) and leads to underestimation in soil moisture at 40 cm depth

(high negative MBE in the Table 2.2). The EnKF reduces the open loop overestimated soil moisture at 5 and 20 cm and is closer to the measured values, but this change affects the soil moisture profile at 40 and 60 cm. As a result, the initially underestimated soil moisture at 40 cm and 60 cm was further reduced to give a lower correlation coefficient and higher error (Table 2.2). This local phenomenon, due to the highly heterogeneous soil characteristics, is ignored in large-scale data assimilation studies. However, this finding for the small field scale study points to issues that should be considered for future soil moisture data assimilation studies.

As expected from the previous studies (Walker et al., 2001b; Zhang et al., 2006), the EnKF is shown to be superior to the DIR except for the results of deeper layers at AS1. Walker et al. (2001b) illustrated the difference between the DIR and the EnKF graphically and provided a concise description of both techniques. They describe two limitations of the DIR. First, with the DIR, surface information from observations can be transferred into the deeper layers only through physical infiltration and exfiltration processes while the Kalman Filter modifies the entire soil profile using the covariances of both the surface measurements and model profile prediction. Second, DIR changes the soil moisture profile according to the difference between the measured and simulated soil moisture. The DIR results in this study revealed that the updated (reinitialized) simulated soil moisture, using the surface observations at the end of day, returns to soil moisture values very near those of the open loop after several sub-hourly simulation loops. However, the results of the EnKF are closer to measured values than the DIR (Figure 2.4) resulting in better statistical results. The above-mentioned characteristics of the DIR

approach had less of an effect than the EnKF at the 40 and 60 cm depths at AS1 (Table 2.2).

Use of the EnKF proved superior to the DIR data assimilation technique, which is consistent with findings from previous studies. Even though the DIR method is easy to implement by substituting the observed values into the simulation results directly, it can adjust the results only within the observational depth, while the EnKF corrects the entire soil profile. This study demonstrates the potential for reliable application of the EnKF for surface soil moisture assimilation. However, in spite of satisfactory model performance in this study with application of the EnKF, some precautions should be considered. One of the main reasons for the uncertainty with the EnKF appear to be the presence of bias in the model forecast and violation of the basic assumption of the EnKF (zero mean noise) because of nonlinear processes imbedded in the model and the bounded nature of soil moisture (porosity and residual water content of soil). To overcome these limitations of the EnKF with the real observations, it is desirable to adopt an appropriate bias-correction algorithm for the EnKF (De Lannoy et al., 2007a; De Lannoy et al., 2007c; Ryu et al., 2009). Second, randomly generated ensemble numbers may interfere with the numerical algorithm to solve the Richards' equation in the RZWQM and lead to an infinite routine. Therefore, cautious assignment of variances for the EnKF and proper restriction of random numbers are necessary.

There appear to be inconsistencies between the field-measured soil moisture and laboratory measured soil hydraulic properties as observed in Figure 2.4. For both sites, field-measured soil moisture at the 5 cm depth during the dry period (less than 0.1 m^3m^{-3})

were lower than the laboratory-measured or model estimated water content at -1500 kPa, considered the wilting point in Table 2.2. The field measured soil moisture content is indicative of high evaporative demand during the dry-down period. The RZWQM, however, does not allow the simulated soil moisture to fall below the wilting point when using laboratory measured values as input. Even though the model-estimated inputs for wilting point were used, the simulated soil moisture during the dry period did not go below 0.146 for AS1, and 0.153 for AS2 (Table 2.1). This implies that calibration and optimization of these values should be conducted for future model runs taking into consideration that laboratory measured hydraulic properties are determined based on point-scale soil cores that do not always characterize actual field conditions.

2.4.2 Effect of Update Interval

The effect of update interval was investigated for DIR and EnKF by varying the update interval from 1 to 14 days. AS1 simulation in Figure 2.5 showed that for the 5 cm and 20 cm depths, less frequent update intervals longer than four days did not produce significant benefit from the assimilation scheme even though assimilated simulation results are slightly better than ones of the open loop (0.816 and 0.785, R values for 5 cm and 20 cm, respectively). Interestingly, deep layers (40 and 60 cm depths) have worse results with the frequent update (daily and every two day updates). This seems to be because of the large forecasting errors as Walker et al. (2001a) showed and that the accuracy of the forecasting model is more important than the temporal resolution of observed data.

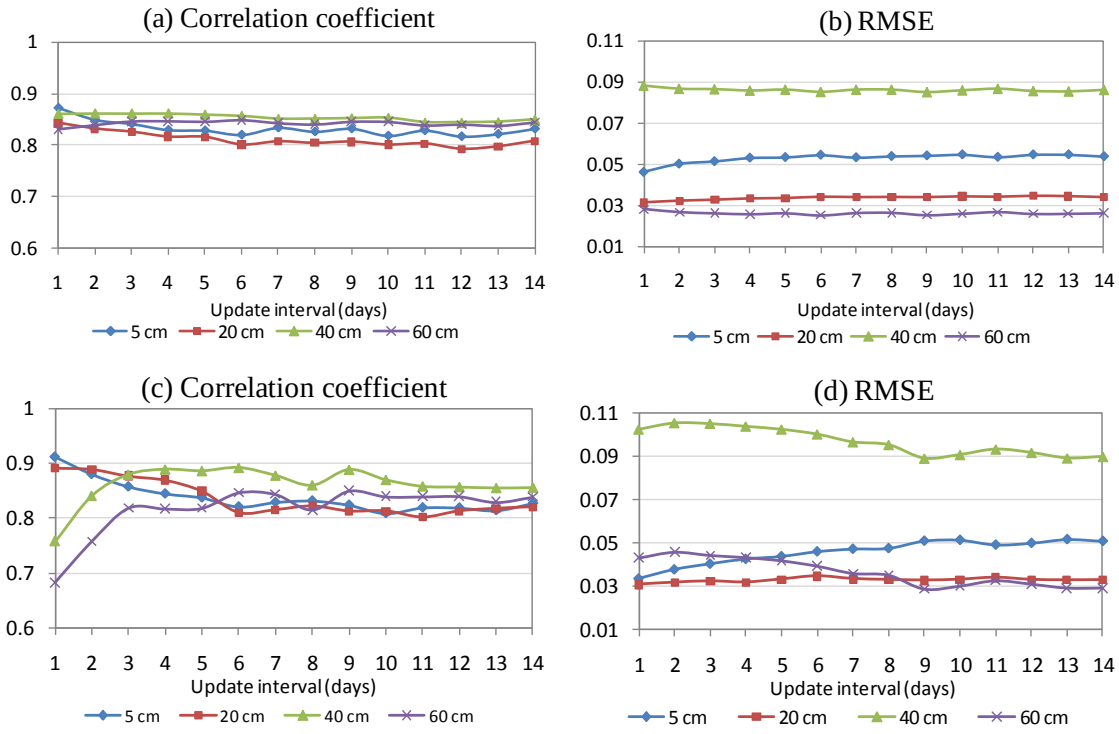


Figure 2.5 Effects of Update Intervals with DIR (a and b) and EnKF (c and d) for AS1

On the effect of update interval with AS2 data, even less frequent update interval improved model predictions compared to open loop simulations (Figure 2.6). For instance, an update interval of two weeks for the 5cm depth raised the correlation coefficient to 0.754 with DIR and 0.784 with EnKF compared to 0.690 with the open loop. In addition, the effect of update interval is more obvious with AS2 simulations. As the update interval increases, errors increased and correlation coefficient decreased apparently. The results of 60 cm depth do not vary much with different update intervals. This is due, in part, to relatively small changes in soil water content at this depth.

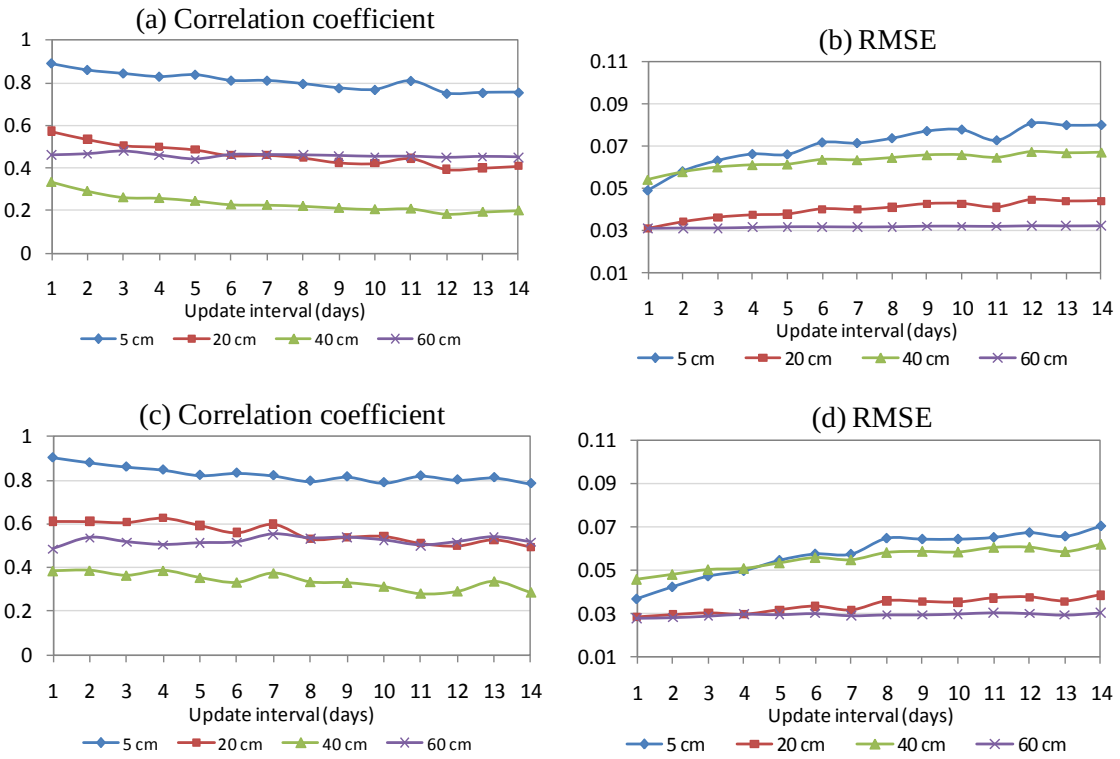


Figure 2.6 Effects of Update Intervals with DIR (a and b) and EnKF (c and d) for AS2

Two previous studies (Walker et al., 2001b; Zhang et al., 2006) investigated the effect of update interval on soil moisture profile retrieval in a desktop study using a one dimensional soil moisture equation and synthetic data. In the work of Walker et al. (2001b), full retrieval of profile soil moisture (1 m) was achieved approximately after 12 days with the DIR and 12 hours with the EnKF for the observation depth of 4 cm when the system was updated once every hour. Their conclusion, in regards to observation interval, was that the frequent observations are more important than observation depth using the EnKF. Zhang et al. (2006) also showed similar results for the update interval impact with a required time for full retrieval being 16 hours with the EnKF and 12 days for the DIR when updated once every hour. With a more realistic "daily" update interval,

they found that the full profile retrieval took 15 days with the EnKF while the DIR failed to retrieve the full profile. The studies mentioned above were based on synthetic experimental data, whereas the current study explores the effect of update interval using actual soil moisture measurements and different governing equations in the RZWQM to find the optimal and effective update interval.

Contrary to the previous synthetic studies (Walker et al. 2001b; Zhang et al. 2006), data assimilation in this study involves much more complexities due to the more complicated physical processes in the RZWQM and limited knowledge on the model and observation errors. In general, it is found that more frequent updates contribute to improved model predictions and that the EnKF is superior to the DIR method of data assimilation.

2.4.3 Effect of Ensemble Size

As mentioned in section 2.3.2, the EnKF is based on the concept of Monte Carlo methods and calculates error statistics from an ensemble of system states at the time of update. Therefore, it requires a sufficient size of ensemble numbers to obtain satisfactory estimates. However for practical and realistic applications of the EnKF, it is useful to know an appropriate size of the ensemble.

In this study, different numbers of ensemble sizes from 50 to 500 were applied to determine how the assimilation results were affected by ensemble size. With a relatively smaller ensemble size of 50, compared to the default value of 100, model performance improved with data assimilation compared to the open loop results; the correlation

coefficient increased by 12% and 13% and RMSE was reduced by 38% and 13% for 5 cm and 20 cm depth, respectively for AS1 (Figure 2.7). For the AS2 simulation which initially had relatively poor prediction from the open loop simulation, an even smaller ensemble sizes (e.g., 50) produced better results. The improvement with the EnKF at AS2 with 50 ensemble numbers is more apparent than with the DIR and open loop; RMSE were reduced by 59% from the open loop results and 24% from the DIR results for the 5 cm depth (not shown).

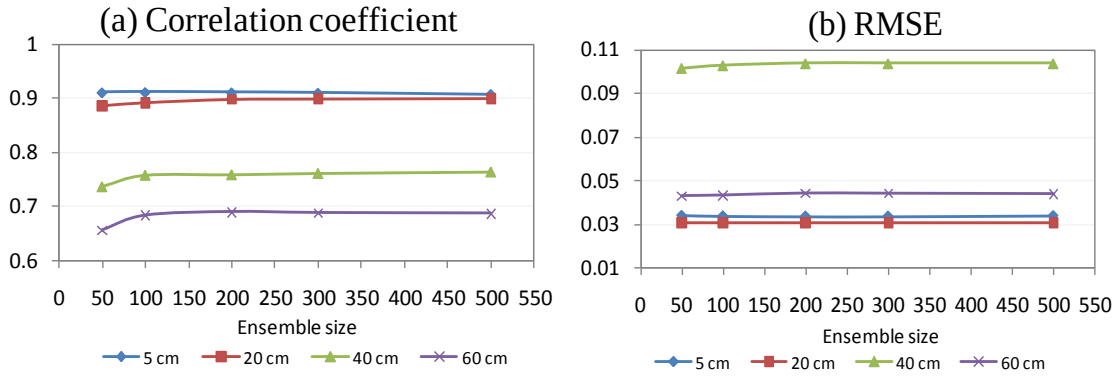


Figure 2.7 Effects of Ensemble Size on the (a) Correlation Coefficient and (b) RMSE for AS1

The sensitivity of EnKF to ensemble size can be seen in Figure 2.7. For AS1 data, the EnKF showed little difference in the efficiency due to ensemble size for upper layers. In the case of deeper layers, smaller ensemble size (50) produced poorer results compared to larger ensemble sizes. Simulations using AS2 data did not show significant differences in predictions between different ensemble sizes (not shown).

Earlier studies for surface soil moisture assimilation with the EnKF (Reichle et al., 2002a; Reichle et al., 2002b) show that the EnKF estimates converge to the true states with

different numbers of ensembles based on synthetic experimental results using a land surface model. Reichle et al. (2002b) presented that an ensemble size greater than ten outperforms the extended Kalman Filter. The work of Reichle et al. (2002a) showed that larger ensemble size improved results even though only 30 ensembles reduced actual errors by 55% compared to the open-loop simulation. In addition, they concluded that for robust estimation of the error covariance, at least 500 ensemble numbers were required. Recently, Zhang et al. (2006) suggested that 40 ensemble members are enough to represent the error covariance for their one dimensional model.

2.5 Conclusion

In this study, field measured surface soil moisture was assimilated into a one dimensional physically based model, the RZWQM, using two different data assimilation techniques, the direct insertion method and the Ensemble Kalman Filter. A key difference of this study from previous studies is the use of real observational data instead of synthetic data for the EnKF assimilation into a field scale water quality model. In addition, this study provides a more practical and operational application of surface soil moisture assimilation using the EnKF by investigating the effect of update intervals and ensemble size. This is a significant aspect in terms of the eventual use of satellite soil moisture products and frequency of observations.

Overall, the results of this study indicate that daily assimilation of surface soil moisture resulted in more accurate soil moisture estimation in the upper, more dynamic, layers (5 and 20 cm). However, improvement of soil moisture prediction in deeper layers (40 and

60 cm) is less certain. Since the observed data are clustered in the top soil layers, there are limitations for the corrected information from the observed surface soil moisture to propagate to the deep layers. Nevertheless, unsuccessful predictions from the open loop for AS1 and AS2 were greatly improved for all layers with both the DIR and the EnKF, except for deeper layers in AS1.

Unique characteristics of soil profile properties and soil moisture distribution at AS1 caused inconsistent assimilation results for the deep layers. That is, the data assimilation techniques, especially the EnKF gave somewhat inadequate results at the 40 cm and 60 cm depths. Therefore, for more successful application of the EnKF, it is recommended to adjust or calibrate the model parameters before data assimilation so that the model prediction can simulate the unique field characteristics properly.

Even though soil hydraulic properties measured in the laboratory and estimated by the RZWQM model were used for the simulation, the open loop results of AS2 highly overestimated the measured soil moisture, thus failing to capture dynamic variations in the soil profile. Considering uncertainties in the laboratory analysis and the high spatial variability of field soil characteristics, it is recommended that the model parameters be optimized through calibration. In this study, however, the model was not calibrated because the purpose of this experiment was to investigate the fundamental effect of different data assimilation techniques on model output compared to the open loop simulation. However, it is also found a need for a proper rescaling step to remove systematic errors between model forecasts and observed data which exist even after calibration.

The investigation into the effect of update interval in the range of 1 to 14 days found that shorter update intervals improve open loop simulation results better than the long update intervals. This linear relationship between the update interval and performance of the data assimilation was more apparent in the upper layers with the AS2 simulation. Results of AS1 do not show much benefit from more frequent update because of its unique soil water distribution in the deep layers.

Theoretically, the EnKF requires large enough ensemble numbers for satisfactory results. In this study, however, the test of the effect of the ensemble size showed that larger ensemble sizes did not produce any significant improvement in the results, and an ensemble of 100 members produced results that were comparable to results from larger ensembles. Especially, the top layer at 5 cm depth did not show any sensitivity to the different ensemble numbers. For the deeper layers, a smaller ensemble size of 50 deteriorated the statistical results slightly but still produced better results than the open loop and the DIR for AS2.

Although data assimilation was shown to contribute to better estimation of profile soil moisture, there are some additional points to consider in future studies. First, data assimilation breaks up the water balance of the system. Contrary to the above-mentioned synthetic experiments, the RZWQM includes very sophisticated physical processes; varying boundary conditions due to surface interaction with the atmosphere and fluctuating water table depth, and plant water uptake and tile drainage as sink terms. Artificial interruption from the surface soil moisture assimilation affects all these processes and further work is necessary to determine to what extent data assimilation

impacts them. Therefore, future investigations involving data assimilation will include the assessment of other hydrologic variables such as ET, runoff, as well as water quality parameters and crop yield.

CHAPTER 3 SOIL MOISTURE DATA ASSIMILATION AT WATERSHED SCALE

3.1 Introduction

One of the key variables in understanding land surface hydrologic processes is soil moisture because it controls the infiltration-runoff mechanism and energy exchange at the land-atmosphere boundary. Therefore, estimating soil moisture has been a long-standing research topic for various purposes in many areas: weather forecast in atmospheric science, flood or drought prediction in hydrology, water quality management in environmental science, irrigation operations in agricultural engineering and soil erosion in soil science (Walker et al., 2001b).

Since the 1990s, remotely-sensed soil moisture data has become much more available while overcoming limitations of traditional *in situ* point measurements of soil moisture (Jackson et al., 1995; Njoku and Entekhabi, 1996; Wagner et al., 1999; Kerr et al., 2001; Njoku et al., 2003; Entekhabi et al., 2008). Remotely-sensed data provide soil moisture estimates for the top soil layer (~5cm). However, information on the moisture condition in the root-zone and subsurface layers is more critical for understanding and simulating many hydrologic processes including evapotranspiration, surface runoff and subsurface flow. The need for profile soil moisture estimates has motivated researchers to integrate measured surface data and hydrologic models to obtain more accurate estimates of soil

moisture content in the root zone through data assimilation techniques (Reichle et al., 2002b; Ni-Meister et al., 2006a; Reichle et al., 2007; Sabater et al., 2007; Das et al., 2008; Draper et al., 2009). However, those surface soil moisture assimilation studies have been conducted at regional or global scales with land surface models in hydrometeorology for better initialization of soil moisture conditions. Even though great progress in surface soil moisture assimilation studies has been made in land surface atmospheric interactions, there is still a lack of research on utilizing remotely-sensed soil moisture and data assimilation techniques for catchment scale water resource management problems (Troch et al., 2003).

The main objective of this study was to investigate the effect of surface soil moisture data assimilation on hydrological response at the catchment scale through a synthetic experiment. Using intentionally limited rainfall input, I investigate how soil moisture update through surface soil moisture assimilation may compensate for errors in the hydrologic prediction due to the inaccurate rainfall. Simply focusing on the streamflow prediction overlooks the different contributions of each water balance component such as evapotranspiration, infiltration, surface runoff and lateral flow to the streamflow. Therefore, in this study, a catchment scale, continuous time, semi-distributed hydrologic model, the Soil and Water Assessment Tool (SWAT), is used to determine and account for the sources of error in streamflow prediction. In addition, this study also aims to investigate the effects of spatially varying input such as landuse and soil type on data assimilation results.

In this study, one of the most popular data assimilation techniques, the Ensemble Kalman Filter (EnKF) is used to assimilate surface soil moisture observations into the model and a synthetic experiment is conducted assuming that uncertainties in the model and observations are known. Previous studies related to this study are summarized in section 3.1.1. Brief explanations about the EnKF and the SWAT model are described in section 3.2, followed by the illustration of how I conducted the synthetic assimilation experiments in section 3.3. Section 3.4 shows the results of the experiments with discussions.

3.1.1 Previous Studies

This section briefly summarizes previous studies in the context of why surface soil moisture data assimilation at catchment scale is important, specifically for runoff prediction and agricultural applications.

Satellite-based surface soil moisture observations have received considerable attention in runoff or flood forecasts because antecedent soil moisture condition is a critical factor in rainfall-runoff modeling. The recently launched European Space Agency (ESA)'s Soil Moisture and Ocean Salinity **(**SMOS**)** mission (2009) and the upcoming NASA Soil Moisture Active/Passive (SMAP) mission (2014) are designed to better measure soil moisture on a global scale (Entekhabi et al., 2010; Kerr et al., 2010). One of the main application areas of SMAP is to improve flood forecasts using soil moisture measurements at high spatial and temporal resolutions (Entekhabi et al. 2010). Some previous studies have demonstrated the potential of using remotely-sensed soil moisture

to improve streamflow prediction through updating initial soil moisture conditions and finding correlation between soil moisture condition and runoff (Jacobs et al., 2003; Scipal et al., 2005; Weissling et al., 2007).

With regard to assimilating remotely-sensed surface soil moisture into rainfall-runoff models, however, there are few studies to date. Pauwels et al. (2001) assimilated remotely-sensed surface soil moisture data into the TOPMODEL based Land-Atmosphere Transfer Scheme (TOPLATS) using the 'nudging to individual observations' and 'statistical correction assimilation' methods. They applied these methods for both the distributed and lumped versions of the model and concluded that improvement in the discharge prediction can be sufficiently obtained through assimilating the statistics of the remotely-sensed soil moisture data into the lumped model. Crow and Ryu (2009) adopted a smoothing framework for runoff prediction and used remotely-sensed soil moisture to improve both pre-storm soil moisture conditions and external rainfall input. They showed that their smoothing framework improved streamflow prediction, especially for high flow events more than simply updating antecedent soil moisture conditions.

Microwave remote sensing soil moisture observations have high potential for agricultural applications. Some obvious examples are for crop yield forecasting, drought monitoring and early warning, insect and disease control, optimal fertilizer application, and operational decision-making for effective irrigation (Engman, 1991; Lakhankar et al., 2009). However, very few studies exist on the application of remotely-sensed soil moisture for agricultural operations. Jackson et al. (1987) tested the feasibility of using airborne microwave sensors to assess the preplanting soil moisture condition by creating

a soil moisture map which represents overall soil moisture patterns and variations expected over large areas. Recently, soil moisture data from the Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) was incorporated into a global agricultural model by Bolten et al (2009). Their preliminary results showed that surface soil moisture assimilation has potential to improve the crop yield prediction capability of a global crop production decision support system.

Soil moisture information with high spatial and temporal resolution is also required for watershed or field scale agricultural applications. Timely and cost effective operational decisions for on-farm irrigation and trafficability should be based on field-specific information (Jackson et al., 1987). Narasimhan et al. (2005) used the Soil and Water Assessment Tool (SWAT) to produce a long-term soil moisture dataset for the purpose of drought monitoring and crop yield prediction. They selected the SWAT model because of its capability: 1) for simulating crop growth and land management and, 2) for incorporating all available spatial information for the watershed. The SWAT model has also been successfully used to examine the temporal variability of soil moisture over longer time periods (e.g. 30 years) with detailed land surface information (DeLiberty and Legates, 2003).

Based on the aforementioned previous studies, there is a strong need to estimate soil moisture content through assimilating remotely-sensed soil moisture into a long-term, physically based distributed catchment scale hydrologic model. Most of the previous studies that explored data assimilation for runoff simulation used conceptual rainfall-runoff models (Aubert et al., 2003a; Weerts and El Serafy, 2006; Crow and Ryu, 2009;

van Delft et al., 2009) or lumped models (Jacobs et al. 2003) or for short-term periods with real measurements (Pauwels et al. 2001). From an agricultural aspect, soil moisture reserve between rainfall events is also critical for scheduling water supply for crops. Therefore, it is desirable to apply data assimilation techniques to physically based and continuous time hydrological models to address various water resource problems at catchment scales.

Recently, SWAT has been used to assess the capability of the EnKF for catchment scale modeling (Xie and Zhang, 2010; Chen et al., 2011). Xie and Zhang (2010) explored combined state-parameter estimation using different types of measurements based on a synthetic experiment and demonstrated effective update of hydrological states and improved parameter estimation using the EnKF. Chen et al. (2011) conducted both synthetic and real data EnKF experiments. Their results showed improved update for soil moisture in the upper soil layers, but limited success for deeper soil layers and streamflow prediction due to the insufficient vertical coupling strength of SWAT.

3.2 Hydrologic Model and Data Assimilation

3.2.1 Soil and Water Assessment Tool (SWAT)

The Soil and Water Assessment Tool (SWAT) is categorized as a physically based basin-scale, continuous time and semi-distributed hydrologic model. Spatially distributed data related to soil, landuse, topography and weather are used as input to SWAT. The SWAT

model has the capability to simulate plant growth, nutrients, pesticides and land management as well as hydrology on a daily time step and has been proven as an effective tool for assessing water resource and nonpoint-source pollution problems after being applied to watersheds of different scales and characteristics (Gassman et al., 2007). In addition, it is considered as one of the promising models for long-term simulations in predominantly agricultural watersheds (Borah and Bera, 2003), similar to the study area used in the present study, Upper Cedar Creek Watershed in Indiana. Considering the above factors, SWAT is used to accomplish the objectives of this study. The 2005 version of the SWAT model is used in this study, and detailed information of SWAT 2005 can be found in Neitsch et al. (2002).

In the SWAT model, a watershed is first subdivided into sub-basins based on topography, and each sub-watershed is further divided into hydrologic response units (HRU) based on soil and landuse characteristics. The soil profile can be subdivided into multiple layers for up to 2 meters. Hydrological processes in the SWAT, including soil moisture accounting, are based on the water balance equation (Equation 3.1).

$$SW_t = SW_{t-1} + \sum_{i=1}^{t} \left(P_i - Q_{surf,i} - ET_i - Q_{loss,i} - Q_{gw,i} \right) \quad (3.1)$$

where SW_t is the soil water content above the wilting point at the end of day *t*. P_i is the amount of precipitation on day *i* and $Q_{surf,i}$, ET_i, $Q_{loss,i}$ and $Q_{gw,i}$ are the daily amounts of surface runoff, evapotranspiration, percolation into deep aquifer, and lateral subsurface flow, respectively. All components are estimated as mm of H_2O (mmH_2O).

SWAT simulates surface runoff using either the modified SCS curve number (CN) method or Green Ampt Mein-Larson excess rainfall method (GAML) depending on the availability of daily or hourly precipitation data, respectively (Neitsch et al. 2002). In this study, the modified SCS CN method is used with daily precipitation data. The SCS CN method computes accumulated surface runoff (Q_{surf}) using the empirical relationship between rainfall (P_i), the initial abstraction (I_a) and retention parameter (S_i) as shown in Equation 3.2. SWAT follows the typical assumption of $I_a = 0.2 \cdot S_i$

$$Q_{surf,i} = \frac{(P_i - I_a)^2}{(P_i - I_a + S_i)} = \frac{(P_i - 0.2S_i)^2}{(P_i + 0.8S_i)} \qquad (3.2)$$

S_i is a function of the daily curve number(CN_i). Both S_i and CN_i vary spatially with the soil type, landuse management type, and slope.

$$S_i = 25.4\left(\frac{1000}{CN_i} - 10\right) \qquad (3.3)$$

Since the daily curve number varies with the antecedent soil moisture condition, SCS defines three different curve numbers: CN_1-dry, CN_2-average moisture, and CN_3-wet. This classification, however, is too coarse to reflect antecedent soil moisture condition accurately. Thus SWAT adopted a new equation to compute S_i as a function of soil profile water content (SW_i).

$$S_i = S_{\max,i}\left(1 - \frac{SW_i}{SW_i + \exp(w_1 - w_2 \cdot SW_i)}\right) \qquad (3.4)$$

where $S_{\max,i}$ is the maximum retention parameter that can be achieved on any given day. w_1 and w_2 are shape coefficients determined from the amount of water in the soil profile at field capacity and when fully saturated, and retention parameters for CN_1 and CN_3 conditions (Neitsch et al., 2002).

USDA-SCS (USDA-SCS, 1972) states that the SCS curve number method is designed to estimate "direct runoff" which is composed of channel runoff, surface runoff and subsurface flow, excluding base flow. However, SWAT 2005 uses the CN method to estimate "surface runoff" (Equation 3.2) and uses other equations to compute subsurface lateral flow and groundwater flow (Neitsch et al. 2002). Subsequently, the sum of the surface runoff, subsurface lateral flow and groundwater flow generates streamflow. Thus, according to Neitsch et al. (2002), this study uses the CN method in estimating only surface runoff (Equation 3.2).

The excess water available after initial abstractions and surface runoff infiltrates into the soil. The flow through each layer is simulated using a storage routing technique. Unsaturated flow between layers is indirectly modeled using the depth distribution of plant water uptake and soil water evaporation. Saturated flow is directly simulated and it is assumed that water is uniformly distributed within a given layer. When the soil water in the layer exceeds field capacity and the layer below is not saturated, downward flow occurs and its rate is governed by the saturated hydraulic conductivity. A kinematic storage routing technique is used to simulate lateral flow in the soil profile based on slope, hillslope length and saturated conductivity. Upward flow from a lower layer to an upper layer is simulated by the soil water to field capacity ratios of the two layers.

3.2.2 Ensemble Kalman Filter (EnKF)

In hydrology, data assimilation techniques have been used to improve model predictions by combining uncertain observations and imperfect information from hydrological processes represented in hydrological models (Walker and Houser, 2005). Among various data assimilation methods, the standard Kalman Filter is a sequential data assimilation method for linear dynamics minimizing the mean of the squared errors in state estimates. For nonlinear applications, the extended Kalman filter has been applied (Entekhabi et al., 1994; Walker and Houser, 2001; Draper et al., 2009), but this method requires high computational cost for the error covariance integration and is known as very unstable if the nonlinearities are severe (Miller et al., 1994; Reichle et al., 2002a). Evensen (1994) introduced the Ensemble Kalman Filter (EnKF) and proved that it could successfully handle strongly nonlinear systems with low computational cost unlike the extended Kalman Filter. The EnKF has gained popularity in hydrologic data assimilation and a number of previous studies have demonstrated its performance in improving hydrological predictions (Reichle et al., 2002a; Reichle et al., 2002b; Zhang et al., 2006; Clark et al., 2008; Komma et al., 2008; Xie and Zhang, 2010). In this study, the well-proven EnKF method is used to assimilate surface soil moisture observations into the SWAT model.

Hydrological processes such as infiltration, evapotranspiration and drainage to the groundwater system influence soil water content in the root zone in a highly nonlinear manner. The state variable, soil moisture (θ_k) is a vector containing soil moisture values

for each soil layer in a HRU and is described with a nonlinear model operator, $f_k(\cdot)$ at time step k.

$$\theta_{k+1} = f_k(\theta_k) + w_k \quad (3.5)$$

In the Equation 3.5, w_k represents all uncertainties in the forcing data and model formulation due to the numerical approximation and imperfect knowledge of the hydrologic processes.

The surface soil moisture measurement (d_k) has error (v_k) due to the errors in the observational instruments or procedures. It is explained using the observation operator (H_k) and true state (θ_k) as follows:

$$d_k = H_k \theta_k + v_k \quad (3.6)$$

The EnKF is based on the concept of the statistical Monte Carlo method, where the ensemble of system states marches in state space and the mean of the ensemble is considered as the best estimate of the state. The initial ensemble of state vectors is generated with initial error vector (e_i) with N ensemble members ($i=1, \ldots..,N$).

$$\theta_0^{i+} = \theta_0 + e_i \quad (3.7)$$

The EnKF approach assumes that the error terms, w, v and e, are white noise (uncorrelated in time) and have the Gaussian distribution with zero mean and covariance Q, R, and P, respectively.

Once the initial ensemble is created, the ensemble of state vectors integrates forward in time through the nonlinear model operator (Equation 3.8) and is updated using the Kalman Gain (K_k) whenever the observations are available (Equation 3.9).

$$\theta_k^{i-} = f_{k-1}\left(\theta_{k-1}^{i+}\right) + w_{k-1}^{i} \quad (3.8)$$

$$\theta_k^{i+} = \theta_k^{i-} + K_k\left[z_k - H_k\theta_k^{i-} + v_k^{i}\right] \quad (3.9)$$

where superscripts '-' and '+' refer to the forecasted and updated state variables respectively. In the update step, the observation is perturbed with random errors (v_k^i), following Burgers et al. (1998). The Kalman gain (K_k) works as a weighting factor between uncertainties in model prediction and observations, and is determined by the state error covariance (P_k^-) and covariance matrix of the model predicted observation ($H_k P_k^- H_k^T$).

$$K_k = P_k^- H_k^T\left[H_k P_k^- H_k^T + R_k\right]^{-1} \quad (3.10)$$

Unlike the traditional Kalman Filter, the EnKF does not propagate the state error covariance (P_k^-) explicitly. Instead, the EnKF simply estimates the error covariance using the forecasted ensemble of state and its mean (Equation 3.11), thus improving computational efficiency.

$$P_k^{-1} \approx \frac{1}{N-1}\sum_{i=1}^{N}\left[\left(\theta_k^{i-} - \theta_k^-\right)\left(\theta_k^{i-} - \theta_k^-\right)^T\right]^{-1} \quad (3.11)$$

where $\theta_k^- = \frac{1}{N}\sum_{i=1}^{N}\theta_k^{i-}$

The average of the updated ensemble members is determined as a best estimate of the state variable.

3.3 Methodology

3.3.1 Study Area and Data for SWAT Simulation

The study area for this work is the Upper Cedar Creek Watershed (UCCW) which is located in the St. Joseph Watershed in northeastern Indiana (Figure 3.1). The predominant landuse in the UCCW is agricultural, with major crops of corn and soybeans, and minor crops of winter wheat, oats, alfalfa, and pasture. The area receives approximately 94 cm of annual precipitation and has average daily temperatures ranging from -1 ℃ to 28 ℃.

The United States Department of Agriculture, Agricultural Research Service (USDA-ARS) National Soil Erosion Research Laboratory (NSERL) has established an environmental monitoring network in UCCW as a part of the Source Water Protection Project and the Conservation Effect Assessment Project. The environmental monitoring network in UCCW has been operational since 2002. More details about the network can be found in Flanagan et al. (2008). The UCCW environmental monitoring network provides considerable meteorological data (ten minute rainfall, air temperature, solar radiation, wind speed and relative humidity) and ten minute soil moisture observations at four different depths (5, 20, 40 and 60 cm). Recently, UCCW has been selected as one of several USA core sites to be used for calibration and validation of future SMAP data. Substantial *in situ* data, and its importance for future remote sensing validation work makes UCCW an ideal test bed for this study.

In addition to the data from the monitoring network, daily precipitation and temperature data are available from nearby three National Climatic Data Center (NCDC) weather stations for the model simulation period, from 2003 to 2010 (Figure 3.1).

For the SWAT simulation, the sub-basins and stream network were delineated from the 30 meter Digital Elevation Model (DEM) from the National Elevation Dataset. ArcSWAT, an ArcGIS interface for SWAT, was used for delineating the sub-basins and creating other input files for the model. By using a critical threshold area of 500 ha, the UCCW is divided into 25 sub-basins. Further division of sub-basins into HRU units requires soil and land use information. Soil information is retrieved from the Soil Survey Geographic Database (SSURGO) available from the USDA Natural Resources Conservation Service. Major soil types in the UCCW and their areas are summarized in Table 3.1. Landuse information was obtained from the 2001 National Land Cover Data (NLCD 2001) produced by the Multi-resolution Land Characteristics Consortium. The NLCD 2001 data for the UCCW is reclassified to 15 different landcover classes for this study (Table 3.2). By using a 10% threshold for both landuse and soil, and 0% threshold for slope, a total of 209 HRUs were generated for the UCCW using ArcSWAT.

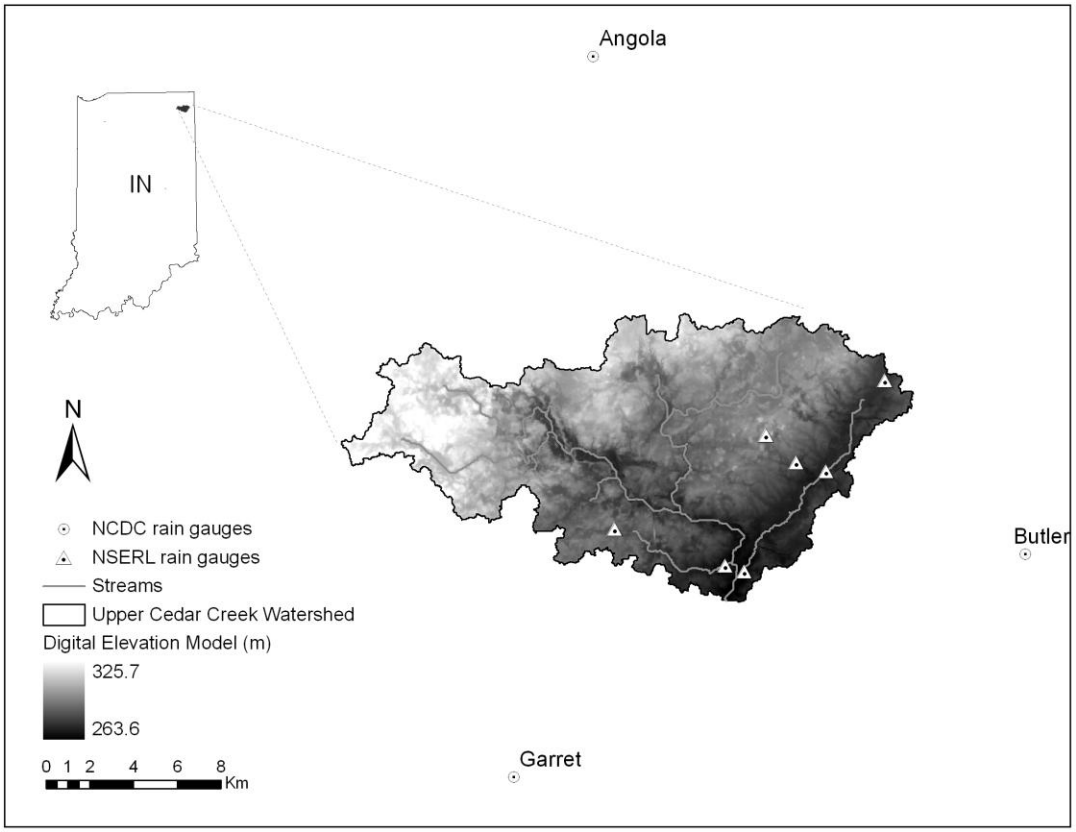


Figure 3.1 Study Area: Upper Cedar Creek Watershed in Indiana, USA.

Table 3.1 Soil Types in UCCW

Soil type		Area [km^2]	Proportion (%)	Hydrologic soil group
GnB2	Glynwood loam	52.69	26.67	C
BaB2	Blount silt loam	43.89	22.21	D
Pe	Pewamo silty clay	35.59	18.01	C
SrB2	Strawn loam	30.71	15.55	B
RaB	Rawson sandy loam	7.43	3.76	C
MrC3	Morley silty clay loam	5.38	2.72	D
StC3	Strawn clay loam	4.83	2.44	B
Hw	Houghton muck	4.59	2.32	A
Re	Rensselaer loam	4.14	2.09	B
BoB	Boyer sandy loam	3.84	1.94	A
Se	Sebewa sandy loam	1.47	0.75	B
All others		3.01	1.54	

Table 3.2 Landuse in UCCW

Landuse classification		Area	
		[km^2]	Proportion (%)
AGRR	Agricultural	113.67	57.54
HAY	Hay	30.75	15.56
FRSD	Forest, deciduous	23.59	11.94
URLD	Residential, low intensity	8.38	4.24
WETF	Wetlands-forested	7.99	4.04
URMD	Residential, medium intensity	4.03	2.04
RNGB	Range, brush	3.82	1.94
RNGE	Range, grasses	2.89	1.46
WATR	Water	0.91	0.46
URHD	Residential, high intensity	0.76	0.38
WETN	Wetlands, non-forested	0.41	0.21
FRSE	Forest, evergreen	0.16	0.08
UIDU	Industrial	0.10	0.05
FRST	Forest, mixed	0.08	0.04
SWRN	Barren land	0.02	0.01
	Sum	197.57	99.99

3.3.2 Experimental Setup

Synthetic experiment

Data assimilation study requires reliable observed data and decent knowledge on the uncertainties in the model and observations. In the present study, a synthetic experiment is performed to better understand how soil moisture data assimilation affects various hydrologic processes in the watershed scale SWAT model. The synthetic experiment assumes that errors in the model and observations are known, which allows us to focus only on the effect of data assimilation. Thus, this synthetic experiment at watershed scale should contribute to the practical application of the forthcoming remotely-sensed soil moisture data.

In this study, it is assumed that surface soil moisture observations for each HRU are available and the synthetic observed data was created by adding random observation errors with zero mean. Considering that the typical microwave remote sensing data have a penetrating depth up to 5cm, I also assume that the synthetically observed soil moisture is available for the top 5cm and the input SSURGO soil database is modified so that each HRU has the first two layers in the 5cm depth interval.

Experimental design

Design of this study includes three separate scenarios all run for the same time period: a true run, open loop and the EnKF. First the experiment starts with the implementation of the "true" state by running the model with all available rainfall data from the three NCDC stations and the seven NSERL rain gauges (Figure 3.1). To represent our imperfect knowledge of the true hydrologic processes, the model is subsequently run with an intentionally poor set of initial conditions and with "limited" forcing data from only the NCDC rain gauges. Hereafter, this is called the "open loop" scenario. The third scenario, referred to as "EnKF" includes model integration with all the same input including rainfall and model parameters as the open loop, but updating soil moisture by assimilating daily (synthetic) observed surface soil moisture through the EnKF. By limiting precipitation input, which is the driving force of soil moisture and streamflow, while keeping other model parameters unchanged, the EnKF scenario enables the determination of the influence of the surface soil moisture assimilation on model predictions of profile soil water content and other hydrological responses.

Conventionally, precipitation data measured at point rain gauges have been used for hydrological modeling, despite its limitations in representing the spatial distribution and variability of actual precipitation. In reality, many watersheds rely on the precipitation data from the sparsely located rain gauges inside or around the watershed as the open loop scenario of this study represents. Therefore, in this study, I investigate how the surface soil moisture assimilation may compensate for the errors from the inaccurate precipitation input by comparing the aforementioned three scenarios (true, open loop and EnKF).

SWAT assigns precipitation values at the subbasin level based on precipitation data measured at rain gauges. Each subbasin takes the precipitation value from the gage station that is closest to the centroid of the subbasin. Unfortunately, there is no NCDC rain gauge inside the UCCW. Therefore, when the precipitation data from the only three NCDC rain gauges are used (open loop and EnKF), SWAT assigns overestimated precipitation into the watershed (Figure 3.2). The average precipitation of the true scenario and the open loop scenario (and EnKF) are 2.66 mm day^{-1} and 3.21 mm day^{-1}, respectively during the analysis period (from June 2008 to May 2009).

SWAT calibration

In order to create the true scenario, the SWAT model was calibrated using the daily streamflow measured at the outlet of the UCCW. The model ran for three years (April 2003 to April 2006) for warm-up and two years (May 2006 to May 2008) for calibration. Based on the sensitivity analysis results, 18 parameters were chosen for calibration

including CN2, ESCO, SURLAG, TIMP, Ch_K2, SOL_AWC, BLAI, ALPHA_BF, SOL_Z, RCHRG_DP, SMTMP, EPCO, CH_N2, REVAPMN, SLOPE, SOL_ALB, CANMX, SOL_K. They were calibrated using the autocalibration tool, Parasol (Parameter Solutions) method, implemented in the ArcSWAT. After the calibration, the model was validated for one year (June 2008 to May 2009). The coefficient of determination (R^2) and the Nash and Sutcliffe model efficiency coefficient (E_{NS}) are 0.46 and 0.44 for the calibration period and 0.42 and 0.37 for validation period, respectively with the daily streamflow predictions. Although these statistical metrics seem to indicate marginal model performance, they are within an acceptable range at the daily time step (Moriasi et al. 2007). This validation period will be used for data assimilation experiment of this study. For the true scenario, the model was run from 2003 to 2009 with the calibrated parameters, and the simulation results only from the last one year (June 2008 to May 2009) were used for analysis. For the open loop and EnKF scenarios, the model was run from January 2008 to May 2009 with an intentionally short warm-up period to simulate poor initial conditions and the same calibrated parameters as the true scenario. Again the last one year simulation results (June 2008 to May 2009) were used for analysis.

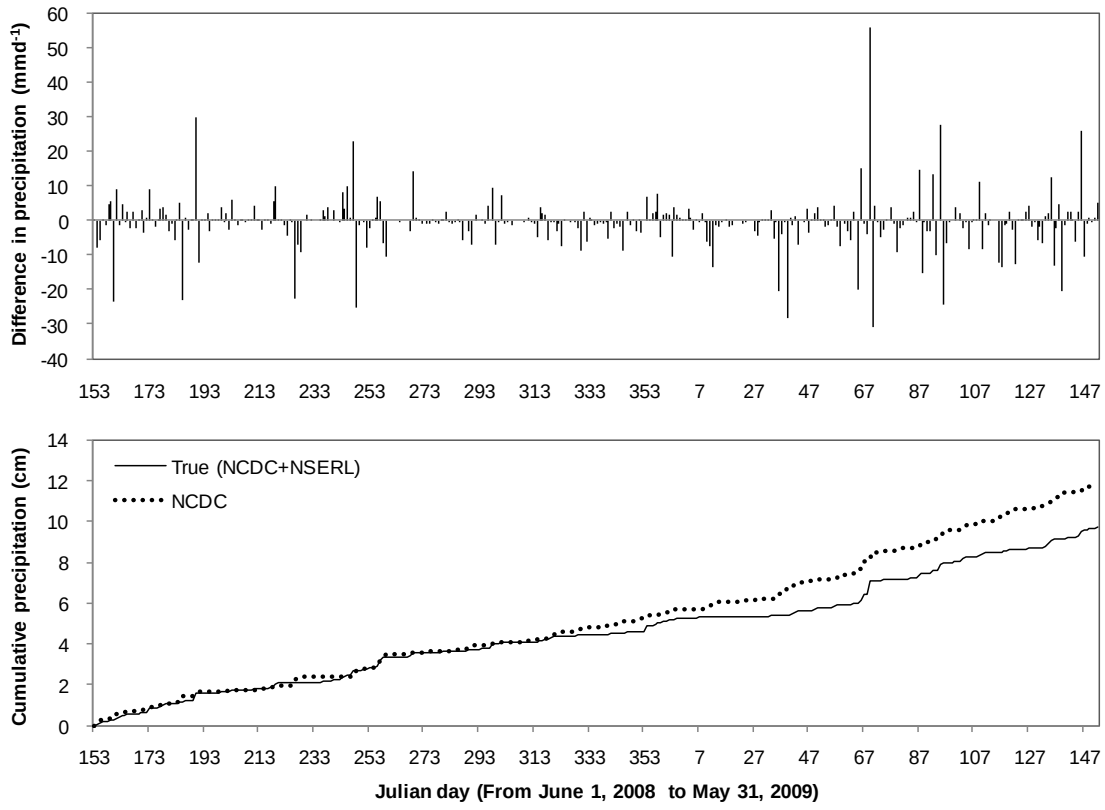


Figure 3.2 Comparison of two different precipitation inputs
(a) difference in precipitation rate (true precipitation from all available rain gauges less precipitation from the NCDC rain gauges) (b) cumulative total precipitation

Evaluation metrics

Evaluation of the surface soil assimilation through the EnKF is first conducted by comparing time series of soil water storage and other representative hydrological responses obtained from the true, open loop and EnKF scenarios. Root mean square error (RMSE), mean bias error (MBE) and correlation coefficient (R) are used to compare the results quantitatively.

$$RMSE = \sqrt{\frac{\sum (\quad - S)^2}{n}} \qquad (3.12)$$

$$MBE = \frac{\sum (T - S)}{n} \quad (3.13)$$

$$R = \frac{\sum (T - \bar{T})(S - \bar{S})}{\sqrt{\sum (T - \bar{T})^2}\sqrt{\sum (S - \bar{S})^2}} \quad (3.14)$$

where T and S are true and simulated values respectively, and n is the number of values (days). $\bar{T}$ and $\bar{S}$ are averages of true and simulated values respectively.

In addition, distributed soil moisture maps are shown to better visualize and illustrate the spatial differences in the results and to determine the effect of spatially varying input.

3.3.3 Implementation of the EnKF into SWAT

Ensemble simulation with SWAT

The SWAT model runs on a daily time step. This study assumes that daily soil moisture observations are available, and therefore soil moisture condition is reinitialized for each day through the analysis (update) step of the EnKF. The framework of hydrologic processes in the SWAT and additional routines due to the EnKF are illustrated in Figure 3.3. The default SWAT routines in Figure 3.3 are excerpted from Neitsch et al. (2002). Each HRU runs independently and the one dimensional EnKF scheme is applied to each HRU separately.

Ensemble simulation starts on day 153 (June 1) of 2008 and 100 initial ensemble members are created by adding random noise to the initial condition (soil moisture simulation results from day 152). SWAT represents soil water content as a concept of soil water storage in mmH_2O (vs $m^3\ m^{-3}$) and therefore the range of soil moisture values for

each layer varies with the thickness of the soil layer. The initial perturbation noise is assigned to the normalized soil moisture values (mmH_2O per 1cm). The initial perturbation noise has a Gaussian distribution with zero mean and 0.001 mmH_2O of standard deviation, and the magnitude of the standard deviations are scaled by the thickness of each layer according to Ryu et al. (2009). Note that forcing variables (e.g. precipitation) and model parameters are not perturbed explicitly to generate ensemble members in this study. Adding model error (w_k in Equation 3.5) takes into account all uncertainties raised from forcing variables and parameter estimation, as well as model physics.

Each of the ensemble members of soil moisture goes through separate hydrological processes and generates different subsequent variables such as Leaf Area Index and evapotranspiration. All subsequent variables from different ensemble members are also averaged after the ensemble forecasting step and the average values are considered as the best estimations of the day.

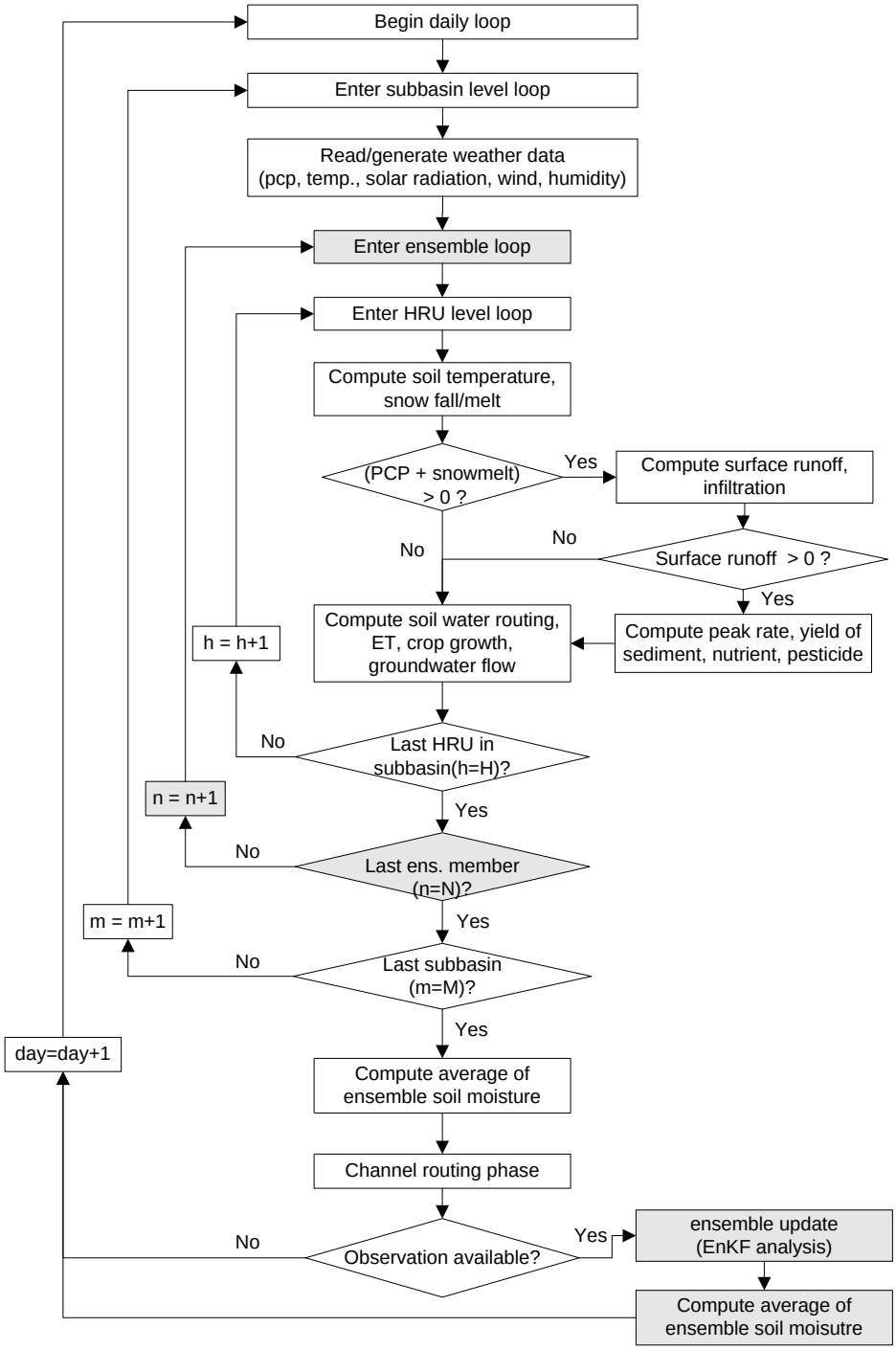


Figure 3.3 Schematic of Surface Soil Moisture Data Assimilation with the SWAT Additional routines due to the EnKF are shown in gray color. H, N and M refer the number of HRUs in a subbasin, total ensemble size (100) and total subbasin number (25) in the UCCW, respectively.

Specification of model and observation errors

At the end of each day, the predicted ensemble soil moisture values are updated through the EnKF analysis step. In the analysis step, model errors (w_k in Equation 3.5) with zero mean Gaussian distribution are added. The standard deviations of the model errors are determined by the current soil moisture content from the true scenario using a scaling factor of 0.01. Observations are also perturbed with the observation errors (v_k in Equation 3.9) which have zero mean Gaussian distribution and standard deviation of 0.01.

Procedures for soil moisture perturbation by adding system variance (model error) are described as follows:

1) Compute a weight (wl_j) for each layer (j) in order to take account for the different thicknesses of each layer (Ryu et al., 2009).
2) Compute model noise (w_k in Equation 3.5) by applying a scaling factor (0.01) and a weight (wl_j) for each layer.

$$w_j = wl_j \cdot 0.01 \cdot \theta_{true,j} \cdot r_j$$

where $\theta_{true,j}$ is soil moisture estimation from the true scenario and r_j is a random number from a Gaussian distribution with zero mean and variance one for $j=1,\ldots,n$ where n is the number of soil layers.

For soil moisture perturbation, a constant scaling factor is applied to all layers every daily time step. Vertical correlation between perturbations is also an important factor in determining how surface soil moisture assimilation affects deeper layers and other dependent variables. Chen et al. (2011) showed the impact of error coupling with the

SWAT model by comparing the results from zero vertical error correlation and perfect vertical error correlation. In the present study, perfect vertical error correlation was applied.

Several past studies have attempted to specify observation and model error statistics. Clark et al. (2008) and Xie and Zhang (2010) used the scaling factor approach for straightforward estimation of model and observation error. Reichle et al. (2002b) assigned temporally constant standard deviations of errors. The scaling factor approach used in this study allows standard deviation of the model errors to vary with time, which is a better representation of the real uncertainties than the time invariant standard deviation of errors. Since the synthetic experiment of this study uses the same model parameters as the true scenario, the source of model error is mainly from errors in the precipitation. Therefore, in this study, a simple but more operational approach of using the scaling factor is adopted for model error estimation. This approach overestimates the real errors when the soil moisture content is high, but is advantageous to test the robustness of the EnKF (Xie and Zhang, 2010). The disadvantage of this approach, however, is that it may enhance the nonlinear impact of the saturation (or wilting point) threshold by inflating (or decreasing) the standard deviation of soil moisture predictions. In spite of convenient application of the scaling factor approach, more sophisticated approaches such as an adaptive filtering seem to be desirable for future real data assimilation studies.

Time invariant standard deviation of observation errors is adopted in this study because errors in remotely-sensed soil moisture come from vegetation cover, surface roughness,

soil properties, radio frequency interference (RFI), and retrieval algorithms, all of which are not proportional to the surface soil moisture condition (Schmugge et al., 2002; Entekhabi et al., 2010). In addition, if the same scaling factor approach is applied to the observation error, it will assign unreasonably high weight on the observation accuracy when the soil is very dry because SWAT defines soil water content excluding the amount of water held at wilting point, with the minimum soil water content in the SWAT being zero.

3.3.3 Additional steps in analysis procedure

The bounded nature of soil moisture between wilting point and saturated soil water content makes the application of the EnKF more complicated. Reichle et al. (2002a) mentioned that in an operational perspective, the violation of its inherent assumption, Gaussian distribution, would have the greatest impact when the soil is very dry and there is high skewed forecast error. Crow and Wood (2003) showed that the skewed model ensembles and a non-Gaussian error structure negatively impacted the EnKF's performances. Because of this unique characteristic of soil moisture, some additional steps are added after the analysis step. When the soil is very dry or wet (close to the wilting point or saturated condition), the best estimates of the soil moisture (average of ensemble) after the analysis step may exceed the actual physical limits of the soil moisture. If the average of the ensemble of the soil moisture becomes negative, the best estimate of the ensemble is adjusted to 10^{-6} mmH_2O. In the case of an ensemble average

higher than the maximum soil moisture limit, saturated soil moisture content minus 10^{-6} mmH_2O replaces the best estimate.

The concept of a simple bias correction method adopted by Ryu et al. (2009) is implemented in this study to take account of the effect of the bounded range of soil moisture. Mean bias is computed using the unperturbed soil moisture prediction and the ensemble is corrected by subtracting the mean bias from the perturbed soil moisture ensemble. After this adjustment, all ensemble members exceeding minimum or maximum soil water content are replaced by 10^{-6} mmH_2O (almost wilting point) or saturated water content. This boundary truncation might shift the average of ensemble again. Therefore, these bias corrections and boundary truncations are repeated 10 times to reduce the remaining bias. Boundary truncation and the simple bias correction approach are applied at each update time step for all state variables (soil moisture vector). A simple boundary truncation might cause the mean of the state variable to be shifted. For example, for a wet soil moisture condition, the simple boundary truncation may shift the mean of soil moisture higher than the actual mean from the analysis step. Therefore, this repetition of boundary truncation and bias correction may be advantageous in generating less mass balance error by decreasing the biases (Ryu et al. 2009).

3.4 Results and Discussion

3.4.1 Effect of Surface Soil Moisture Assimilation on Hydrologic Processes

This section describes how surface soil moisture data assimilation affects subsequent hydrological processes. First, in this study, inaccurate precipitation is the main source of error in soil moisture predictions as well as other hydrological processes. Figure 3.4 shows the daily simulation results of surface (a) and profile (b) soil moisture in the watershed which is the sum of area-weighted soil water content from all HRUs. Soil moisture prediction errors are significant especially during the winter time (from day 353 of the first year to day 67 of the second year) due to the cumulatively overestimated precipitation shown in Figure 3.2b. There exist distinct discrepancies in soil moisture predictions between the true and the open loop scenario.

Soil moisture update through the EnKF pulls the inaccurate soil moisture prediction in the open loop scenario closer to the true state. The correlation coefficient increases from 0.585 to 0.747 and from 0.906 to 0.942 for surface and profile soil moisture respectively (Table 3.3). The RMSE and MBE also decrease in both surface and profile soil moisture predictions with the EnKF. The improved results with the EnKF support the previous studies (Das et al. 2008; Draper et al. 2009; Reichle et al. 2002b; Sabater et al. 2007) further demonstrating the potential of current and forthcoming remotely-sensed surface soil moisture data to improve the profile soil moisture estimations for land surface hydrology through data assimilation. The results of this study also show that the errors in

soil moisture estimation due to inaccurate precipitation can be partially compensated by assimilating surface soil moisture observations.

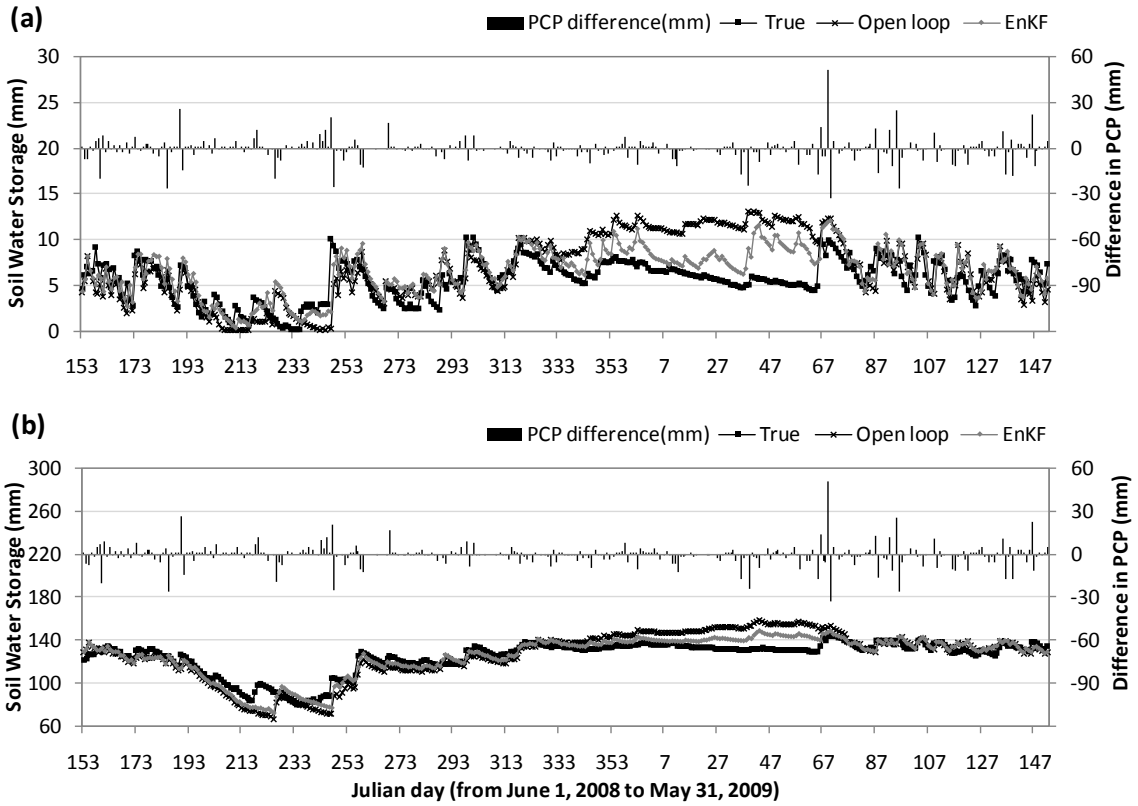


Figure 3.4 Watershed Average Soil Moisture Prediction
(a) Surface (~5cm) and (b) Soil Profile
(Precipitation difference = true precipitation – open loop precipitation)

Temporal variations in the RMSE of the subsequent variables are shown in Figure 3.5 where daily RMSE is based on the errors of all HRUs. It is apparent that reduced errors in surface soil moisture prediction (Figure 3.5a) are clearly identified for almost every time step. In spite of the prompt decrease of errors in surface soil moisture prediction, the magnitude of improvements in the profile soil water content varies with time as shown in Figure 3.5b with reduction in errors with the EnKF being distinct in winter (from day 353

to 67). Limited success in updating profile soil water content for some periods may be caused by the weak model vertical coupling of soil moisture in SWAT subsurface physics (Chen et al., 2011). Another reason is that non-linearities in model physics and the bounded nature of soil moisture result in suboptimal update with the EnKF by violating the Gaussian assumption. When optimality of ensemble perturbation is checked as in Reichle et al. (2002a), non-symmetric ensemble distribution is found with the soil moisture condition close to saturation or wilting points (not shown). In addition, consistently lower forecast and analysis error variances compared with actual error were found, resulting in estimates less than optimal.

Table 3.3 Statistical Results of the Simulated Watershed Average Soil Water Content

	R		RMSE		MBE	
	Open loop	EnKF	Open loop	EnKF	Open loop	EnKF
SOL_SW (5cm)	0.585	0.747	3.214	1.931	-1.275	-0.879
SOL_SW	0.906	0.942	11.176	6.832	-1.533	-0.317
SHALLST	0.929	0.923	0.127	0.097	-0.054	-0.041
DEEPST	0.986	0.995	2.444	2.438	2.444	2.438
INFLPCP	0.339	0.339	3.572	3.572	-0.152	-0.171
QDAY	0.458	0.455	3.059	3.025	-0.452	-0.442
LATQ	0.844	0.860	0.035	0.034	-0.013	-0.014
RCHRG	0.927	0.921	0.219	0.168	-0.092	-0.071
GW_Q	0.929	0.923	0.217	0.165	-0.092	-0.071
GWSEEP	0.930	0.924	0.001	0.001	0.000	0.000
CNDAY	0.976	0.977	2.125	1.617	0.641	0.358
ET	0.696	0.736	0.852	0.849	-0.111	-0.200

The improved soil moisture prediction affects other subsequent hydrological responses. Table 3.3 shows that the EnKF scenario reduces errors in other hydrological variables compared to the open loop scenario even though the correlation coefficient for some

variables may be slightly lower (e.g., SHALLST, QDAY, RCHRG, GW_Q and GWSEEP).

Similar to the profile soil moisture in Figure 3.5b, the depth of water in shallow aquifer (SHALLST) in Figure 3.5c and GW_Q (groundwater contribution to streamflow) in Figure 3.5f are not noticeably affected except during winter conditions. Little differences in RMSE are observed in the prediction of QDAY (surface runoff) in Figure 3.5d and LATQ (lateral flow) in Figure 3.5e. Since accurate prediction of surface runoff (QDAY) and groundwater contribution (GW_Q) are critical to streamflow prediction, it is expected from these results that surface soil moisture assimilation may not improve streamflow prediction significantly. Further discussion of these finding is provided in section 3.4.2. Application of EnKF reduced the RMSE in evapotranspiration prediction throughout the experiment period (Figure 3.5g). This is because evaporation occurs mainly on the soil surface and the improved surface soil moisture condition contributes to greater accuracy in simulating evaporation. Finally, the curve number for a day (CNDAY) is computed based on the profile soil water content. Therefore, the trend of error reduction in the curve number prediction (Figure 3.5h) is very similar to the trend of the profile soil moisture (Figure 3.5b).

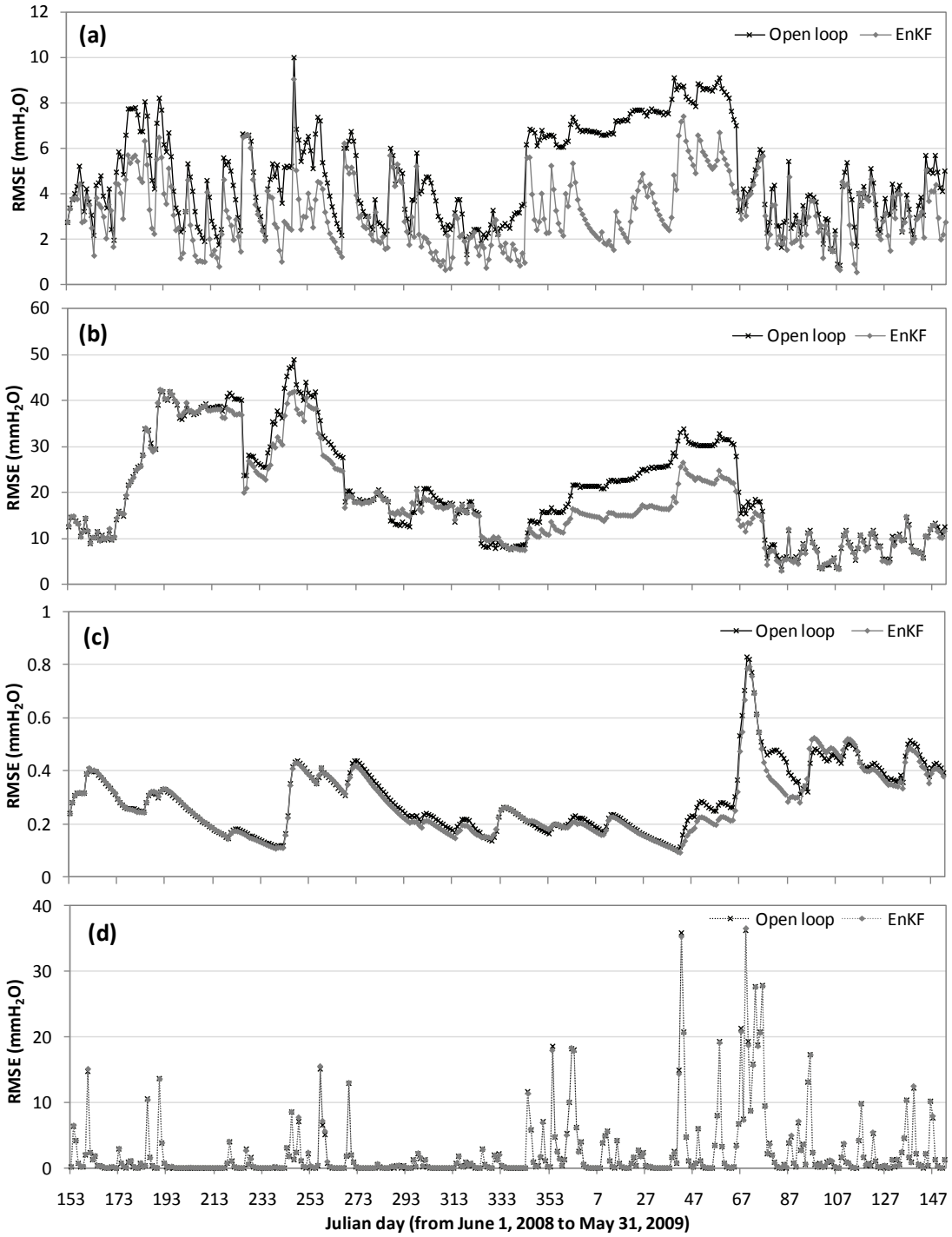


Figure 3.5 RMSE of Subsequent Variables
(a) surface (~5cm) soil water content, (b) profile soil water content, (c) depth of water in shallow aquifer (SHALLST), (d) surface runoff loading to main channel for day in HRU (QDAY), (e) amount of water in lateral flow in HRU for the day (LATQ), (f) groundwater contribution to streamflow from HRU for the day (GW_Q), (g) Evapotranspiration (ET), (h) curve number for current day (CNDAY).

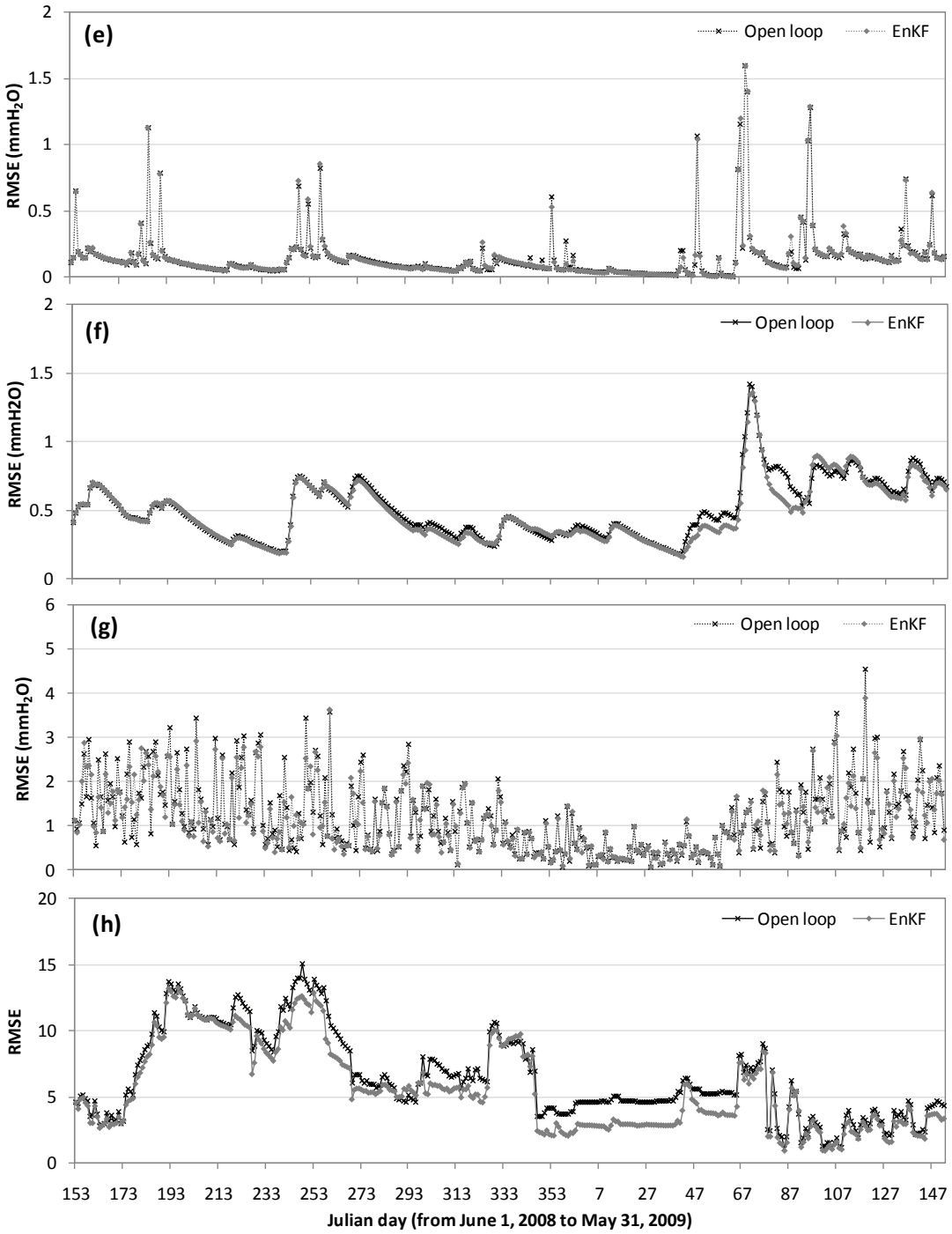


Figure 3.5 Cont. RMSE of Subsequent Variables
(a) surface (~5cm) soil water content, (b) profile soil water content, (c) depth of water in shallow aquifer (SHALLST), (d) surface runoff loading to main channel for day in HRU (QDAY), (e) amount of water in lateral flow in HRU for the day (LATQ), (f) groundwater contribution to streamflow from HRU for the day (GW_Q), (g) Evapotranspiration (ET), (h) curve number for current day (CNDAY).

3.4.2 Streamflow Prediction

The answer to the question "Can surface soil moisture data assimilation improve streamflow prediction?" depends on the accuracy of precipitation input and antecedent soil moisture condition. In order to answer to this question, four simple and different cases can be considered accounting for only antecedent soil moisture conditions and the accuracy of precipitation data, and assuming that the EnKF always moves soil moisture predictions closer to the true values.

1) Model predicted antecedent soil moisture is less than the true soil moisture ($\theta_{predicted} < \theta_{true} \approx \theta_{EnKF}$) and current precipitation is overestimated.

2) Model predicted antecedent soil moisture is greater than the true soil moisture ($\theta_{predicted} > \theta_{true} \approx \theta_{EnKF}$) and current precipitation is overestimated.

3) Model predicted antecedent soil moisture is less than the true soil moisture ($\theta_{predicted} < \theta_{true} \approx \theta_{EnKF}$) and current precipitation is underestimated.

4) Model predicted antecedent soil moisture is greater than the true soil moisture ($\theta_{predicted} > \theta_{true} \approx \theta_{EnKF}$) and current precipitation is underestimated.

Cases 2 and 3 would provide improvement in streamflow prediction by updating soil moisture through the EnKF. In Case 2, the overestimated antecedent soil moisture condition will aggregate the error from the overestimated precipitation and therefore, updated (corrected to lower) soil moisture with the EnKF will generate less error than the open loop. The same principle can be applied to Case 3 where improved runoff prediction

with the EnKF is expected compared to the open loop. However, Cases 1 and 4 will generate more errors with the EnKF than with the open loop. In Case 1, the underestimated antecedent soil moisture condition counterbalances the error in the overestimated precipitation. Therefore, improved runoff prediction occurs with the open loop rather than with the EnKF in this case. The same principle is applied to Case 4. Therefore, a hypothesis, "Assimilating surface soil moisture observation into a hydrologic model will improve streamflow (or surface runoff) prediction" is valid only if I have accurate precipitation information. Crow and Ryu (2009) pointed out this limitation of updating solely the antecedent soil moisture condition and designed an assimilation system that simultaneously updates soil moisture and corrects rainfall input by taking into account soil moisture observations.

Case 1 occurs in the results of this study during the summer, especially around day 259. The EnKF improves soil moisture prediction (Figure 3.4b and Figure 3.5b). However, because of inaccurately overestimated precipitation, improved soil moisture with the EnKF (higher soil moisture prediction than the open loop results) results in overestimated runoff and streamflow (Figure 3.5d and Figure 3.6). The lower model performance using the EnKF is primarily due to inaccurate precipitation. Therefore, in this case, surface soil moisture assimilation does not appear to improve streamflow prediction.

Results during winter time represent Case 2 where both soil moisture and precipitation are overestimated. Application of the EnKF improves streamflow prediction during winter time (Figure 3. 6, day 42 to 87). This is because the open loop continues overestimating soil moisture during winter, but the EnKF improves the soil moisture

prediction (lowers the overestimated soil moisture in Figure 3.4). Even though the improvement in runoff prediction (QDAY) due to the improved soil moisture with the EnKF is not significant (Figure 3.5d), the improved groundwater contribution (GW_Q) also contributes to the improved (reduced) streamflow in Figure 3.5f.

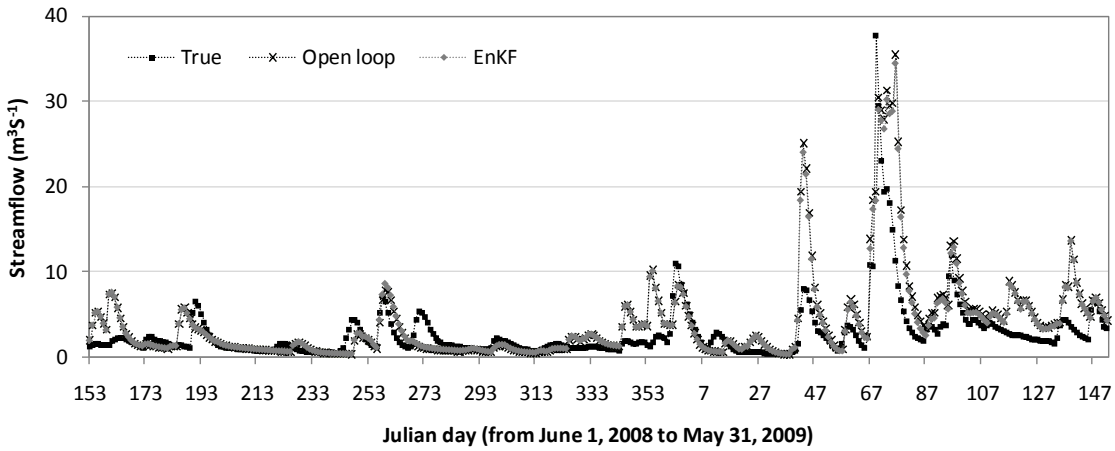


Figure 3.6 Streamflow Prediction

In this study, streamflow prediction does not show significant improvement even after applying surface soil moisture assimilation. The primary reason for this is that most of the errors in streamflow prediction are due to inaccurate precipitation input (Figure 3.6, Figure 3.7 and Table 3.4). Second, SWAT model physics does not have sufficient vertical coupling strength to constrain soil water content in deeper layers or in the root zone using the surface soil moisture observations (Chen et al., 2011). Since surface runoff generation depends on profile soil water content, failure to significantly improve estimates of profile soil water content impede successful surface runoff (or streamflow) prediction. Therefore, improvement in streamflow prediction is not expected when the improvement of profile soil water content with the EnKF is of little or marginal consequence as shown in Figure

3.5b. Lastly, the unsuccessful improvement in streamflow prediction may be attributed to the limited sensitivity of surface runoff (or streamflow) prediction to the change in soil moisture with the CN method implemented in SWAT. Surface runoff is the main contributor to streamflow especially with high rainfall intensity. Therefore, if the updated soil water content with the EnKF is not reflected properly in the surface runoff estimation, improved streamflow prediction cannot be expected.

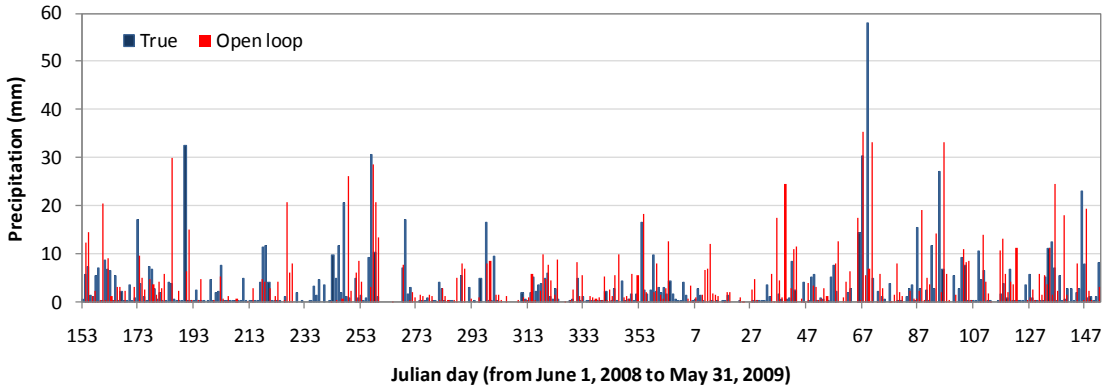


Figure 3.7 Area Weighted Sum of Daily Precipitation

Table 3.4 Errors in Precipitation and Outflow of the Watershed during Simulation Period

	True	Open loop	EnKF
Input precipitation (mm year^{-1})	937.02	1175.37	1175.37
Outflow (mm year^{-1})	438.47 (47%)	641.94 (55%)	630.85 (54%)
Error in precipitation (mm year^{-1})	-	-238.36	-238.36
Error in streamflow (mm year^{-1})	-	-203.46	-192.37

In order to test model sensitivity in these regards, the relationship between surface runoff generation and different soil moisture conditions for various curve numbers is illustrated in Figure 3.8 based on Equations 3.2, 3.3 and 3.4. The soil type, GnB2, of which water storage at field capacity and saturation are 101 mm and 201 mm respectively for a soil

depth of 1246 mm, is used to create Figure 3.8. In the case of high curve numbers (CN_2=93), a 20 mm reduction in soil moisture from 200 to 180 mm results in only 2 mm reduction in surface runoff. This decrease in runoff prediction will result in a 4.57 m^3sec^{-1} streamflow decrease, which is not sufficient to overcome the significant streamflow overestimation during the winter in Figure 3.6.

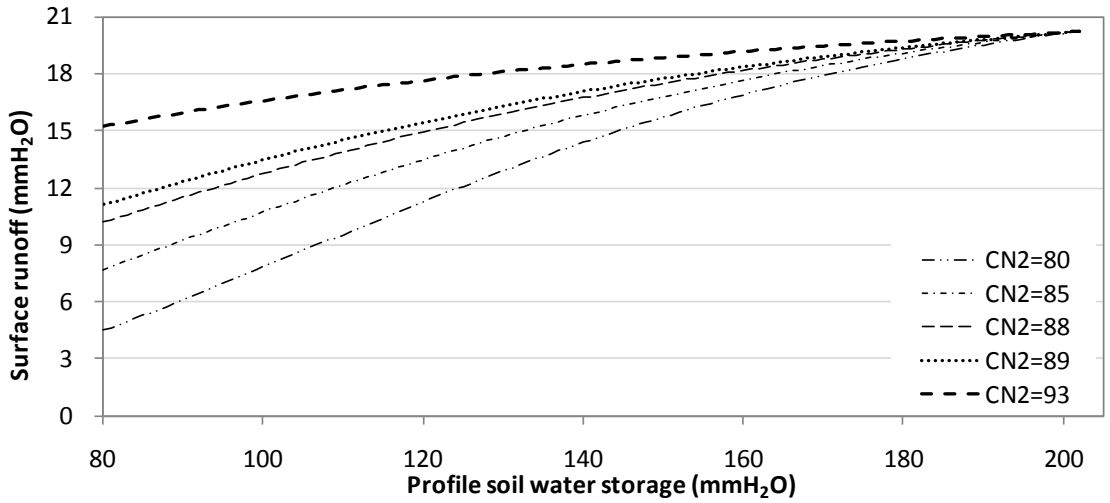


Figure 3.8 Relationship between Surface Runoff and Soil Water Condition with Various CN2 Numbers in Curve Number Method When Rainfall is 23 mm/day.

As Figure 3.8 shows, the sensitivity of surface runoff to soil water content varies with the curve number, which is a function of soil type and landuse. This study also shows that the degree of improvement in surface runoff prediction with the EnKF is highly related to soil type and landuse. The areas which have low runoff error with the EnKF consist of soil types BoB, Hw, SrB2, StC3 and Se (Table 3.1). Hydrologic soil groups of those soil types are A or B (Table 3.1). Furthermore, the combinations of those soil types and landuse (HAY and FRSD) provide the smallest errors because of their hydrologic soil group (A or B) and low CN_2 (35 ~ 60). On the contrary, the main three soil types (BaB2,

GnB2 and Pe) which cover 67% of the area in the UCCW produce high errors in surface runoff prediction regardless of landuse type. Their CN_2 is relatively high (77 ~ 89). For most of the areas in UCCW the EnKF, therefore, cannot decrease the error in the runoff prediction because surface runoff is not very sensitive to the soil moisture change with high CN_2 values. That is, the slope of the graph is small for high CN_2 values in Figure 3.8. Especially when soil moisture is high (close to saturation, 200mm in the case of GnB2), the slope is very small. During the winter time, highest soil moisture estimation errors exist and the EnKF improves profile soil moisture up to 20 mm in Figure 3.4b, but in this period, the soil moisture condition is very wet, so its improvement is difficult to reflect in surface runoff.

In this study, surface soil moisture assimilation was minimally successful in improving streamflow predictions. However, if I consider only streamflow results, traditional calibration methods can improve streamflow simulation considerably, even in the presence of inaccurate precipitation. In SWAT, streamflow prediction has a higher sensitivity to changes in certain parameters rather than changing soil moisture conditions. Parameter adjustments through the calibration can change streamflow prediction effectively by increasing the E_{NS}. However, conventional calibration methods do not directly take into account the uncertainties in precipitation or model structure, or the possibility of model parameters that change temporally. Data assimilation techniques usually focus on updating state variables (soil moisture in this study) on the premise that model parameters are pre-specified. Therefore, new frameworks which can address both the limitations of conventional parameter calibrations and data assimilation are required

for future advances in hydrologic modeling. Moradkhani et al. (2005), for example, showed the possibility of estimating both model states and parameters simultaneously using a dual state-parameter estimation method based on the EnKF.

3.4.3 CN Method vs. Green Ampt Method

Even though the CN method has been widely used in various hydrologic models, it has limitations. Garen and Moore (2005) stresses some issues with the broad use of the CN method. Since the CN method was developed as an event model for the prediction of flood streamflow conditions, they argue that it is questionable to apply the CN method for a continuous model and daily flow of ordinary magnitude. In addition, they also mention that daily time step might not be appropriate to simulate the infiltration excess mechanism which is designed for hourly (or subhourly) time steps.

SWAT 2005 provides another option to estimate surface runoff, the Green Ampt Mein-Larson excess rainfall method (GAML). This method is based on the Green & Ampt infiltration method and requires subdaily precipitation input. The CN method has been more widely used than the GAML method because of the difficulties in obtaining subdaily precipitation data and the uncertain benefit of using the GAML compared to the CN method. While Kannan et al. (2007) and King et al. (1999) found no significant advantages of using the GAML instead of the CN method with SWAT, Jeong et al. (2010) showed that the GAML outperforms the CN method and suggested that the quality of subdaily precipitation data and the size of study area influence the results. In addition,

Jeong et al. (2010) also showed that even the subhourly simulation performs poorly under low or medium flows.

In this study, a simple experiment is conducted to test if the GAML method has a greater potential than the CN method for improving runoff (and eventually streamflow) prediction with surface soil moisture data assimilation. As mentioned in section 3.4.2, sensitivity of surface runoff generation to the soil moisture change determines how successfully the soil moisture update through the EnKF will contribute to the improvement of the surface runoff prediction. For the two main soil types, GnB2 and SrB2, surface runoff was computed by using the CN method and the GAML method for one day (June 21, 2008) when precipitation was 23mm day^{-1}. For the GAML method, 10 minute precipitation data were used. The amount of water held in soil profile at field capacity is 101.5 and 240 mmH_2O for GnB2 and SrB2, respectively. Water content at saturation is 210 and 400 mmH_2O for GnB2 and SrB2, respectively.

Figure 3.9 shows the different runoff-soil moisture relationships from the CN and GAML methods. The CN method maintains a smooth relationship between runoff and soil moisture, however, the GAML method produces drastic changes when soil moisture is near field capacity (101.5 and 240 mmH_2O for GnB2 and SrB2 respectively). This is because the GAML code implemented in SWAT2005 replaces any soil moisture value that is greater than field capacity soil moisture with the field capacity value in computing infiltration rate. Therefore, under higher soil moisture conditions, the GAML method produces much less surface runoff than the CN method, which is appropriate in reducing the overestimated streamflow. However, the slope of the graph for GAML is much

smaller than the slope of the CN method above field capacity. Therefore, the impact of the improved soil moisture with the EnKF might be difficult to see with the GAML method when the soil moisture is above field capacity. However, soil moisture less than the field capacity has a high sensitivity to changes in soil moisture. The two different soil types show different shapes of the relationship because of their different field capacity and saturated water content values. Therefore, the effectiveness of the soil moisture assimilation with the GAML depends on the soil type and soil moisture condition. In using either the CN method or GAML method, it is very difficult and somewhat complicated to determine precisely how changes in soil moisture affect streamflow prediction.

Successful improvements of streamflow prediction through soil moisture data assimilation can be expected only if the model is based on correct linkage between soil moisture conditions and surface runoff generation. Even though SWAT 2005 adopts a modified CN method which accounts for antecedent soil moisture condition continuously, its effectiveness is not thoroughly proven. In addition, application of the CN method has been questioned by many studies and its modification for runoff simulation has been introduced (Michel et al., 2005; Mishra and Singh, 2006; Chung et al., 2010; Sahu et al., 2010), specifically with SWAT (Kannan et al., 2008; Kim and Lee, 2008; Wang et al., 2008; White et al., 2010). In addition, various factors such as quality of rainfall data, rainfall intensity, soil type, landuse and identification of critical source area make streamflow prediction more complicated. Therefore, more careful selection and development of runoff simulation procedures effectively taking account of soil moisture

variations should be required to enhance streamflow prediction with future soil moisture assimilation studies.

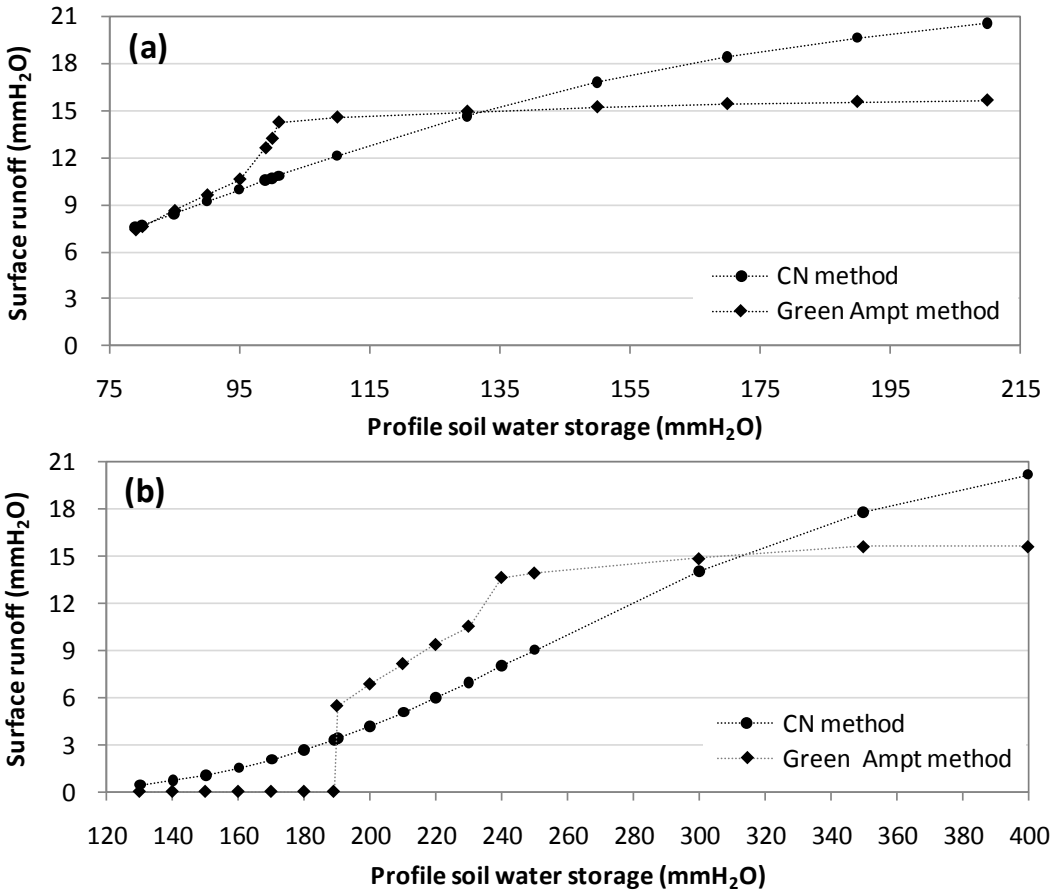


Figure 3.9 Comparison of CN Method and Green Ampt Method for the Relationship between Surface Runoff and Soil Water Condition: (a) soil type: GnB (b) soil type: SrB.

3.4.4 Spatial Variation in Soil Moisture Prediction

In this section, the impact of spatially varying input, specifically landuse and soil type, on surface soil moisture assimilation is illustrated. To visualize the spatial distribution of output, an HRU map is created by overlaying landuse and the SSURGO soil map generated from ArcSWAT. Since a single slope is defined in the initial SWAT setup, the slope is not taken into account in creating the HRU map. HRU level outputs are assigned to each HRU to show spatially distributed results.

Precipitation as a forcing variable is the most important factor for successful soil moisture estimation. Figure 3.10 shows time-averaged RMSE of precipitation. Time-averaged RMSE results of surface and profile soil moisture prediction are shown in Figure 3.11 and 3.12 respectively. The RMSE distribution of the open loop and the EnKF, especially for profile soil moisture distribution, coincides with the distribution of precipitation in general. That is to say, higher precipitation errors within an area result in greater errors in soil moisture estimation.

The EnKF reduces errors in surface and profile soil moisture compared to the results of the open loop. The average RMSE errors in surface soil moisture estimates in Figure 3.11 are 5.05 and 3.43 mmH_2O for the open loop and the EnKF, respectively. For profile soil moisture estimation, the open loop and the EnKF have 22.77 and 19.77 mmH_2O of average RMSE, respectively in Figure 3.12.

Within a subbasin where constant precipitation is assigned, types of landuse and soil determine the magnitude of errors in the soil moisture estimation. One distinct example is

shown in subbasin 2 in Figure 3.11a. Even though the subbasin has the same amount of precipitation throughout the area, some of the areas (light blue in the Figure 3.11a) have highly underestimated surface soil moisture (much higher RMSE) than others. The areas consist of HRU 11(FRSD, Hw), 18(AGRR, Hw) and 21(WETF, GnB2). The soil type Hw is classified as hydrologic soil group A (Table 3.1) which has a high infiltration rate.

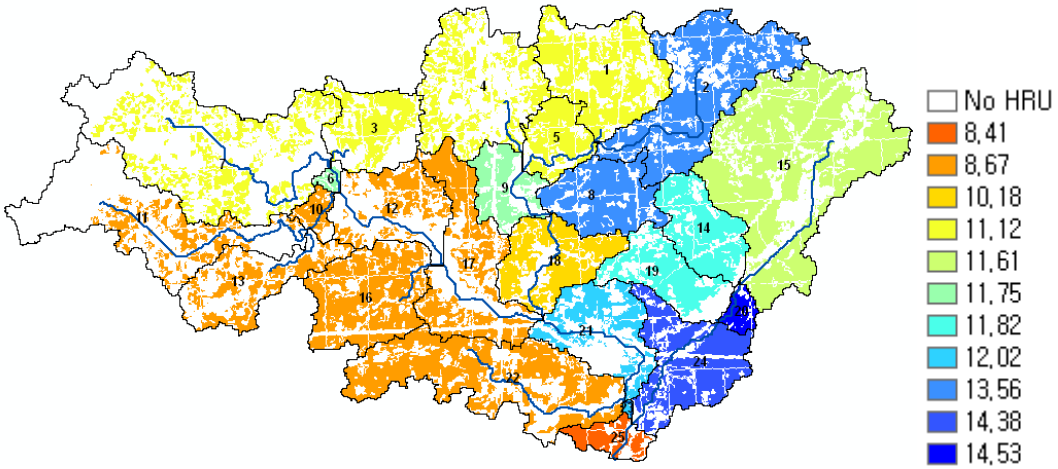


Figure 3.10 Time Averaged RMSE of Precipitation
(The numbers indicate subbasin number and unit is mm)

Figure 3.13 compares the results of different combinations of landuse and soil types. Figure 3.13b is the reference because it consists of the major landuse (AGRR) and soil type (GnB2). Comparing Figure 3.13b and 3.13c explains the effect of different soil types on soil moisture estimation. Although landuse type and precipitation are same, Figure 3.13c which has more infiltration because of the soil type Hw, exaggerates both the underestimated and overestimated errors compared to the soil type GnB2 in Figure 3.13b.

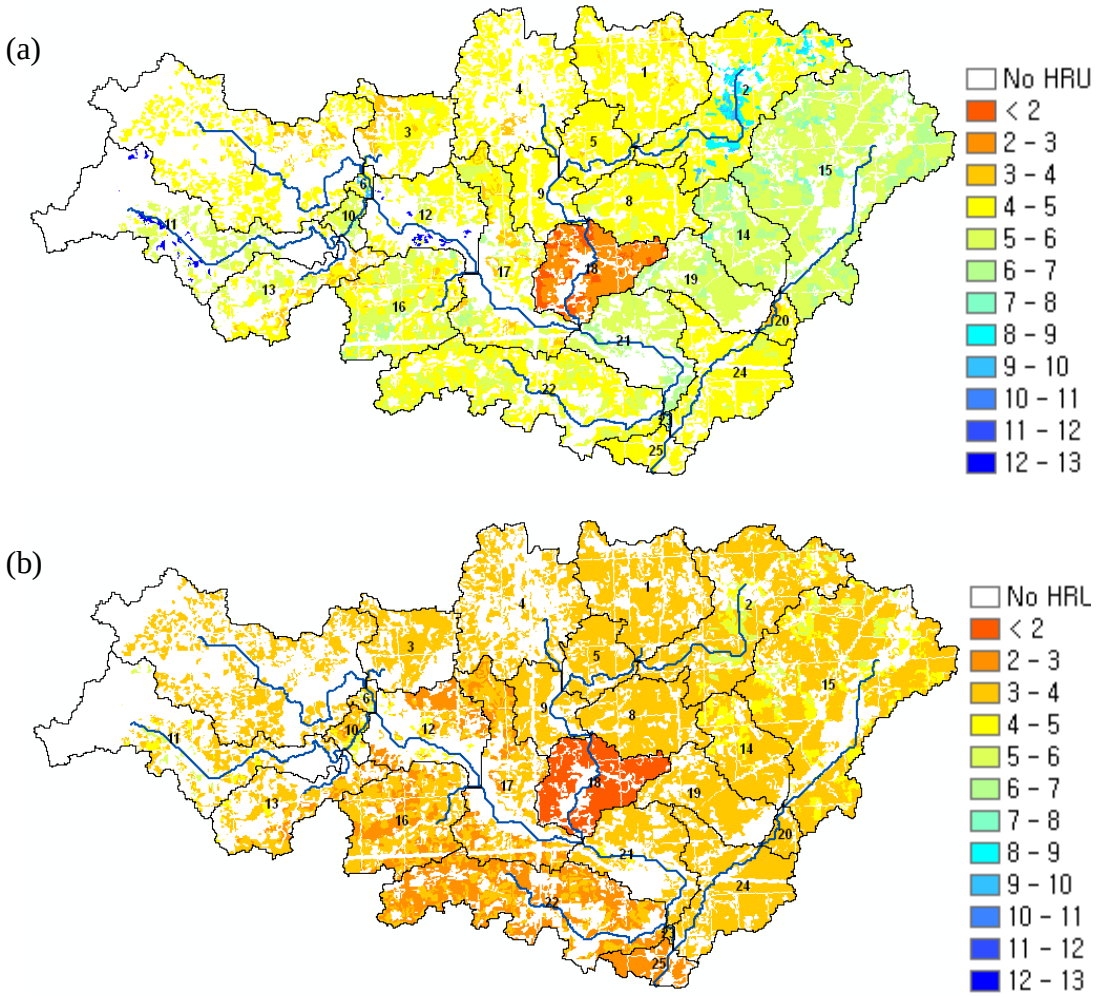


Figure 3.11 Time Averaged RMSE of Surface (~5cm) Soil Moisture
from (a) Open loop and (b) EnKF
(The numbers on the map indicate subbasin number and unit is mmH_2O)

Types of landuse also affect the soil moisture variations. Figure 3.13a and 3.13d show that landuse types, forest (FRSD) and wetland (WETF), produce greater errors than agricultural area. The proportion of precipitation intercepted by canopy is large in the forest area. Therefore, the actual precipitation that reaches the ground (after canopy interception) in this area may be much smaller than other landuse areas, which will

exaggerate errors in soil moisture estimation. Further explanation regarding the errors due to different landuse types follows in section 3.4.5.

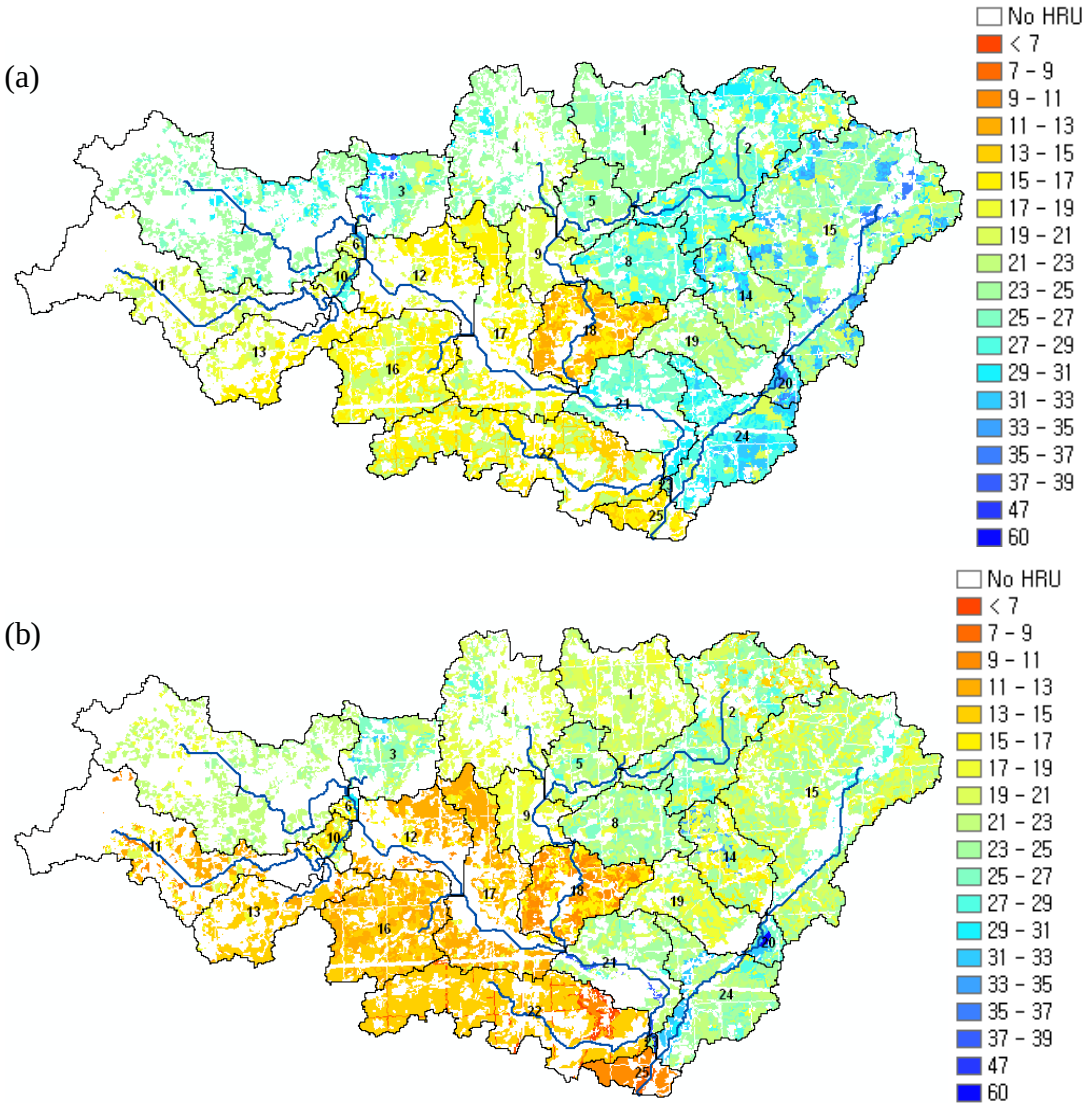


Figure 3.12 Time Averaged RMSE of Profile Soil Moisture
from (a) Open loop and (b) EnKF.
(The numbers on the map indicate subbasin number and unit is mmH_2O)

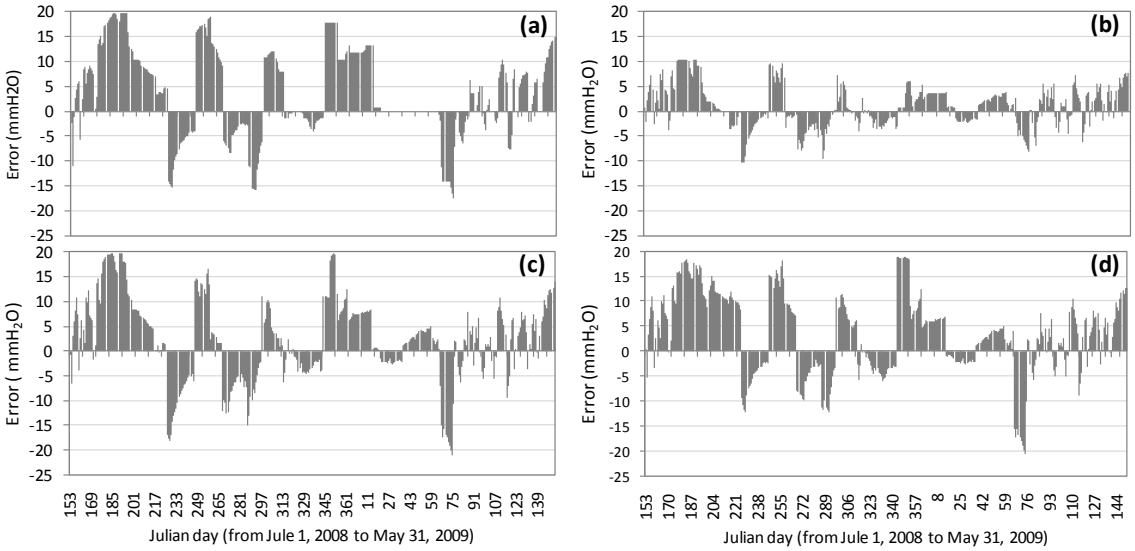


Figure 3.13 Surface Soil Moisture Estimation Error of Open Loop Scenario (error = true – open loop): (a) HRU 11 (FRSD, Hw), (b) HRU 17 (AGRR, GnB2), (c) HRU 18 (AGRR, Hw), (d) HRU 21 (WETF, GnB2).

3.4.5 Spatiotemporal Variation in Soil Moisture Prediction

Aforementioned results show significant effects of errors in precipitation on hydrologic responses. In this section, the effect of precipitation is excluded by selecting a drydown period when no precipitation is received throughout the watershed. Therefore, the effectiveness of the EnKF can be further evaluated.

Figure 3.14c and 3.14d show the antecedent surface soil moisture condition in terms of deviations (errors) from the true condition on day 258 before the precipitation events on day 259 shown in Figure 3.14a and 3.14b. The western areas of the UCCW have overestimated soil moisture (negative error) and eastern areas have underestimated states (positive error) on day 258. On day 259 in Figure 3.14e and 3.14f, highly overestimated precipitation in subbasins 1, 3, 4, 5, 6 and 7 do not make much differences in soil moisture status because soil moisture is already overestimated (close to saturated

condition) and the overestimated precipitation becomes surface runoff instead of infiltrating into the soil and increasing soil moisture. However, slightly overestimated precipitation in subbasins 2, 8, 14, 15, 18, 19, 20, 21 and 24 infiltrates into soil and decreases the errors in the previously underestimated soil moisture.

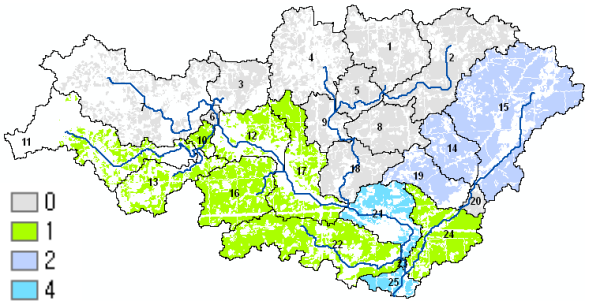
After the precipitation on day 259, there is no precipitation at all throughout the watershed until day 263. While errors in surface soil moisture estimation in the open loop do not change significantly day by day, the EnKF results show noticeably decreasing errors in one or two days. In other words, errors in the EnKF results become close to zero sooner than the open loop results. In regard to the profile soil moisture variations, the EnKF produces better results (smaller errors in general) than the open loop results even though the magnitudes of improvements are not as much as the ones with the surface soil moisture (not shown).

Interestingly, some areas in the open loop give soil moisture results that remain highly underestimated (high positive errors with blue color) even after the precipitation on day 259 (Figure 3.14g, i and k). Those areas correspond to certain distributions of landuse type, such as forest (FRSD) and hay (HAY). A large portion of initial precipitation is intercepted by canopy in this area and thus, actual precipitation generating surface runoff and infiltration on the ground becomes very small. Table 3.5 shows the amount of initial precipitation that is intercepted by the leaves of plants; less than 20% of initial precipitation reaches the ground surface in the forest area in the case of the open loop and the EnKF. The amount of canopy interception is computed based on the values of leaf area index, amount of water held in canopy storage and potential leaf area index. The

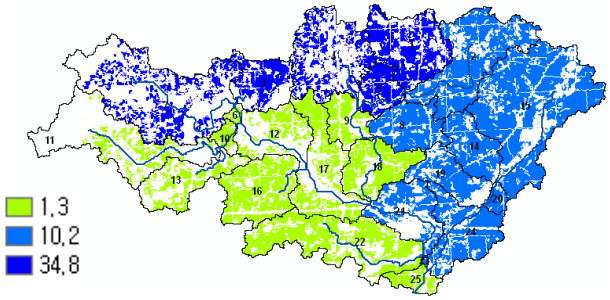
difference in the results between the open loop and the EnKF arise from the differences in the simulation results of those variables. In SWAT, those variables related to crop growth are, therefore, very important in simulating surface runoff and subsurface flow as well as soil moisture because actual precipitation being a forcing variable is affected greatly by the canopy interception.

Contrary to the open loop, surface soil moisture assimilation through the EnKF prevents this long-lasting underestimation for certain landuse types (Figure 3.14h, j and l). The results of profile soil moisture estimation in the open loop scenario also show the same problem, high underestimation with a specific landuse type (not shown). These long-lasting underestimated (dry) soil moisture conditions will affect simulation of crop growth and other hydrologic processes such as plant water uptake and subsurface flow in the long term. The errors due to specific landuse types can be minimized to a certain degree by optimizing parameters through calibration. However, considering that most calibrations for a distributed model are achieved by focusing on the streamflow and overall water balance, it is difficult to overcome this issue for specific soil types. However, the soil moisture update through the EnKF enables us to overcome this problem and to prevent further accumulation of errors in hydrologic predictions by improving soil moisture estimates.

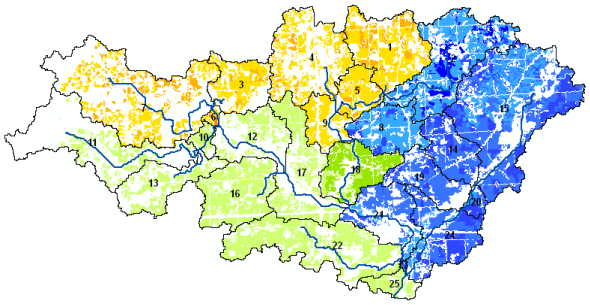
(a) True precipitation on day 259



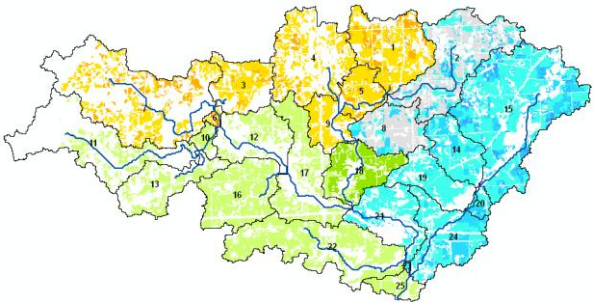
(b) Open loop precipitation on day 259



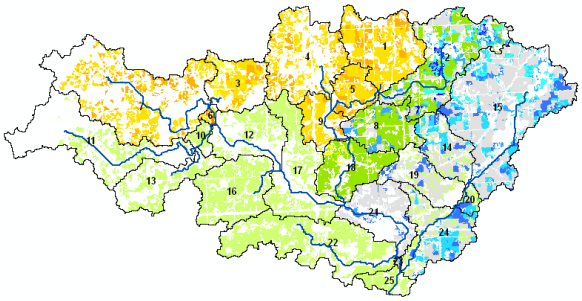
(c) DAY=258, Open loop



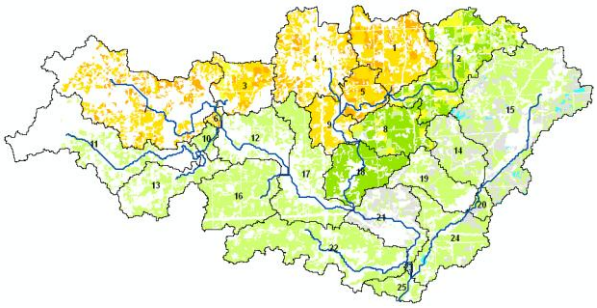
(d)D AY=258, EnKF



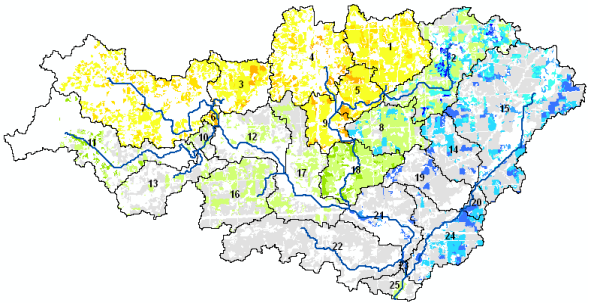
(e) DAY=259, Open loop (after precipitation)



(f) DAY=259, EnKF (after precipitation)



(g) DAY=260, Open loop



(h) DAY=260, EnKF

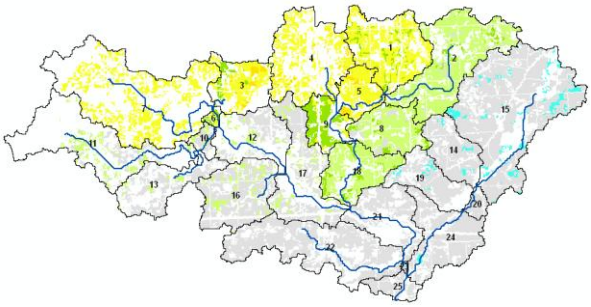


Figure 3.14 Surface Soil Moisture Error Variation during Drydown Period
Precipitation ((a) and (b)) in mm and surface soil moisture error variation ((c) ~ (l))
(error = true – open loop or EnKF).in mmH_2O

(i) DAY=261, Open loop

(j) DAY=261, EnKF

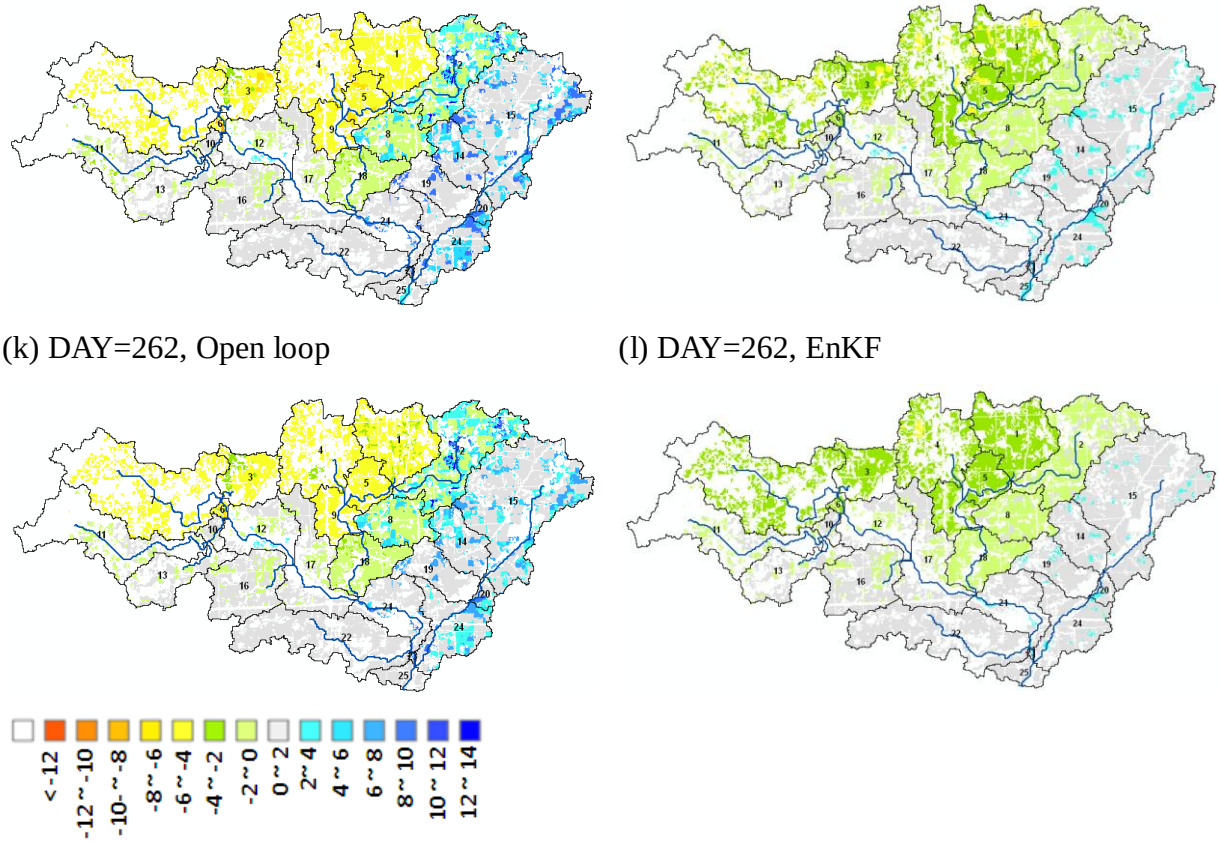


Figure 3.14 Cont. Surface Soil Moisture Error Variation during Drydown Period Precipitation ((a) and (b)) in mm and surface soil moisture error variation ((c) ~ (l)) (error = true – open loop or EnKF).in mmH_2O

Table 3.5 Effect of Canopy Interception (DAY29)

HRU (Landuse)	Initial precipitation		Precipitation after canopy interception		
	True	Open loop/EnKF	True	Open loop	EnKF
118,119 (FRSD)	2.4	10.2	0	1.83	1.83
115,116,117 (AGRR)			1.989	10.151	10.08

3.5 Conclusion

In this study, a synthetic experiment was conducted to investigate how surface soil moisture assimilation affects hydrological processes in the Upper Cedar Creek Watershed using the SWAT hydrologic model and the Ensemble Kalman Filter. This study compared three scenarios: 1) a true scenario with no errors in model, precipitation and soil moisture observations; 2) an open loop scenario with limited precipitation information; and 3) the EnKF scenario with the same imperfect information as the open loop but assimilating observed surface (~5cm) soil moisture every day using the EnKF data assimilation technique.

Soil moisture update through the EnKF improved surface and profile soil moisture estimations compared to the open loop. In addition, the EnkF improved predictions of other subsequent hydrological variables with reduced errors even though the magnitude of the improvements varied according to different variables and rainfall accuracy. The most significant improvements in the prediction of those variables were found during the winter months due to large errors in cumulatively overestimated precipitation.

The capability of surface soil moisture assimilation for improving streamflow prediction was constrained by the accuracy of precipitation and characteristics of the rainfall-runoff mechanism in SWAT. Highly overestimated or underestimated streamflow peaks corresponded to the inaccurate rainfall events. Updated soil moisture conditions after applying the EnKF do not always lead to improvements in surface runoff (or streamflow) predictions. This is because in using the curve number method, sensitivity of surface

runoff generation to changes in soil moisture is affected by various factors such as soil type, landuse type, rainfall intensity, CN_2 and antecedent soil moisture conditions. A simple experiment in this study showed that both the CN method and the Green-Ampt method in the SWAT have some limitations in reflecting soil moisture updates into the surface runoff routine. Therefore, it is essential to have accurate precipitation input to achieve improved streamflow predictions through surface soil moisture assimilation using SWAT. This fact was the motivation behind several previous studies for exploring methods to improve precipitation input using soil moisture retrievals (Crow and Bolten, 2007; Crow et al., 2009; Crow et al., 2011). Second, greater effort is needed in developing better excess rainfall simulation algorithms to overcome the limitations of the current CN method so that continuous model processes under various rainfall intensities are taken into account.

Distributed RMSE maps from the surface and profile soil moisture predictions illustrated the effects of different landuse and soil types on soil moisture estimation. The EnKF results in much less error in soil moisture estimation than the open loop scenario throughout the watershed. However, depending on the soil and landuse type, the magnitude of reduced errors with the EnKF varied. When I removed the effect of inaccurate precipitation and examined a certain drydown period, the EnKF scenario performed much better than the open loop scenario. The EnKF returned the soil moisture condition close to the true state more quickly than the open loop scenario reducing the prediction errors that resulted from inaccurate or limited precipitation data. In addition, the EnKF was shown to overcome the problem of consistently underestimated soil

moisture in some areas compared with the open loop due to a certain soil type and landuse.

The synthetic data modeling experiment conducted in this study assumed that surface soil moisture observations for each HRU in the SWAT model were available and that model/observational errors were known. However, for real world applications, further studies are required to answer the following questions. 1) How can we best use remotely-sensed soil moisture observations in coarse resolution in time and space for a watershed scale hydrologic model? This question will lead to more studies on developing temporal and spatial downscaling methods (Kaheil et al., 2008; Merlin et al., 2010). 2) How can we determine the uncertainties in precipitation, models and observations? Various approaches have been presented in previous studies with the development of data assimilation techniques, some of which being mentioned in this paper. Although this study demonstrates the potential of remotely-sensed surface soil moisture measurements and data assimilation for applications of watershed scale water resources management, future studies using actual observed data are necessary to effectively transfer the science to practical applications.

CHAPTER 4 ESTIMATION OF REPRESENTATIVE SOIL MOISTURE CHARACTERISTICS AT FIELD SCALE

4.1 Introduction

Soil moisture is a key variable in understanding the hydrologic processes and energy fluxes at the land surface. In spite of new technologies for *in situ* soil moisture measurements, and increased availability of remotely-sensed soil moisture data, scaling issues between soil moisture observations and model grid sizes pose a problem for full utilization of the available data (Western et al., 2002; Robinson et al., 2008; Verstraeten et al., 2008). For example, the scale of point measurements (support area less than 1 m^2) typically does not match the size of model grids (support area greater than 100 m^2). In addition, proper linkage of soil moisture estimates across different scales of observations and model predictions is essential for the validation of current and upcoming space-borne surface soil moisture retrievals, as well as the successful application of data assimilation techniques (Grayson and Western, 1998; Crow et al., 2005; Cosh et al., 2006; Crow and van den Berg, 2010; Jackson et al., 2010b; Miralles et al., 2010; Loew and Schlenz, 2011).

Efforts to obtain the spatially representative soil moisture mean and variance from a few sampling points were initiated by Vachaud et al. (1985) using the concept of temporal

stability. Temporal stability studies have gained considerable attention in the context of validation of remotely-sensed soil moisture products and have been applied to a range of spatial scales, different measurement depths and various time scales (Mohanty and Skaggs, 2001; Cosh et al., 2004; Cosh et al., 2006; Starks et al., 2006; Choi and Jacobs, 2008; Heathman et al., 2009; Joshi et al., 2011). Temporal stability of soil moisture conditions is affected by many factors including topography, soil texture, vegetation and precipitation (Grayson and Western, 1998; Mohanty and Skaggs, 2001; Joshi et al., 2011). Therefore, it is uncertain as to whether a temporally stable location maintains its rank across time and space or if the rank of surface soil moisture measurements is consistent with time-stable observations at deeper depths in the profile.

Grayson and Western (1998) introduced an offset method to estimate spatial soil moisture average from the most temporally stable site having a non-zero relative mean difference. Even though the most temporally stable site can be determined, in reality, there are cases where it might not be feasible to install a permanent soil moisture sensor at that location for long-term monitoring. For example, in agricultural fields such as the study area in this paper, the permanent soil moisture sensors are installed on the boundary of a field in order not to impede landowner agricultural practices such as crop spraying and tillage. Therefore, it is necessary to investigate how well soil moisture observations from the permanent sensors represent field average conditions and how to transform the data to obtain representative information of the field soil moisture variability.

De Lannoy et al. (2007b) explored several statistical methods to upscale point measurements to field average soil moisture. They concluded that cumulative distribution

function (CDF) matching and a simple linear relationship were the most successful upscaling methods to obtain the field average. They also identified converting equations or parameters from those methods to scale up observations from a specific point to the field average. The converting equations or parameters were termed observation operators according to Drusch et al. (2005a). However, the work of De Lannoy et al. (2007) has some limitations for practical applications because they did not test if the observation operators can be transferrable in time or to other similar or adjacent sites.

Variance of soil moisture within an area, as well as the spatial average, are important statistical metrics for soil moisture data assimilation in providing observation error parameters, land surface modeling and validation of remotely-sensed soil moisture products (Charpentier and Groffman, 1992). The relationship between variance and mean has been investigated in several studies showing that the trend of the relationship may differ depending on the particular study (Famiglietti et al., 1999; Brocca et al., 2007). As Vachaud et al. (1985) suggested, it would be very attractive if we could determine variances of soil moisture within an area by measuring soil moisture only at certain locations without dense sampling.

This study aims to link two different scales of soil moisture estimates by upscaling single point soil moisture measurements to field average for representing field-scale agricultural watersheds (~ 2 ha) located within the Upper Cedar Creek Watershed (UCCW) in northeastern Indiana. Estimation of field average soil moisture from point measurements can be utilized for calibration and validation of field-scale hydrologic models such as the Root Zone Water Quality Model. Since field-scale deterministic models can incorporate

complicated soil-water-plant interaction processes from laboratory and field studies, they have provided a fundamental basis for watershed-scale hydrologic, water quality and agricultural system models (Abrahamson et al., 2006). In addition, several previous studies have shown great potential in using surface soil moisture data assimilation with one-dimensional hydrologic models at the field scale to provide better predictions of profile soil moisture (Walker et al., 2001b; Walker et al., 2001a; Heathman et al., 2003a; Han et al., 2011). Therefore, this upscaling approach is a prerequisite for successful data assimilation applications at the field scale by providing more accurate, less biased field representative soil moisture observations.

A couple of methods are investigated, in this study, to induce the soil moisture field average and variance from single point measurements, especially focusing on the CDF matching method which has been shown to be an effective observation operator (Reichle and Koster, 2004a; Drusch et al., 2005a). The main objectives of this study are: 1) to find appropriate way(s) for upscaling point soil moisture measurements to field averages by comparing different observation operators, 2) to explore whether the observation operator from CDF matching is transferable in time and space (between different fields and between different soil layers), 3) to find the most influential factors affecting the temporal or spatial transferability of the observation operators, and 4) to estimate spatial variability (standard deviation) of soil moisture within the study areas using the same point measurements and CDF matching method.

This study provides an alternate means for overcoming the scale differences between *in situ* point scale soil moisture measurements and coarse resolution remote sensing

observations. The two field study sites in this research work, AS1 and AS2, have permanent soil moisture sensors installed at the outlet of the field scale watershed. The permanent sensors are part of the soil moisture monitoring network in the UCCW which will be included in the validation program for soil moisture products from the upcoming (2014) NASA Soil Moisture Active/Passive Mission (SMAP). Therefore, it is expected that this study can be a first step to resolve the scaling issue between *in situ* point measurements and the remote sensing grid size and for successful calibration and validation of SMAP and similar space-borne operational products.

In the next section, limitations in applying temporal stability analysis as a method to obtain spatially representative values are discussed based on previous studies by Heathman et al. (*in press*) and Heathman et al. (*in review*) which used the same dataset from the same field sites as the present study. By addressing the limitations of temporal stability analysis, which has been widely applied in literature, this study aims to advance our ability to more adequately obtain spatially representative soil moisture estimates.

4.2 Related Work – Temporal Stability Analysis

The work in this investigation is an extension of two previous studies that involved a comprehensive analysis of temporal stability conducted in two agricultural fields (AS1 and AS2) to determine whether measurements of soil moisture data from permanent sensor locations (edge-of-field) could be used to adequately represent actual field averages that were obtained from a number of temporary "in-field" sensors (Heathman et al., *in press;* Heathman et al., *in review*). Detailed descriptions of the study areas and

datasets are provided in section 4.3. The previous studies found that sensor locations do not necessarily maintain their rank in mean relative difference (MRD) for different time periods and depths. Nonetheless, most sensors consistently overestimated (positive MRD) or underestimated (negative MRD) the field averages for two different years with some exceptions. In regards to consistency of rank in MRD between two different depths (5cm vs. 20cm), the majority of sensor locations in field AS2 maintained their consistency in having positive or negative MRD. However, five out of eleven sensors at AS1 had opposite signs of MRD, which means that they overestimated the field average at surface and underestimated it at the 20 cm depth, or vice versa.

According to the results in the two previous studies there were no single measurement locations found to represent field average soil moisture conditions from one year to the next or for two different soil layers at both AS1 and AS2. In addition, the most time-stable sites do not remain the same for different years (e.g. #17 in 2009 and #8 in 2010 for 5cm data at AS1) and different depths (e.g. #8 for 5cm and #19 for 20cm at AS1 in 2010) as shown in Table 4.1. Several other studies investigated temporal stability at multiple soil depths found it difficult to determine one representative location for all soil layers (Starks et al., 2006; De Lannoy et al., 2007b; Heathman et al., 2009). Starks et al. (2006) compared the results of temporal stability analysis of soil moisture observations from two different remote sensing campaigns in the Little Washita river watershed, the Southern Great Plains 1997 (SGP97) and the Soil Moisture Experiment 2003 (SMEX03). They found that no single site represented the watershed average soil moisture consistently for any depth interval during the combined experiment periods.

However, for the two agricultural fields in the present study, if we consider that one single location could be selected for practical purposes, then the second most temporally stable site may serve as a substitute for the primary site provided there is little statistical difference between the two sites. Then, #19 and #21 can be determined as the most temporally stable sites for AS1 and AS2 respectively, considering all data from two separate years and two different depths.

Even though sites #19 and #21 were determined to be the most temporally stable sites for AS1 and AS2 respectively, there are problems in estimating field averages from these sites using, for example, the constant offset approach by Grayson and Western (1998). This is due to the signs of their respective constant offsets being opposite for the data in 2009 and 2010. For example, #19 at AS1 had a constant offset of -0.1003 in 2009 and 0.0586 in 2010, and #21 at AS2 had values of -0.0368 in 2009 and 0.2061 in 2010. This indicates that they underestimated the field averages in 2009 and overestimated them in 2010. Therefore, even though they are considered to be the most temporally stable sites, they need different offset values for each year in order to adequately estimate field averages. This finding contradicts the argument of Grayson and Western (1998) that a constant offset can be used irrespective of the time of year to estimate the field average using point measurements which have a non-zero mean relative difference. This fact places limitations on the practical applications of temporal stability analysis for upscaling point soil moisture measurements to field averages since additional spatial samplings are still required to estimate the offset values which need to be updated for different time periods. Thus we must reconsider whether the temporal stability approach can achieve its

original purpose: minimizing the numbers of point observations to characterize the spatiotemporal variability of soil moisture.

Most of the previous studies conducting temporal stability analysis across different scales and depths did not attempt to validate the most temporally stable sites for different time periods, considering differences in vegetation type and rainfall amounts that could occur over time. For instance, in the Corn Belt region in the Midwestern United States crops are typically rotated between corn and soybean every year or every two years. Therefore, the effect of different crop types (or other factors which temporally vary and affect the stability of soil moisture conditions) should be accounted for in determining temporally stable sites for the purpose of validating long-term remote sensing soil moisture retrievals or in evaluating long-term hydrologic model predictions.

As I have noted, the application of temporal stability analysis has received considerable attention for investigating soil moisture variability and in validating remotely-sensed soil moisture data (Mohanty and Skaggs, 2001; Cosh et al., 2004; Cosh et al., 2006; Starks et al., 2006; Choi and Jacobs, 2008; Heathman et al., 2009; Choi and Jacobs, 2011; Joshi et al., 2011). However, the aforementioned limitations of the temporal stability analysis (low possibilities of finding a single representative location for different periods and different soil layers, and uncertainty in the transferability of offset constants in time and space) still remain. Due to limitations involving temporal stability studies, De Lannoy et al. (2007b) suggested that certain transformation functions or upscaling of point measurements may offer a viable way to obtain spatial averages within a satellite pixel or a model grid cell in order to limit the representativeness error of point measurements.

Thus, the CDF matching method, an alternative upscaling method, is tested in the present study as a possible means for obtaining temporally and spatially transferrable observation operators

4.3 Study Area and Data Collection

Soil moisture observations for this study were collected at two small agricultural watersheds, AS1 (2.23ha) and AS2 (2.71ha) that are located in the Upper Cedar Creek Watershed (UCCW) in northeastern Indiana (Figure 4.1). The two watersheds have permanent soil moisture sensors and weather stations installed as part of an environmental monitoring network in the UCCW established by the United States Department of Agriculture, Agricultural Research Service (USDA-ARS) National Soil Erosion Research Laboratory (NSERL). The monitoring network was initiated in 2002 by the Source Water Protection Project and the Conservation Effect Assessment Project, and are planned to be used for the validation of upcoming NASA SMAP soil moisture products as one of the validation sites. For more detailed information regarding the UCCW environmental network please see Flanagan et al. (2008).

Both AS1 and AS2 were planted in soybeans in 2009 and corn in 2010. The two fields have different management practices; AS1 is in no-till, with AS2 having rotational tillage in years when corn is planted. The slope of the fields ranges from 5 to 10%. Major soil types in each field are a Glynwood (GnB2) silt loam (Fine, illitic, mesic Aquic Hapludalfs) and a Blount (BaB2) silt loam (Fine, illitic, mesic, Aeric Epiaqualfs) for AS1 and AS2, respectively (Figure 4.1).

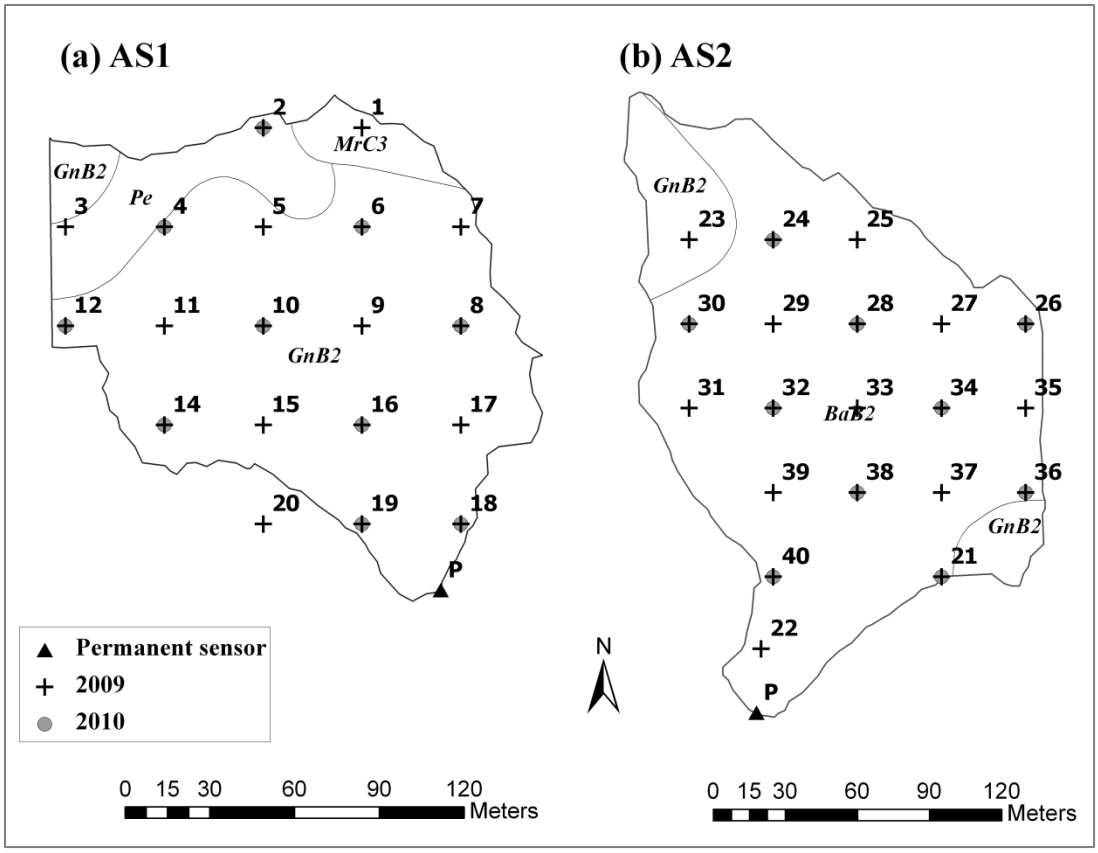


Figure 4.1 Study Area

Temporary spatial soil moisture measurements were obtained during the summers of 2009 and 2010 between the last crop spraying and harvest so as not to interfere with landowner agricultural operations. Twenty temporary soil moisture sensors, with 35m grid spacing, were installed across the watersheds at the surface 5cm depth in 2009, and two sensors at two depths (5cm and 20cm) were installed, with a 70m grid spacing, at ten locations in each field in 2010 (Figure 4.1). The 2010 sensor locations were coincident with the 2009 locations based on 70m spacing. Some of the temporary sensors in 2009 did not function properly because of data transmission errors. Therefore, 18 and 15 point measurements out of 20 sensors were obtained in 2009 for AS1 and AS2, respectively.

Also due to data transmission errors, AS2 did not report data at several sites in August and September of 2009. Therefore, AS2 had a shorter measurement period from July 16 to August 8 in 2009. However, measurements were obtained in AS1 from July 15 to September 21, 2009. In 2010, after modifications in data transmission technology, soil moisture was spatially measured from June 30 to September 21 from all ten temporary sensor locations in AS1 and AS2.

The permanent soil moisture sensors are located near the outlet of the two watersheds. They provide soil moisture measurements at four different depths (5cm, 20cm, 40cm and 60cm) every 10 minutes. The temporary sensors measured soil moisture every 30 minutes and transmitted to a field data logger. The original 10 minute measurements from the permanent sensors were resampled so that they would match the temporary sensor 30-minute measurements. Out of four different measurement depths from the permanent sensors, only 5cm and 20cm measurements were used in combination with the temporary sensor measurements for data analysis in this study.

Soil moisture sensors used for this study were Stevens SDI-12 Hydra Probes, a type of frequency domain reflectometry that measures soil water content based on changes in soil dielectric permittivity. More details regarding the accuracy of the soil moisture measurements, sensor calibrations and soil properties information can be found in Han et al. (2011) and Heathman et al. (*in press*).

4.4 Methodology

As Figures 4.2 and 4.3 show, soil moisture observations from the permanent sensors do not represent the field average appropriately in either field. Even though the permanent sensor measurements and the field averages are generally in agreement with the rainfall dynamics, their relationships are complicated. For example, the AS1 permanent sensor overestimated the field averages for the first half of the experiment period and underestimated them during September, 2009 (Figure 4.2). Given these problems with the permanent sensor measurements, it would be desirable to transform the soil moisture observations from the permanent sensors so that they match the field averages. This transformation can be achieved using observation operators (Drusch et al., 2005a). This section explains the different observation operators used in this study to estimate field average conditions and standard deviations.

4.4.1 Estimation of Field Averages

CDF matching method

In this study, the cumulative distribution function (CDF) matching method is the primary approach used for estimating field averages based on point soil moisture measurements from the permanent sensors. The CDF matching method has been used in many geophysical studies for the purpose of bias correction or rescaling of different sets of observations. For example, the concept of CDF matching was applied for rainfall estimation, specifically to adjust the relationships between reflectivity and rain rate (Atlas et al., 1990) and to quantify uncertainties in satellite rainfall retrievals (Anagnostou et al.,

1999). Wood et al. (2002) also used a similar approach to correct biases in model forecasted rainfall and temperature for long-range hydrologic forecasting. With regard to soil moisture variables, Liu et al. (2009) and Liu et al. (2011) merged soil moisture retrievals from different satellites after rescaling them by the CDF approach. Reichle and Koster (2004a) and Drusch et al. (2005a) adopted the CDF method to define observation operators for rescaling satellite-based soil moisture observations into model climatology. In a field scale soil moisture study, De Lannoy et al. (2007) used the CDF matching method as one of the upscaling methods to estimate field averages from point soil moisture measurements. While Liu et al. (2009) and Liu et al. (2011) defined different scaling linear equations for separate percentile segments of CDF curves, Drusch et al. (2005) and De Lannoy et al. (2007b) derived a polynomial equation for their entire dataset. In the present study, the same approach as Drusch et al. (2005) is adopted to find appropriate observation operators.

The CDF matching method is used to adjust the CDF curve of one variable to the curve of another reference variable. In this study, the CDF curves of the point measurements are scaled to those of the field averages. Higher order moments of point measurement distribution functions, including mean, standard deviation and skewness, can be corrected through the CDF matching method. Procedures for the CDF matching method are summarized below.

1) Rank the two datasets of permanent sensor measurements ( ) and the field average ( ).

2) Compute differences between each corresponding element from the ranked datasets

(4. 1)

where and are the ranked field averages and point measurements, respectively, for $i=1...n$, where n is the number of the measurements.

3) Find a optimum polynomial fit to the ranked point measurements and corresponding differences using the least square optimization. In this study, the third order polynomial fit was chosen for all datasets according to De Lannoy et al. (2007) and the principle of parsimony.

(4.2)

where is the estimated difference for

4) Compute the estimated field average from the estimated difference.

(4.3)

The polynomial equation (Equation 4.2) is the observation operator for each dataset and is used to remove the systematic differences between the point soil moisture measurements and the field averages.

Other upscaling methods

Besides the CDF matching method, other linear transformation methods are applied to find the best observation operators. First, a simple linear relationship between point measurements and field averages is used.

(4.4)

where is the estimated field average corresponding to point measurements ( )

The second alternative method is to use the point measurements from the most temporally stable site and a constant offset (obtained from the temporal stability analysis as Grayson and Western (1998) introduced in order to estimate field averages from single point measurements.

—— (4.5)

where $\theta_{st,i}$ is soil moisture measurements from the temporally stable site at time *i*. The constant offset (is a temporal mean relative difference from the temporally stable site, defined by Vachaud et al. (1985).

– ——— (4.6)

where is an average of spatially distributed *m* point measurements at time step *i* (–).

The mean relative difference approach has been applied in several previous studies to find representative field average values once the temporally stable site, having a non-zero mean relative difference, is found (Grayson and Western, 1998; Starks et al., 2006; Heathman et al., *in press*). However, even though one of the temporary sensors in this study is found to be the most temporally stable site, the information from the site cannot be used for long-term monitoring because the sensor equipment must be removed from the field after the experiment period to allow landowner farming operations to continue. Therefore, the mean relative difference method using the permanent sensor measurements, instead of the "in-field" temporally stable site, is also tested as a third

alternative observation operator.

(4.7)

where the constant offset (is a temporal mean relative difference from the permanent sensor measurements defined as Equation 4.6.

Even though the mean relative difference approach can predict field averages with high correlation coefficients, as shown in Starks et al. (2006), it does not guarantee unbiased estimation. In addition, the amplitude of the estimated field average will be smaller than that of the actual soil moisture measurement if the mean relative difference is positive, and the amplitude will be larger if the difference is negative, due to the multiplicative relationship (De Lannoy et al., 2007b). In order to avoid these problems of the mean relative difference approach, absolute difference relationship is also used based on the additive relationship.

(4.8)

where is a temporal mean of absolute differences between the permanent sensor measurements (and spatial averages (, –

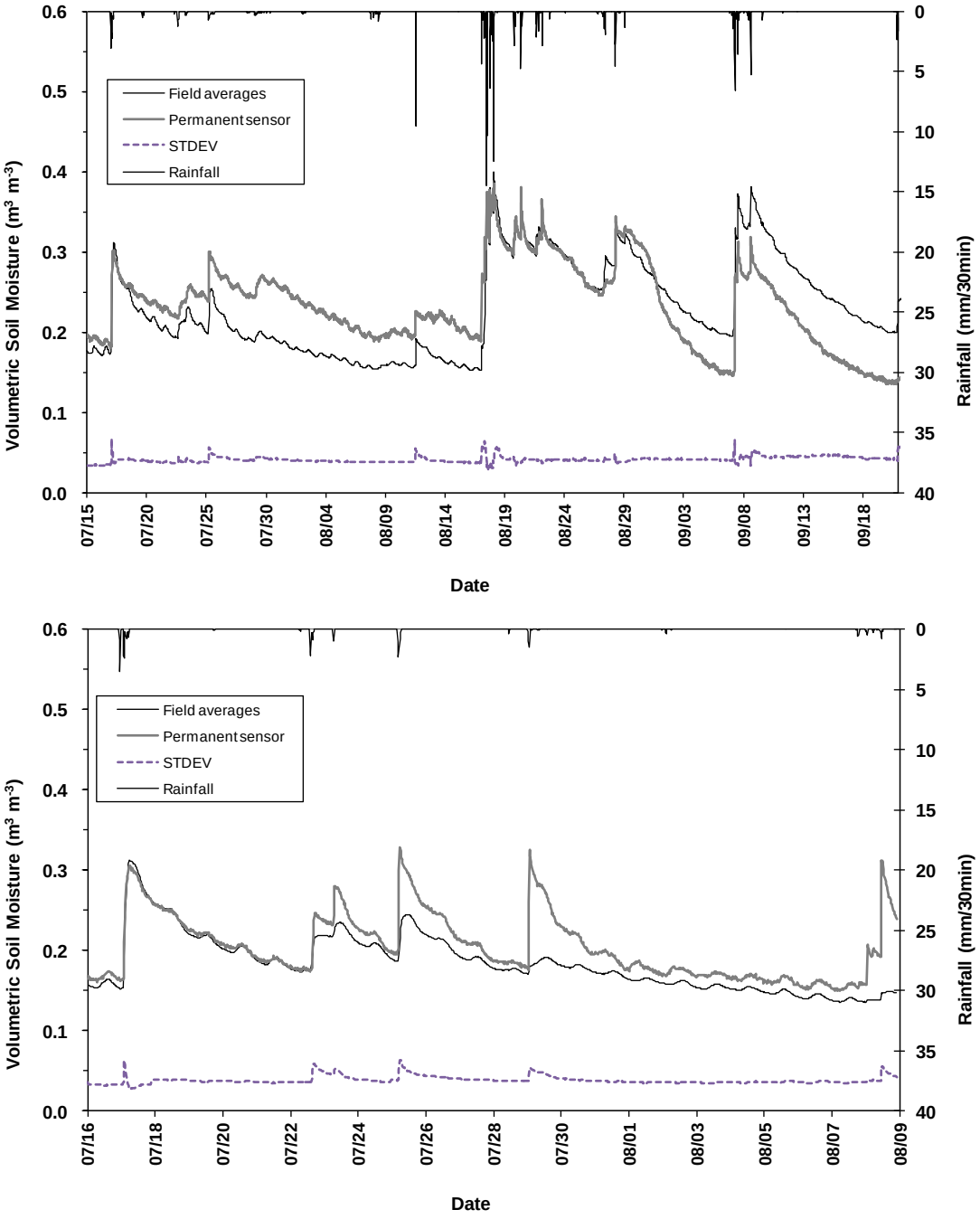


Figure 4.2 Soil Moisture Observations from the Permanent Sensor, Field Averages and Standard Deviations in 2009 at (a) AS1 and (b) AS2

4.4.2 Estimation of Soil Moisture Variability

The same principle of the aforementioned CDF matching method is applied to estimate spatial standard deviations of soil moisture within the study areas using point measurements from the permanent sensors. Vachaud et al. (1985) used the concept of temporal stability to identify the value of one or two standard deviations, as well as the mean, based on the assumption that particular locations maintain their rank in the cumulative probability function for different observation times. In the present study, a hypothetical location is assumed to be a temporally stable site and to provide soil moisture differing from the mean by one standard deviation (
consistently. The means ( ) and standard deviations ( ) of soil moisture observations from the spatially distributed sensors are assumed to represent the true variability of the soil moisture in the watersheds. Using the mean and standard deviations obtained from all the sensors, the soil moisture time-series data for the hypothetical location (that represent consistently wet conditions (mean plus one standard deviation) are computed (. Then, the same CDF matching procedures described in the previous section are repeated to compute by rescaling the permanent sensor measurements (to . The third order polynomial equations are selected for observation operators again for consistency. Once is determined, the predicted standard deviation at time step *i* ( ) is obtained by subtracting the field average ( ) estimated by the CDF matching and the permanent sensor data.

(4.9)

4.5 Results and Discussions

4.5.1 Upscaling Point Measurements to Field Averages

Results of CDF matching method

Figure 4.3 shows how the original CDF of permanent sensor measurements is adjusted to the CDF of field averages for each dataset of AS1 and AS2 (5cm in 2009, and 5cm and 20cm in 2010). In 2009, the permanent sensor at AS1 overestimated the field averages for some periods and underestimated them for other periods, whereas the permanent sensor at AS2 consistently overestimated the field averages (Figure 4.3 (a) and (b)). In 2010, permanent sensors at both AS1 and AS2 highly underestimated the field averages for both 5cm and 20cm depth measurements.

The observation operators were determined through least square fits of a third order polynomial with the coefficients listed in Table 4.1. It is found that the observation operators could be used to rescale the CDFs of point measurements to field averages successfully. Even though the relationships between ranked differences and ranked point observations from the six data sets were very different from each other, the third order polynomial was consistently selected for observation operators in order to test their transferability between data sets (Figure 4.4). In addition, it was found that other higher order polynomial equations did not provide significantly improved fitting results.

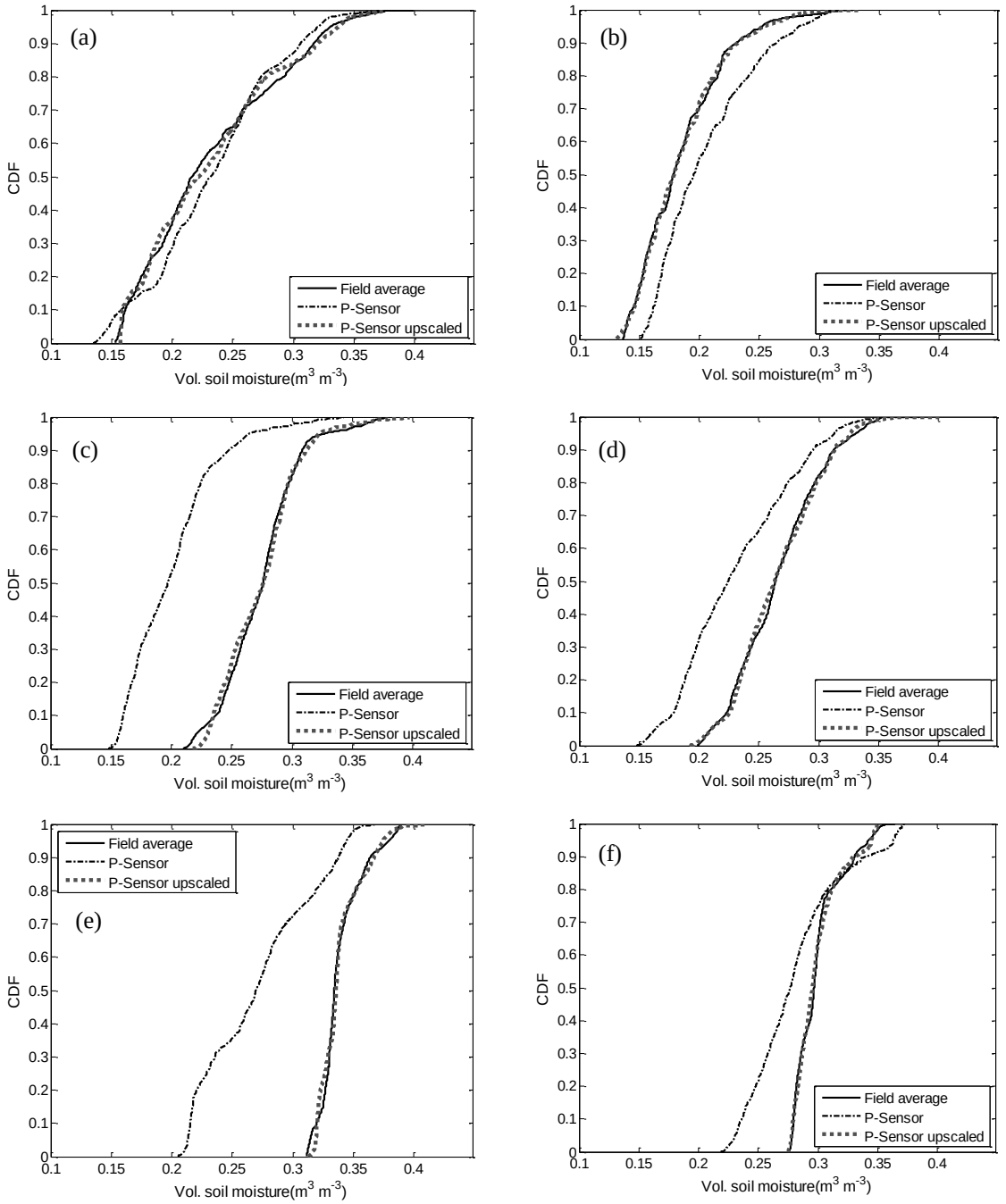


Figure 4.3 CDFs of Soil Moisture
(a) AS1, 5cm in 2009, (b)AS2, 5cm in 2009, (c) AS1, 5cm in 2010, (d) AS2, 5cm in 2010, (e)AS1, 20cm in 2010, (f)AS2, 20cm in 2010

Table 4.1 Coefficients of Transforming Equations for Upscaling from the Permanent Sensor Measurements to Field Averages

Method*	coefficients	5 cm (2009)		5 cm (2010)		20 cm (2010)	
		AS1	AS2	AS1	AS2	AS1	AS2
LR (Equation 4.4)	b_0	0.064	0.040	0.117	0.125	0.244	0.164
	b_1	0.718	0.706	0.787	0.613	0.351	0.483
AMD (Equation 4.8)		0.002	0.020	-0.074	-0.035	-0.068	-0.018
MRD-ST (Equation 4.5)		0.061 (#17)**	-0.037 (#21)**	0.102 (#8)**	0.134 (#30)**	0.128 (#19)**	0.183 (#21)**
MRD-P (Equation 4.7)		0.0367	0.1164	-0.274	-0.137	-0.2041	-0.065
CDF (Equation 4.2)	a_0	0.494	-0.481	-0.342	-0.127	-0.774	0.449
	a_1	-31.595	47.843	32.537	13.772	57.008	-7.235
	a_2	25.165	-31.842	-24.271	-10.665	-45.440	7.727
	a_3	-6.310	6.773	5.654	2.437	11.222	-3.138

* LR: linear relationship
AMD: absolute mean difference
MRD-ST: relative mean difference from the most temporally stable site
MRD-P: relative mean difference from the permanent sensor
CDF: CDF matching method

** Numbers in parenthesis indicate locations of the most temporally stable sensors

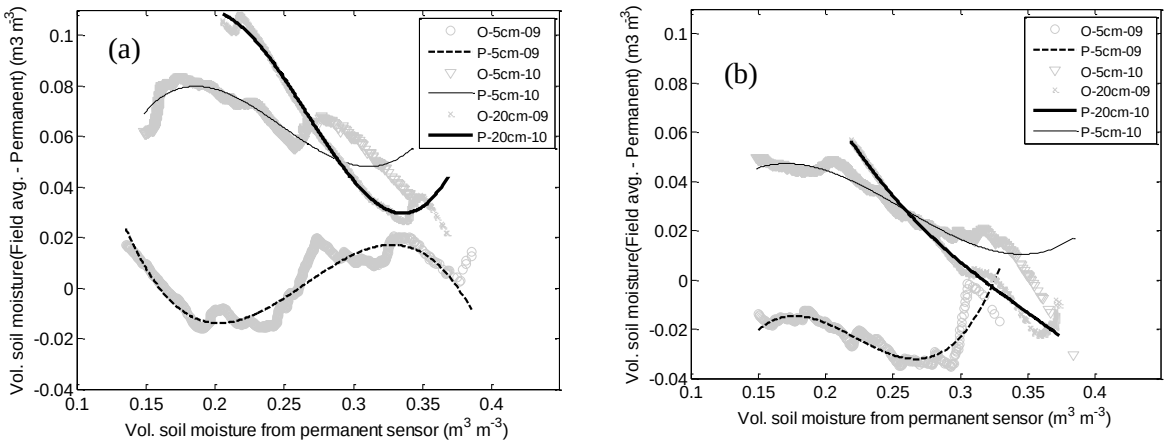


Figure 4.4 Observation Operators for the CDF Matching (a) AS1, (b)AS2 (O and P indicate observed and predicted data respectively)

Figure 4.4 shows that the observation operators were non-linear with the permanent point measurements for all data sets in the same manner that remotely-sensed soil moisture observations were non-linear with the model predictions in Drusch et al (2005). Shapes of the observation operators for AS1 and AS2 are similar for the 2010 data, whereas they are very different for the 2009 data. Due to more stable soil moisture conditions at deeper depths, the data sets from the 20cm sensors in 2010 show less non-linearity compared to the data for the 5cm sensors for both AS1 and AS2. Even though the operators for the deeper depths appear linear for most of the soil moisture ranges, a high degree of non-linearity is associated with higher soil moisture values, which is consistent with highly variable and unstable soil moisture conditions in response to rainfall as discussed in a later section.

The significant difference in the shapes of CDFs and observation operators between different data sets implies that there is a low probability that the observational operators are temporally and spatially transferable. In spite of having the same rainfall and crop

types (the two field sites are only 500 meters apart), their soil moisture variabilities are very different. This fact suggests high local heterogeneity in these two fields. More details about possible factors affecting the transferabilities of the observation operators are discussed in later sections.

Comparison of different observation operators

Table 4.2 summarizes the results of the different upscaling methods explained in the methodology section. Pearson's linear correlation coefficient (R), root mean square error (RMSE) and mean bias error (MBE) are used to evaluate the performances of the five different observation operators. They are defined as follows:

(4.9)

(4.10)

(4.11)

where and are the means of estimated field averages and permanent sensor measurements, respectively.

The CDF matching method shows the best results for all data sets: correlation coefficients are close to 0.99, RMSE is improved more than an order of magnitude from the original data set (P) and bias is successfully reduced to zero. Successful application of the CDF matching method was also presented by De Lannoy et al. (2007b).

Table 4.2 Performance Comparison of Different Upscaling Methods

	Method*	5 cm (2009)			5 cm (2010)			20 cm (2010)		
		R	RMSE	MBE	R	RMSE	MBE	R	RMSE	MBE
AS1	P	0.649	4.62E-02	0.002	0.910	7.57E-02	-0.074	0.882	7.43E-02	-0.068
	LR	0.649	4.37E-02	0.000	0.910	1.30E-02	0.000	0.882	8.30E-03	0.000
	AMD	0.649	4.61E-02	0.000	0.910	1.51E-02	0.000	0.882	2.99E-02	0.000
	MRD-ST	0.991	9.80E-03	0.001	0.935	1.14E-02	-0.001	0.908	7.90E-03	0.000
	MRD-P	0.649	4.60E-02	-0.006	0.910	2.51E-02	0.001	0.882	4.10E-02	0.002
	CDF	0.997	4.80E-03	0.000	0.991	4.10E-03	0.000	0.989	2.70E-03	0.000
AS2	P	0.789	3.20E-02	0.020	0.801	4.49E-02	-0.035	0.944	2.76E-02	-0.018
	LR	0.789	2.19E-02	0.000	0.801	2.12E-02	0.000	0.944	6.40E-03	0.000
	AMD	0.789	2.49E-02	0.000	0.801	2.78E-02	0.000	0.944	2.06E-02	0.000
	MRD-ST	0.978	7.50E-03	0.000	0.985	6.20E-03	0.000	0.887	1.00E-02	0.000
	MRD-P	0.789	2.32E-02	-0.001	0.801	3.30E-02	0.001	0.944	2.31E-02	0.001
	CDF	0.997	2.60E-03	0.000	0.997	2.80E-03	0.000	0.995	2.00E-03	0.000

* P: Original data from permanent sensor, LR: linear relationship, AMD: absolute mean difference, MRD-ST: relative mean difference from the most temporally stable site, MRD-P: relative mean difference from the permanent sensor, CDF: CDF matching method

The mean relative difference (MRD) method using the most temporally stable points (MRD-ST) provides the second best results (Table 4.2). Heathman et al. (*in press*) determined that locations #17 and #21 are the most temporally stable sites (having smallest standard deviations of MRD) for AS1 and AS2, respectively for the same 2009 dataset. They also showed that the observations from those temporally stable sites can be successfully converted to field averages using a constant offset (Equation 4.5) resulting in a coefficient of determination (R^2) greater than 0.95. For the 2010 data sets, locations #8 and #19 were found to be the most temporally stable for 5cm and 20cm depth, respectively at AS1, and locations #30 and #21 for 5cm and 20cm depth, respectively at AS2 (Heathman et al., *in review*). This method improved overall error statistics, especially R for 5cm depth measurements. However, R for 20cm data at AS2 for 2010 decreased (= 0.887) even when compared to the original permanent sensor measurements (= 0.944). This coincides with the low R^2 (= 0.786) of linear regressions between the actual field averages and the estimated field averages using the constant offset value, compared to R^2 of other data sets close to 0.9 (Heathman et al., *in review*)

The disadvantage of the MRD-ST method is that it does not remove bias completely. Although this method improves R and RMSE values, with values comparable to the CDF matching method, some data sets (5cm in 2009 and 5cm in 2010 for AS1) still have non-zero MBE. In addition, Figure 4.5 shows that this method is not successful in predicting several peak soil moisture conditions (e.g. August 7, September 7 and 8 for AS1 and July 23 for AS2).

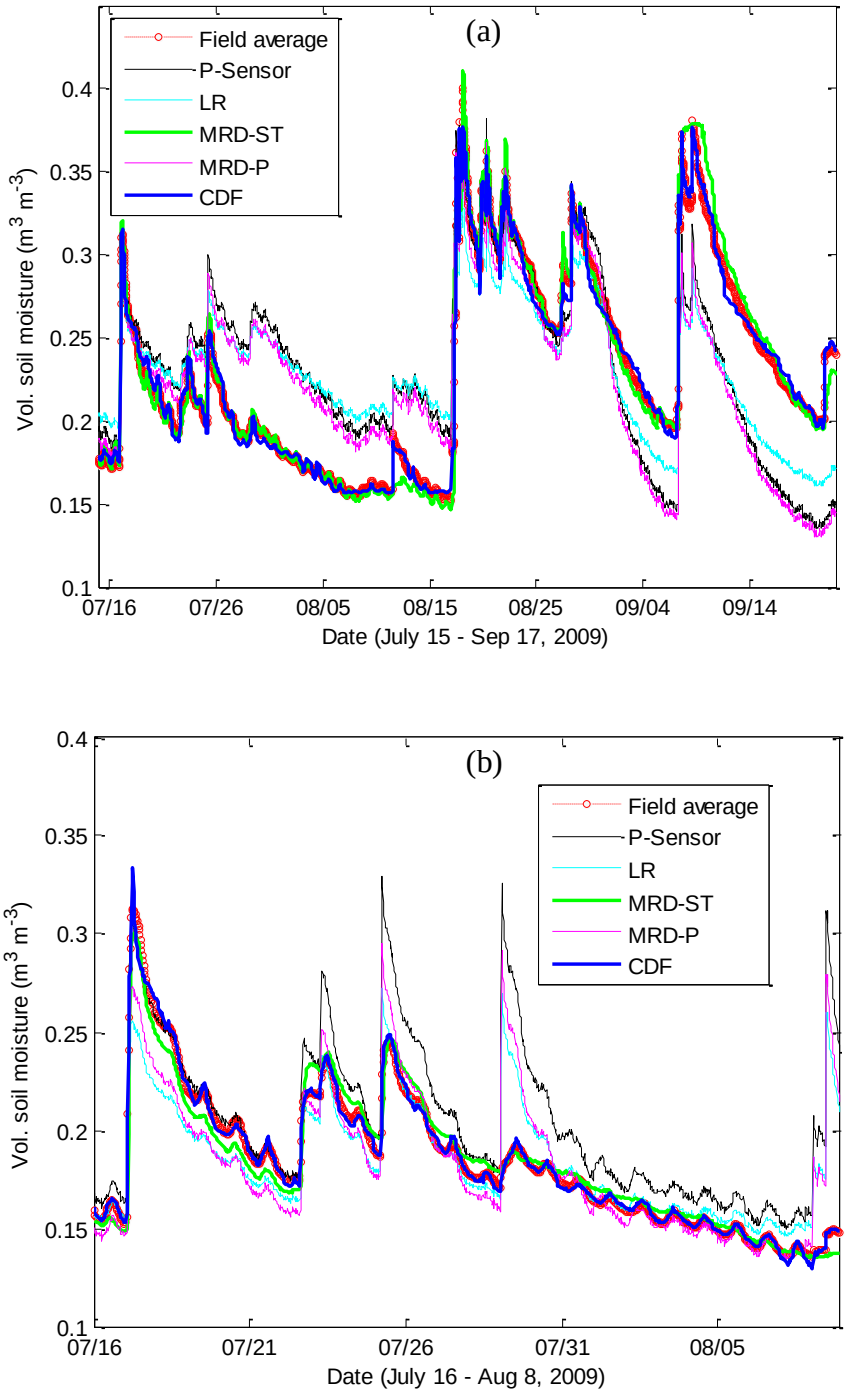


Figure 4.5 Time Series of Soil Moisture for Permanent Sensor Data, Field Average and the Transformed Permanent Sensor Data for (a) AS1 (b) AS2
(P-Sensor: Permanent sensor measurements, LR: linear relationship, MRD-ST: relative mean difference using the most temporally stable site, MRD-P: relative mean difference using the P-Sensor, CDF: CDF matching)

The mean relative difference method using the data from the permanent sensors (MRD-P in Table 4.2) resulted in little improvement compared to the original data (P); Low R of the original data remains the same, RMSE is improved slightly, and MBE is decreased but still remains non-zero for all data sets. These results might be expected based on the previous temporal stability studies because permanent sensor measurements had high standard deviations of MRD for both AS1 and AS2 (Heathman et al., *in press;* Heathman et al., *in review*). Equation 4.7 which adjusts the point measurements to field averages using the constant offset is not suitable for temporally unstable point measurements, such as the data from the permanent sensors in the present study.

The absolute mean difference method (AMD) is not very successful in upscaling the point measurements to the field averages, except in removing bias as intended (Table 4.2). The method using simple linear relationships (LR) produces better results than either AMD or MRD-P, but much less successful results than either the CDF matching or the MRD-ST. Due to relatively stable soil moisture conditions at the 20cm depth, upscaling results are quite good, even for the LR method where for the AS2 data in 2010, the LR method produces less RMSE than MRD-ST.

Figure 4.5 compares the results of different observation operators graphically for the 2009 data sets. Results for the CDF matching and the MRD-ST methods show that upscaling permanent sensor measurements to the field averages was successful, whereas the other three methods fail to do so.

Table 4.1 illustrates coefficients of the different observation operators for each data set. Some of the coefficients from different data sets have similar magnitudes but others have

very different magnitudes indicating little possibility of their transferability between data sets.

4.5.2 Spatial and Temporal Transferability of Observation Operators from CDF Matching

Six data sets obtained from different years (2009 and 2010), sites (AS1 and AS2) and depths (5cm and 20cm) allow us to test whether the observation operators from CDF matching are transferable in time and space. Data0 in Table 4.3 represents the original data from the permanent sensor and Data1 represents upscaled data using the observation operator from the CDF matching with corresponding field averages (which are the same as the CDF results in Table 4.2). The results of these two data sets are used as reference to evaluate the results of other observation operators (Data2, Data3 and Data4) for spatial and temporal transferability tests.

Temporal transferability at two different years (2009 vs. 2010)

Data2 in Table 4.3 are results of applying the observation operators from another time period in order to test temporal transferability of the observation operator. For example, the observation operator obtained from 2009 data was used to upscale the permanent sensor measurements of the year 2010. As expected from the very different shapes of CDF curves in Figure 4.3 and the observation operators in Figure 4.4 for the data from 2009 and 2010, the error statistics with Data2 are not satisfactory, indicating unsuccessful temporal transferability of observation operators for both AS1 and AS2. Although correlations are improved, error terms are increased using observation operators from another time period.

Table 4.3 Spatiotemporal Transferabilities of Observation Operators from the CDF Matching Method

	Data*	5 cm (2009)			5 cm (2010)			20 cm (2010)		
		R	RMSE	MBE	R	RMSE	MBE	R	RMSE	MBE
AS1	Data0	0.649	4.62E-02	0.002	0.910	7.57E-02	-0.074	0.882	7.43E-02	-0.068
	Data1	0.997	4.80E-03	0.000	0.991	4.10E-03	0.000	0.989	2.70E-03	0.000
	Data2	0.948	7.14E-02	0.068	0.942	8.27E-02	-0.082	-	-	-
	Data3	0.956	2.57E-02	-0.019	0.990	3.26E-02	-0.032	0.961	4.33E-02	-0.043
	Data4	-	-	-	0.950	2.53E-02	0.023	0.974	2.10E-02	-0.007
AS2	Data0	0.789	3.20E-02	0.020	0.801	4.49E-02	-0.035	0.944	2.76E-02	-0.018
	Data1	0.997	2.60E-03	0.000	0.997	2.80E-03	0.000	0.995	2.00E-03	0.000
	Data2	0.989	6.12E-02	0.061	0.974	5.84E-02	-0.057	-	-	-
	Data3	0.979	1.76E-02	0.014	0.995	3.31E-02	0.033	0.933	4.48E-02	0.044
	Data4	-	-	-	0.900	2.90E-02	0.017	0.993	1.21E-02	0.005

*Data0: Original data from permanent sensor
Data1: Upscaled results using the observation operator from the same data
Data2: Upscaled results using the observation operator from another year to test temporal transferability (2009 vs. 2010)
Data3: Upscaled results using the observation operator from another site to test spatial transferability (AS1 vs. AS2)
Data4: Upscaled results using the observation operator from another depth to test vertical transferability (5cm vs. 20cm)

There seems to be two major possible reasons for this unsuccessful temporal transferability: different rainfall distribution or characteristics, and different crop type in 2009 and 2010. Although the total amount of rainfall at AS1 was greater in 2009 (225 mm) than in 2010 (130mm) for the same period (July 15 to September 20), the year of 2010 had more evenly-distributed rainfall events, whereas the year of 2009 had a long dry period before heavy rainfall occurred in the middle of August (Figure 4.6). In addition, initial soil moisture conditions in 2010 were very wet due to the heavy rains a few days before installing temporary sensors. Thus, the field averages of surface soil moisture (5cm) in 2010 were wetter than those in 2009 (Figure 4.3). Since surface soil moisture variations depend more on evapotranspiration than lateral flow, different crop type (soybean in 2009 and corn in 2010 at both AS1 and AS2) is also an important factor in considering the failure of temporal transferability. However, in this study, it is difficult to isolate the effect of crop type on transferability due to the difference in the major forcing variable, rainfall in those two years.

Besides these two main factors, rotational tillage practices (only for AS2) and local heterogeneity at the location where the permanent sensors are installed may affect the soil moisture variation between 2009 and 2010. The permanent sensor measurements did not show consistent relationships with the field averages in 2009 and 2010. In 2010, both permanent sensors at AS1 and AS2 underestimated the field average soil moisture conditions consistently, whereas the permanent sensor at AS2 consistently overestimated the field averages in 2009 (Figure 4.3). Since the permanent sensors are located near the outlet and boundary of the watersheds, they seem to be more sensitive to dry or wet

conditions. Whereas AS1 had no-tillage every year, the AS2 field was tilled only in 2010 before planting corn. Less residue cover due to the tillage in 2010 at AS2 may result in drier surface soil moisture condition, but the effect of tillage seems to be blurred by the rainfall difference because AS2 had wetter soil conditions in 2010 than in 2009.

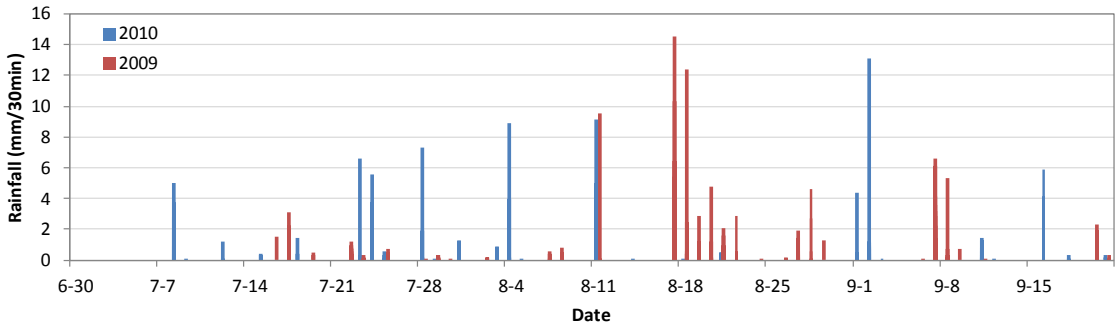


Figure 4.6 30minute Rainfall comparison in 2009 and 2010

Spatial transferability at two different locations (AS1 vs. AS2)

Successful spatial transferability of the observation operators with the 2009 data would not be expected due to significant differences in the shape of CDF curves (Figure 4.3 (a) and (b)) and the observation operators (Figure 4.4) of AS1 and AS2. However, with the 2010 data, the curves are more similar, which is more conducive for spatial transferability.

Error statistics of Data3 in Table 4.3 are results of the application of observation operators from another site in order to examine spatial transferability of observation operators between AS1 and AS2. The results of Data3 show that the observation operators are spatially transferable in general; both R and RMSE are improved compared to the original permanent sensor measurements (Data0) except for the 20cm data for AS2 in 2010. However, the results are not as good as the results of Data1 which are from their

own observation operators. Therefore, when an observation operator for a site is not available, it may be possible to use an alternative observation operator from a nearby site under similar climate, soil and crop conditions, like AS1 and AS2. In spite of possible spatial transferability with improved R and RMSE in Table 4.3, however, careful application of the observation operators from another site should be required because they may highly overestimate wet conditions or underestimate dry conditions.

Compared to 2009 data, the 2010 data have more consistent trends of CDFs between AS1 and AS2. Therefore, the results of spatial transferability in 2010 show greater improvement, especially for AS1; more than 50% of RMSE decrease and R of Data3 is as high as R of Data1 with 5cm data at AS1 in 2010.

Even though temporal transferability was not shown to be successful as mentioned above, spatial transferability may have greater potential. Thus, rainfall characteristics seem to be the most important factor in determining the success of transferability of observation operators, considering AS1 and AS2 are close enough to assume they are under the same climate conditions. Different tillage practices and slightly different soil properties between AS1 and AS2 seem to be less critical than the differences in rainfall; AS2 has slightly more clay content than AS1 (Heathman et al., *in press*), with AS2 being tilled in the spring and AS1 remaining in no-till in 2010. Again, the effect of different crop types is difficult to determine because AS1 and AS2 have the same crop every year.

Spatial transferability at two different depths (5 cm vs. 20 cm)

Data4 in Table 4.3 shows the results of applying the observation operators from different soil depths at the same sites. Since the permanent sensors at both AS1 and AS2 consistently underestimated the field averages at both 5cm and 20cm depths in 2010 (Figure 4.3), the observation operators are expected to be transferable vertically. Table 4.3 shows that Data4 produce better error statistics for all cases than the original permanent sensor measurements (Data0). Since the permanent sensor at AS1 highly underestimated the field averages, the observation operator from another depth can improve the upscaling results. However, data at AS2 did not have significant differences between the permanent sensor measurements and field averages. Therefore, AS2 data show less improvement than AS1.

Compared to the temporal and spatial transferability tests mentioned above, the vertical transferability test has fewer factors that may affect the results. Soil properties at 5cm and 20cm depth are very similar for both AS1 and AS2 (Heathman et al., *in press*). In general, when only one observation operator from soil moisture measurements is available, either for the 5cm or 20cm depth, it can be used to upscale the point measurements from the other depth to field averages as an alternative.

4.5.3 Estimation of Standard Deviations

Relationship between field averages and spatial variability of soil moisture

Identifying the relationship between field average and spatial variability of soil moisture has been an important issue because of the need to: 1) determine optimal numbers of sampling points in order to estimate field averages within certain specified error limits, 2) infer soil moisture variability within a footprint, given remotely-sensed soil moisture for an area, and 3) assess the varying accuracy of the remotely-sensed soil moisture (Famiglietti et al., 1999; Brocca et al., 2007). As Brocca et al. (2007) summarized, previous studies showed contrasting results on the relationship between mean and variance of soil moisture conditions. Some of these studies showed increasing variances with increasing means of soil moisture while others showed inversely proportional relationships between them. However, a number of previous studies showed strong negative relationships between mean soil moisture and the coefficient of variation ((Charpentier and Groffman, 1992; Famiglietti et al., 1999; Jacobs et al., 2004; Brocca et al., 2007; Choi and Jacobs, 2011). This is because the relationship between field average and the variability of soil moisture becomes more evident when the standard deviation is scaled by the mean soil moisture.

In the present study, the field averages and the standard deviations (STDEV) of surface (5cm) soil moisture measurements do not show any distinct relationship at either AS1 or AS2 (Figure 4.7, (a) and (c)). This seems to be because surface soil moisture variation is very dynamic due to the effect of evaporation. Another difference in the data from this

study compared to previous studies is that more frequent measurements (every 30 minutes) were used in the present study compared to daily measurements. When the soil is fully saturated after a significant rainfall event, the STDEV decreases considerably and then increases as the soil dries. The minimum STDEV was found on August 17 when the heaviest rainfall occurred (Figure 4.2). Therefore, variations of the STDEV of soil moisture highly depend on rainfall intensity (characteristics) and antecedent soil moisture conditions. A moderate amount of rainfall (e.g. July 25) or rainfall with dry antecedent soil moisture conditions (e.g. August 11), increases the STDEV, whereas heavy rainfall (e.g. August 17) or rainfall with wet antecedent soil moisture conditions (e.g. September 7) results in decreased STDEV. Unlike the relationship for the surface soil moisture, the relationships between mean soil moisture and STDEV are more evident for the 20cm depth measurements due to relatively homogeneous and more stable conditions in the deeper soil layers (Figure 4.7 (e)).

When relative variability, measured by CV, is applied, more distinguishable relationships between the mean and variability of soil moisture are revealed as shown in Figure 4.7. The negative relationships between the mean and CV of soil moisture were successfully described using an exponential fit, in previous studies (Jacobs et al., 2004; Brocca et al., 2007; Choi and Jacobs, 2011). The results of exponential fit with data from the present study are shown in Table 4.4. Once these relationships (coefficients *a* and *b*) and field averages are known, the STDEV of the soil moisture can be estimated (. The last three columns of Table 4.4 show error statistics of the estimated STDEV from the exponential equations. Since surface (5cm) soil moisture measurements

have relatively low R^2 with the exponential curve fitting, the resulting error statistics are not satisfactory, especially having very low correlation coefficients (R). Although several previous studies adopted the exponential relationship between the mean and CV to determine optimal numbers of sampling points with a certain specified error (Jacobs et al., 2004; Brocca et al., 2007; Brocca et al., 2010), present study did not achieve successful prediction of STDEV due to the weak relationship between the mean and the CV. In the following section, results of an alternative method to estimate STDEV using the CDF matching method are introduced.

Table 4.4 Performance of Standard Deviation Estimation Using Regression Relationship between CV and Field Average

		Coefficients of regression equation			R	RMSE	MBE
		a	b	R^{2*}			
5 cm (2009)	AS1	0.428	-3.641	0.8623	0.432	3.60E-03	0.000
	AS2	0.4498	-4.054	0.6159	0.375	4.50E-03	0.000
5 cm (2010)	AS1	0.3931	-3.149	0.4948	0.305	5.10E-03	0.000
	AS2	1.146	-5.774	0.7209	0.443	8.90E-03	0.000
20 cm (2010)	AS1	21.08	-14.93	0.928	0.931	3.50E-03	0.000
	AS2	2.241	-6.796	0.9413	0.886	3.00E-03	0.000

*R^2 is the coefficient of determination resulted from the exponential regression between CV and field averages.

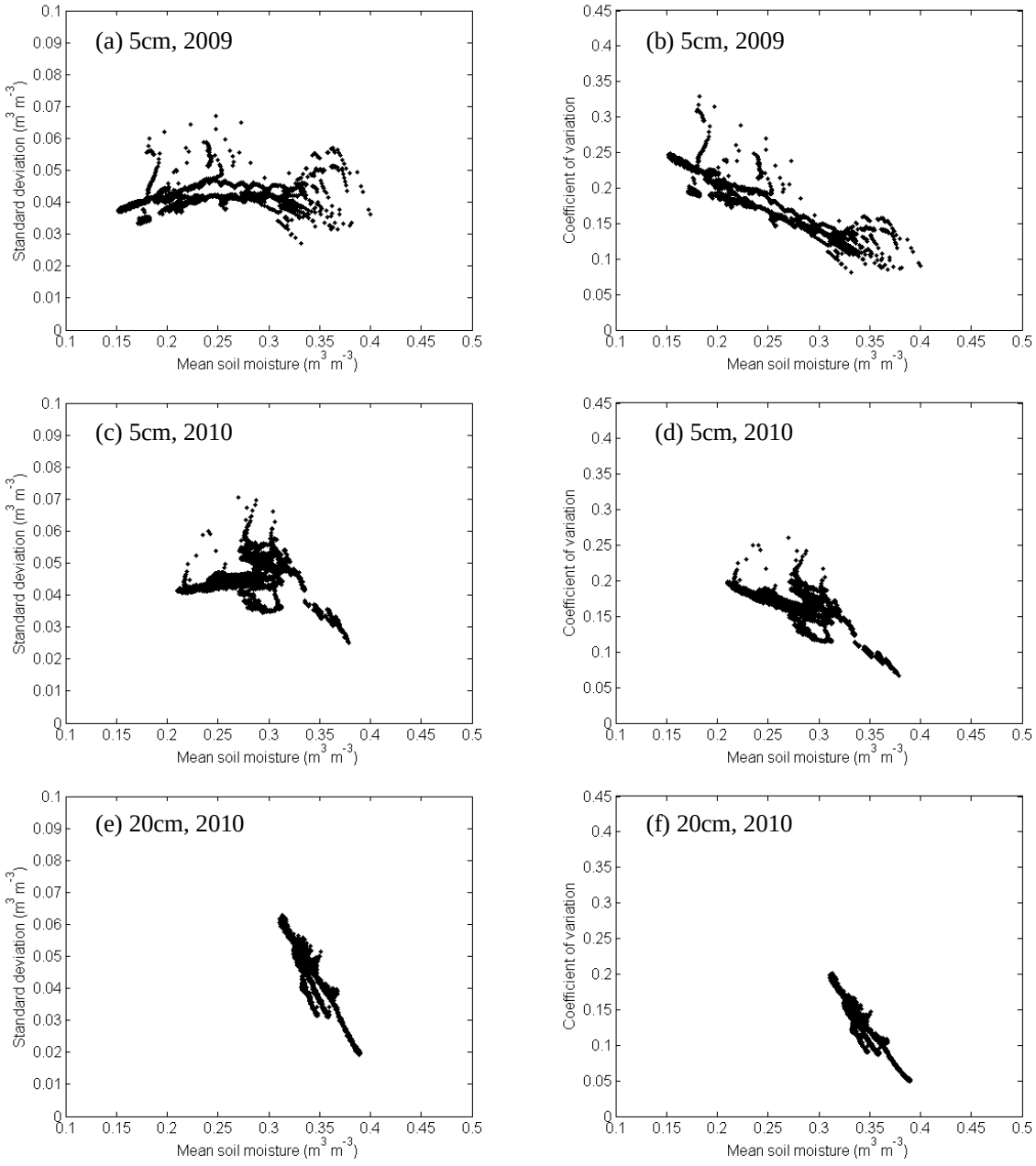


Figure 4.7 Relationship between Mean Soil Moisture and Standard Deviation (a, c, and e) and between Mean Soil Moisture and CV (b, d, and f) at AS1

Estimation of STDEV using CDF matching

Figure 4.8 shows observed (red line) STDEV and estimated (gray line) STDEV by the CDF matching method for six different data sets. The STDEVs of the soil moisture varied dynamically with time due to rainfall and evapotranspiration. In 2009, AS1 and AS2 had STDEVs in similar ranges, whereas in 2010 the variations of STDEVs at the two sites were very different from each other. In 2010, AS2 had higher STDEVs and more dynamic variation of STDEVs than those of AS1 for surface (5cm) soil moisture. Even though AS2 has slightly more clay content than AS1, surface soil moisture at AS2 varied more dynamically than AS1. This is most likely due to the tillage management and thus, having less residue cover in field AS2 (tilled surface) in 2010 may cause more of a response at the soil surface due to rainfall and evaporation. The lowest STDEVs with 5cm data at AS2 in Figure 4.8(d) correspond to major rainfall events. Deeper soil layers in AS1 and AS2 responded in the opposite way compared to the surface soil layer in 2010. STDEVs of the 20cm measurements at AS1 changed more dynamically. For example, STDEVs were reduced abruptly with major rainfall events on August 11 and September 2. However, STDEVs of the 20cm measurements at AS2 rarely changed with time. This may be explained by the fact that AS2 contains more clay than AS1.

The CDF matching method successfully estimated the dynamic variations of STDEVs for all data sets as Figure 4.8 and Table 4.5 show. Correlation coefficients between the observed and predicted STDEVs are higher than 0.9, except for the 5cm measurements at AS1 in 2009 (Table 4.5). Comparisons of error statistics in Table 4.4 and Table 4.5 demonstrate that STDEV prediction by CDF matching is much more successful

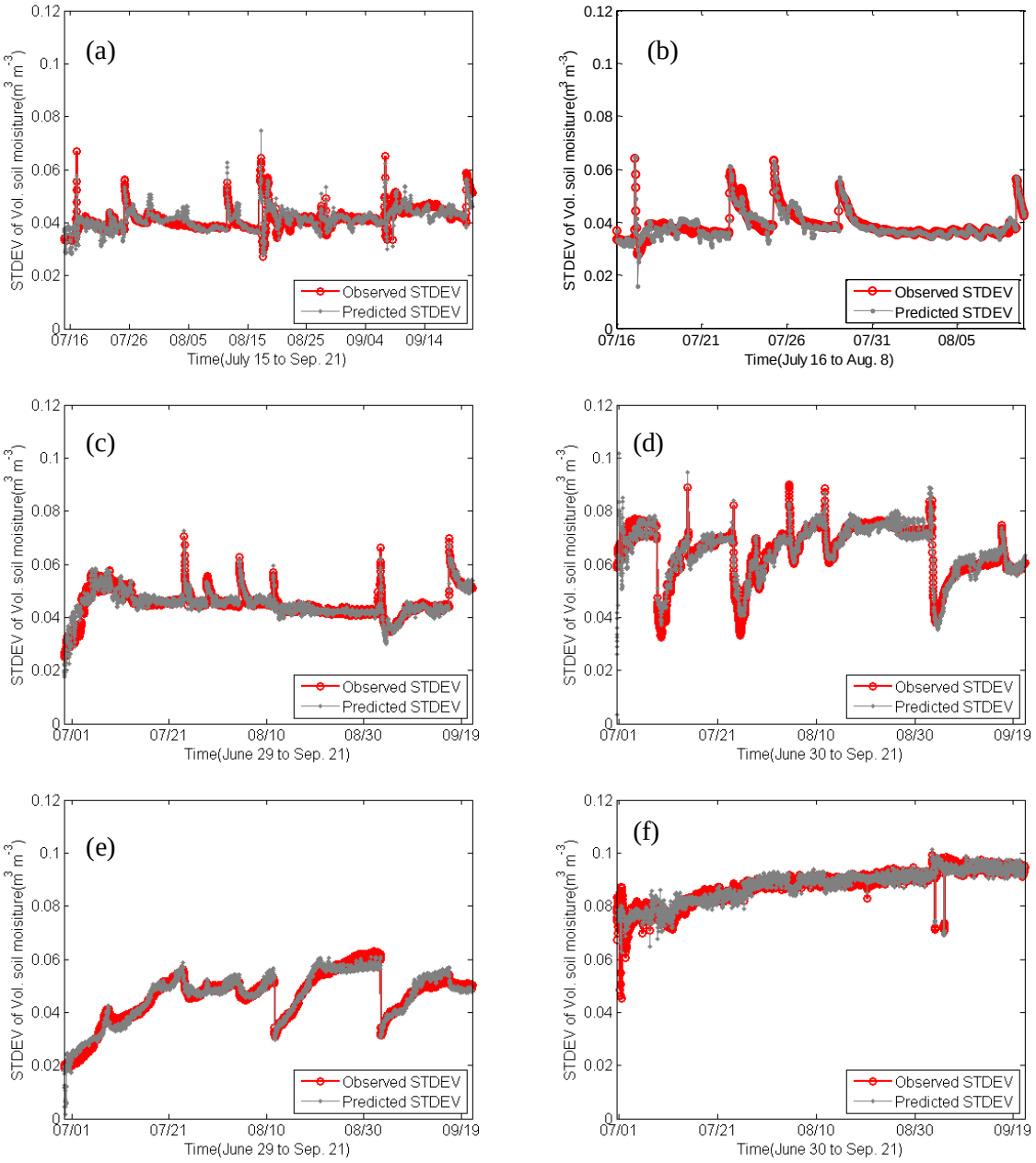


Figure 4.8 Observed and Predicted Standard Deviations of Soil Moisture
(a) AS1, 5cm in 2009, (b)AS2, 5cm in 2009, (c) AS1, 5cm in 2010, (d) AS2, 5cm in 2010, (e)AS1, 20cm in 2010, (f)AS2, 20cm in 2010

compared to applying the exponential relationship between the mean and STDEV of soil moisture. The advantage of this STDEV prediction method using CDF matching is that it can be applied to any soil moisture fields whether or not soil moisture variability has any specific trend or relationship with the field averages. This also shows that the CDF matching method can be used to estimate STDEVs of the field soil moisture as well as field averages using point measurements. Since the third order polynomial fit was adopted consistently in this study, there were high errors with some data, especially with minimum or maximum values. For example, the STDEVs are highly overestimated or underestimated on July 1 in Figure 4.8(d), (e) and (f). In this study, actual 30-minute measurements were used for analysis. Therefore, changes of STDEVs with rainfall events are clearly noticeable. If daily average measurements are used, the dynamic variations would be much less apparent.

Table 4.5 Performance of Standard Deviation Estimation Using the CDF Matching Method

		Coefficients of operation operator ( )				R	RMSE	MBE
		a_0	a_1	a_2	a_3			
5 cm (2009)	AS1	0.50	-30.22	23.94	-5.94	0.835	2.30E-03	0.000
	AS2	-0.24	25.23	-17.21	3.75	0.924	2.00E-03	0.000
5 cm (2010)	AS1	-0.30	26.29	-21.51	5.38	0.943	1.80E-03	0.000
	AS2	-0.09	19.07	-14.15	3.10	0.946	3.20E-03	0.000
20 cm (2010)	AS1	0.07	16.76	-12.94	2.45	0.970	2.30E-03	0.000
	AS2	0.06	12.67	-10.21	2.03	0.965	1.70E-03	0.000

4.6 Conclusions

The primary purpose of this study was to estimate field representative characteristics of soil moisture (field averages and standard deviation) using single point measurements. First, for upscaling the point soil moisture measurements to field averages, some limitations of the temporal stability analysis that has been used widely for this purpose are discussed. Next, alternative upscaling methods, mainly focusing on the CDF matching method, were investigated to test whether they offer means to overcome the limitations of the temporal stability analysis. The difference between the present study and that of De Lannoy et al. (2007b) is that temporal and spatial transferability of the CDF matching method were evaluated for practical applications of the upscaling method. Second, spatial variabilities of soil moisture are estimated from point soil moisture measurements using the CDF matching method.

The results show that the CDF matching method is the best observation operator to estimate the field average soil moisture. The mean relative difference method using the most temporally stable point measurements can also produce satisfactory results, after the CDF matching method. However, the MRD method is not applicable in reality because the permanent sensors cannot be installed “in-field” due to agricultural operations and because the locations of the most temporally stable points are not consistent in time or for different depths. Evaluation of the additional three observation operators, based on simple linear relationships, absolute mean differences, and mean relative differences using the permanent sensor measurements, proved to be unsuccessful in this study.

By comparing data sets from different time periods, areas and soil layers, temporal and spatial transferabilities of the observation operators from the CDF matching were examined. Overall the observation operators are found to be spatially and vertically transferable, but not temporally transferable. Rainfall characteristics and crop type are most likely to be the most significant factors influencing the transferabilities in this study. Other factors such as differences in soil properties, tillage practices and topography may be important as well, but in this study, they appear to be less influential in the transferability of the observation operators than differences due to the rainfall and crop type. Further detailed studies are required to better understand the effect of these factors.

Besides field averages, standard deviations of soil moisture within the fields are also estimated using single point measurements from the edge-of-field permanent sensors. Dynamic variations of standard deviations were successfully predicted through the similar CDF matching method. This approach, which estimates time-series soil moisture variance using point measurements, has potential for various practical applications such as soil moisture data assimilation, land surface modeling, assessing the accuracy of remote sensed soil moisture and establishing optimal *in situ* soil moisture networks.

The CDF matching approach of this study allows estimations of both field averages and variance from point measurements at the permanent soil moisture sensor locations. This approach is expected to contribute to better utilization of *in situ* soil moisture networks which are established for the calibration and validation of remotely-sensed soil moisture products but often times consist of a limited number of permanent soil moisture sensors. The study areas of present study are part of the current monitoring network in UCCW

that are planned to be used for the validation of SMAP soil moisture products at various resolutions (3, 9, 36 km) in the near future. Therefore, findings of this study provide significant insight as to designing better validation schemes for SMAP soil moisture products.

The SMAP Cal/Val Plan contains various aspects of validation approaches from using the *in situ* networks to model-based validations (Jackson et al., 2010a). The temporal stability analysis is one of the tentative methods for validation of SMAP products using *in situ* soil moisture networks. However, the temporal stability analysis has several limitations in determining spatially representative soil moisture values as was discussed in section 4.2. The CDF matching method has been shown to be a successful transforming (upscaling) method in this study and should be considered for the validation of remotely-sensed soil moisture, supplementing other validation methods such as the triple collocation method (Miralles et al., 2010).

Overall, this study examines the capability of different upscaling methods at the agricultural field scale for soil moisture measured in limited time periods (summer periods for two different years). Despite the significant findings of this study, several science questions remain unanswered and are the subject for future research: 1) How can we expand the upscaling approach to watershed scale for the validation of SMAP products at various resolutions? What other factors should be considered for watershed level upscaling?, 2) Are the observation operators spatially transferable to areas beyond the field scale?, 3) How does the length of the temporal sampling affect the observation operators? How long of an observation period is required to obtain temporally

transferable observation operators if they exist?, and 4) Can we separate the effect of crop types on the soil moisture variability from the effect of rainfall characteristics?. These are only some of the questions that may be answered with increased availability of observational data in the near future.

CHAPTER 5 SYNTHESIS

5.1 Soil Moisture Data Assimilation at Field and Watershed Scale

Surface soil moisture observations are assimilated in this dissertation through the Ensemble Kalman Filter with two different hydrologic models at two different scales: the one dimensional physically based model (the Root Zone Water Quality Model) at field scale, and semi-distributed hydrologic model (the Soil and Water Assessment Tool) at watershed scale. The assimilation of surface soil moisture observation with the hydrologic models results in significant improvement of soil moisture estimation in surface or upper soil layers and less improvement for predictions in deeper layers. Overall, the results presented in this work can be utilized to improve applications of soil moisture data assimilation at field and watershed scales and demonstrate the potential of soil moisture data assimilation to advance water quality or crop yield simulations by improving estimates of other hydrological variables such as ET and surface runoff, as well as, profile soil moisture content.

5.2 Estimation of System and Observation Error Statistics for Data Assimilation

Although many previous studies have demonstrated the ability of the EnKF to improve hydrologic predictions, especially profile soil moisture estimation, characterizing *a prior*

model and observation error statistics still remains as a challenge with real observations in using the EnKF (Ni-Meister, 2008). In case of the field scale data assimilation with *in situ* point measurements, the observation errors can be reasonably estimated due to relatively high accuracy of *in situ* measurements. However, it is more complicated to estimate *a prior* observation and model error statistics when the dimension of state variables increases with distributed models. A couple of land data assimilation studies showed that poor estimates of uncertainties in observations and model predictions may degrade the assimilation results compared to the model estimates or the observation alone (Reichle and Koster, 2002; Crow and Van Loon, 2006). In this dissertation, the watershed scale data assimilation study in Chapter 3 avoids this issue by conducting a synthetic experiment with the assumption of known error statistics. However, future study with real observations needs to consider how to determine spatially and temporally dynamic error statistics of observations and model predictions.

5.3 Issues with Soil Moisture Data Assimilation

One of the critical issues in assimilating real observed soil moisture data is bias treatment because the success of data assimilation cannot be achieved without an appropriate bias correction (Reichle, 2008). Systematic bias due to the persistent differences between observations and model predictions should be removed before assimilation, by rescaling the observations into the model climatology (Drusch et al., 2005b). The field scale data assimilation study in this dissertation also showed the systematic bias between model predictions and observations, especially during very dry periods. If a rescaling or bias

correction method had been applied in this study, the results would have been more successful.

The bounded nature of soil moisture between wilting point and saturated soil content also causes skewed model ensembles and unintended bias mostly when the soil is very dry or wet, which impacts negatively on the EnKF's performance (Crow and Wood, 2003; Ryu et al., 2009). Further investigations are required to overcome the non-Gaussian error structure and bias issues for successful soil moisture data assimilation studies.

5.4 Importance of Accurate Forcing Information

The fact that antecedent soil moisture condition is critical for accurate runoff prediction has led many previous studies to focus on improving antecedent soil moisture condition through soil moisture data assimilation (Aubert et al., 2003b; Komma et al., 2008; Brocca et al., 2009). However, the SWAT data assimilation study in this dissertation emphasizes the importance of having more accurate precipitation information rather than improving the initial soil moisture condition because large precipitation error can overwhelm the improvement in the soil moisture estimates through data assimilation. This problem was also mentioned in Crow and Ryu (2009). Securing as accurate precipitation as possible or overcoming the precipitation errors through soil moisture retrievals (Crow and Bolten, 2007; Crow et al., 2009; Crow et al., 2011) will lead to more successful runoff or streamflow predictions from the soil moisture data assimilation.

5.5 Improvement of Rainfall Runoff Mechanism

Successful updates of soil moisture condition may not always results in significant improvement in streamflow predictions. As the SWAT data assimilation study in this dissertation showed, neither the SCS Curve Number method nor the Green Ampt method implemented in SWAT could successfully reflect the soil moisture update through data assimilation and thus enhance streamflow predictions. This finding suggests that more efforts should be made for developing better numerical implementation of the rainfall-runoff mechanism in parallel with advancing the data assimilation scheme itself.

5.6 Linking *In situ* Observations with Remotely-sensed Soil Moisture

In this dissertation, upscaling of point soil moisture measurements is constrained at the field scale. However, to overcome the fundamental scale issues between *in situ* measurement and remotely-sensed products, it is necessary to expand the upscaling effort at least to watershed scale. Currently there are twelve soil moisture monitoring sites installed within the Upper Cedar Creek Watershed. The question on how to use those limited point measurements to obtain watershed representative soil moisture values would involve relating more complicated factors such as the spatial difference in topography and precipitation in future analysis.

Even though the upscaling study in Chapter 4 demonstrates that the CDF matching method can successfully transform the point measurement to field averages, the method is not found to be as temporally transferable. This indicates that a temporally varying

transforming method is still required. Currently soil moisture observation from summer periods of two years are used for this study, but it would also be desirable to investigate in the future how the length of the temporal sampling affects the upscaling results when soil moisture observations are available for longer time periods.

LIST OF REFERENCES

LIST OF REFERENCES

Abrahamson, D.A. et al., 2006. Evaluation of the RZWQM for simulating tile drainage and leached nitrate in the Georgia piedmont. Agronomy Journal, 98(3): 644-654.

Abrahamson, D.A. et al., 2005. Calibration of the root zone water quality model for simulating tile drainage and leached nitrate in the Georgia Piedmont. Agronomy Journal, 97(6): 1584-1602.

Ahuja, L., Rojas, K.W., Hanson, J.D., Shaffer, M.J. and Ma, L., 2000. Root Zone Water Quality Model: Modeling Management Effects on Water Quality and Crop Production. Water Resources Publications LLC, Highlands Ranch, CO.

Anagnostou, E.N., Negri, A.J. and Adler, R.F., 1999. Statistical adjustment of satellite microwave monthly rainfall estimates over Amazonia. Journal of Applied Meteorology, 38(11): 1590-1598.

Atlas, D., Rosenfeld, D. and Wolff, D.B., 1990. Climatologically tuned reflectivity-rain relations and links to area-time integrals. Journal of Applied Meteorology, 29(11): 1120-1135.

Aubert, D., Loumagne, C. and Oudin, L., 2003a. Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall-runoff model. Journal of Hydrology, 280(1-4): 145-161.

Aubert, D., Loumagne, C.i. and Oudin, L., 2003b. Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall-runoff model. Journal of Hydrology, 280(1-4): 145-161.

Baup, F., Mougin, E., Rosnay, P.d., Timouk, F. and Chênerie, I., 2007. Surface soil moisture estimation over the AMMA Sahelian site in Mali using ENVISAT/ASAR data. Remote Sensing of Environment, 109: 473–481.

Bolten, J., Crow, W., Zhan, X., Reynolds, C. and Jackson, T., 2009. Assimilation of a satellite-based soil moisture product into a two-layer water balance model for a global crop production decision support system. In: S.K. Pard (Editor), Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications. Springer-Verlang, London, United Kingdom, pp. 449-464.

Borah, D.K. and Bera, M., 2003. SWAT model background and application reviews, 2003 ASAE Annual International Meeting, Las Vegas, Nevada.

Brocca, L., Melone, F., Moramarco, T. and Morbidelli, R., 2010. Spatial-temporal variability of soil moisture and its estimation across scales. Water Resources Research, 46.

Brocca, L., Melone, F., Moramarco, T. and Singh, V.P., 2009. Assimilation of Observed Soil Moisture Data in Storm Rainfall-Runoff Modeling. Journal of Hydrologic Engineering, 14(2): 153-165.

Brocca, L., Morbidelli, R., Melone, F. and Moramarco, T., 2007. Soil moisture spatial variability in experimental areas of central Italy. Journal of Hydrology, 333(2-4): 356-373.

Burgers, G., van Leeuwen, P.J. and Evensen, G., 1998. Analysis scheme in the Ensemble Kalman Filter. Monthly Weather Review, 126: 1719.

Cameira, M.R., Ahuja, L., Fernando, R.M. and Pereira, L.S., 2000. Evaluating field measured soil hydraulic properties in water transport simulations using the RZWQM. Journal of Hydrology, 236(1-2): 78-90.

Cameira, M.R., Fernando, R.M., Ahuja, L. and Pereira, L., 2005. Simulating the fate of water in field soil-crop environment. Journal of Hydrology, 315(1-4): 1-24.

Charpentier, M.A. and Groffman, P.M., 1992. Soil moisture variability within remote sensing pixels. Journal of Geophysical Research-Atmospheres, 97(D17): 18987-18995.

Chen, F., Crow, W.T., Starks, P.J. and Moriasi, D.N., 2011. Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture. Advances in Water Resources, 34(4): 526-536.

Choi, M. and Jacobs, J.M., 2008. Temporal variability corrections for Advanced Microwave Scanning Radiometer E (AMSR-E) surface soil moisture: Case study in Little River Region, Georgia, U. S. Sensors, 8(4): 2617-2627.

Choi, M. and Jacobs, J.M., 2011. Spatial soil moisture scaling structure during Soil Moisture Experiment 2005. Hydrological Processes, 25(6): 926-932.

Chung, W.H., Wang, I.T. and Wang, R.Y., 2010. Theory-based SCS-CN method and its applications. Journal of Hydrologic Engineering, 15(12): 1045-1058.

Clark, M.P. et al., 2008. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. Advances in Water Resources, 31: 1309-1324.

Cosh, M.H., Jackson, T.J., Bindlish, R. and Prueger, J.H., 2004. Watershed scale temporal and spatial stability of soil moisture and its role in validating satellite estimates. Remote Sensing of Environment, 92(4): 427-435.

Cosh, M.H., Jackson, T.J., Starks, P. and Heathman, G., 2006. Temporal stability of surface soil moisture in the Little Washita River watershed and its applications in satellite soil moisture product validation. Journal of Hydrology, 323(1-4): 168-177.

Crosson, W.L., Laymon, C.A., Inguva, R. and Schamschula, M.P., 2002. Assimilating remote sensing data in a surface flux-soil moisture model. Hydrological Processes, 16(8): 1645-1662.

Crow, W.T. and Bolten, J.D., 2007. Estimating precipitation errors using spaceborne surface soil moisture retrievals. Geophysical Research Letters, 34(8).

Crow, W.T., Huffman, G.J., Bindlish, R. and Jackson, T.J., 2009. Improving Satellite-Based Rainfall Accumulation Estimates Using Spaceborne Surface Soil Moisture Retrievals. Journal of Hydrometeorology, 10(1): 199-212.

Crow, W.T. and Ryu, D., 2009. A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals. Hydrology and Earth System Sciences, 13(1): 1-16.

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

Crow, W.T. and van den Berg, M.J., 2010. An improved approach for estimating observation and model error parameters in soil moisture data assimilation. Water Resources Research, 46.

Crow, W.T., van den Berg, M.J., Huffman, G.J. and Pellarin, T., 2011. Correcting rainfall using satellite-based surface soil moisture retrievals: The Soil Moisture Analysis Rainfall Tool (SMART). Water Resources Research, 47.

Crow, W.T. and Van Loon, E., 2006. Impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture. Journal of Hydrometeorology, 7(3): 421-432.

Crow, W.T. and Wood, E.F., 2003. The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using Ensemble Kalman filtering: a case study based on ESTAR measurements during SGP97. Advances in Water Resources, 26(2): 137-149.

Das, N.N., Mohanty, B.P., Cosh, M.H. and Jackson, T.J., 2008. Modeling and assimilation of root zone soil moisture using remote sensing observations in Walnut Gulch Watershed during SMEX04. Remote Sensing of Environment, 112(2): 415-429.

De Lannoy, G.J.M., Houser, P.R., Pauwels, V.R.N. and Verhoest, N.E.C., 2007a. State and bias estimation for soil moisture profiles by an ensemble Kalman filter: Effect of assimilation depth and frequency. Water Resources Research, 43.

De Lannoy, G.J.M., Houser, P.R., Verhoest, N.E.C. and Pauwels, V.R.N., 2009. Adaptive soil moisture profile filtering for horizontal information propagation in the independent column-based CLM2.0. Journal of Hydrometeorology, 10(3): 766-779.

De Lannoy, G.J.M., Houser, P.R., Verhoest, N.E.C., Pauwels, V.R.N. and Gish, T.J., 2007b. Upscaling of point soil moisture measurements to field averages at the OPE3 test site. Journal of Hydrology, 343(1-2): 1-11.

De Lannoy, G.J.M., Reichle, R.H., Houser, P.R., Pauwels, V.R.N. and Verhoest, N.E.C., 2007c. Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter. Water Resources Research, 43.

DeLiberty, T.L. and Legates, D.R., 2003. Interannual and seasonal variability of modelled soil moisture in Oklahoma. International Journal of Climatology, 23(9): 1057-1086.

Draper, C.S., Mahfouf, J.F. and Walker, J.P., 2009. An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme. Journal of Geophysical Research - Atmospheres, 114.

Drusch, M., Wood, E.F. and Gao, H., 2005a. Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture. Geophysical Research Letters, 32(15).

Drusch, M., Wood, E.F. and Gao, H., 2005b. Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture. Geophysical Research Letters, 32.

Engman, E.T., 1991. Applications of microwave remote sensing of soil moisture for water resources and agriculture. Remote Sensing of Environment, 35(2-3): 213-226.

Entekhabi, D., Nakamura, H. and Njoku, E.G., 1994. Solving the inverse problems for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations. IEEE Transactions on Geoscience and Remote Sensing, 32(2): 438-448.

Entekhabi, D. et al., 2008. The Soil Moisture Active/Passive Mission (SMAP), Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International, pp. III - 1-III - 4.

Entekhabi, D. et al., 2010. The Soil Moisture Active Passive (SMAP) Mission. Proceedings of the Ieee, 98(5): 704-716.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research - Oceans, 99: 10143-10162.

Evensen, G., 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dynamics, 53: 343-367.

Evensen, G., 2004. Sampling strategies and square root analysis schemes for the EnKF. Ocean Dynamics, 54: 539-560.

Evensen, G. and Leeuwen, P.J.v., 1996. Assimilation of Geosat altimeter data for the Agulhas current using the ensemble Kalman filter with a quasigeostrophic model. Monthly Weather Review, 124(1): 85.

Famiglietti, J.S. et al., 1999. Ground-based investigation of soil moisture variability within remote sensing footprints during the Southern Great Plains 1997 (SGP97) Hydrology Experiment. Water Resources Research, 35(6): 1839-1851.

Farahani, H.J., Buchleiter, G.W., Ahuja, L.R., Peterson, G.A. and Sherrod, L.A., 1999. Seasonal evaluation of the root zone water quality model in Colorado. Agronomy Journal, 91(2): 212-219.

Flanagan, D., Huang, C.H., Pappas, E., Smith, D. and Heathman, G., 2008. Assessing conservation effects on water quality in the St. Joseph River Watershed, Proceedings of the AgroEnviron 2008 (Sixth International Symposium AgroEnviron), Antalya, Turkey, pp. 12.

Garen, D.C. and Moore, D.S., 2005. Curve number hydrology in water quality modeling: uses, abuses, and future directions. Journal of the American Water Resources Association, 41: 377-388.

Gassman, P.W., Reyes, M.R., Green, C.H. and Arnold, J.G., 2007. The Soil and Water Assessment Tool: Historical development, applications, and future research directions Transactions of the ASABE 50(4): 1211-1250.

Gelb, A. (Editor), 1974. Applied optimal estimation. The MIT Press, 374 pp.

Grayson, R.B. and Western, A.W., 1998. Towards areal estimation of soil water content from point measurements: time and space stability of mean response. Journal of Hydrology, 207(1-2): 68-82.

Han, E., Merwade, V. and Heathman, G.C., 2011. Application of data assimilation with the root zone water quality model for soil moisture profile estimation in the upper Cedar Creek, Indiana. Hydrological Processes: in press.

Heathman, G., Cosh, M.H., Merwade, V. and Han, E., *in review*. Multi-scale temporal stability analysis of surface and subsurface soil moisture within the Upper Cedar Creek Watershed, Indiana. Catena.

Heathman, G.C. et al., *in press*. Field scale spatiotemporal analysis of surface soil moisture for evaluating point-scale in situ networks. Geoderma.

Heathman, G.C., Larose, M., Cosh, M.H. and Bindlish, R., 2009. Surface and profile soil moisture spatio-temporal analysis during an excessive rainfall period in the Southern Great Plains, USA. CATENA, 78(2): 159-169.

Heathman, G.C., Starks, P.J., Ahuja, L.R. and Jackson, T.J., 2003a. Assimilation of surface soil moisture to estimate profile soil water content. Journal of Hydrology, 279: 1-17.

Heathman, G.C., Starks, P.J. and Brown, M.A., 2003b. Time domain reflectometry field calibration in the Little Washita River Experimental Watershed. Soil Science Society of America Journal, 67(1): 52-61.

Hoeben, R. and Troch, P.A., 2000. Assimilation of active microwave observation data for soil moisture profile estimation. Water Resources Research, 36.

Houser, P.R. et al., 1998. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. Water Resources Research, 34(12): 3405-3420.

Houtekamer, P.L. and Mitchell, H.L., 1998. Data assimilation using an Ensemble Kalman Filter technique. Monthly Weather Review, 126: 796.

Houtekamer, P.L. and Mitchell, H.L., 2001. A sequential ensemble Kalman Filter for atmospheric data assimilation. Monthly Weather Review, 129(1): 123.

Huang, C., Li, X., Lu, L. and Gu, J., 2008. Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter. Remote Sensing of Environment, 112: 888-900.

Jackson, T., Kimball, J., Colliander, A. and Njoku, E.G., 2010a. Soil Moisture Active and Passive (SMAP) Mission - Science Calibration and Validation Plan, SMAP Science Document, No:014. SMAP Science Document, No:014.

Jackson, T.J. et al., 2010b. Validation of Advanced Microwave Scanning Radiometer Soil Moisture Products. IEEE Transactions on Geoscience and Remote Sensing, 48(12): 4256-4272.

Jackson, T.J. et al., 2002. Soil Moisture Retrieval Using the C-Band Polarimetric Scanning Radiometer during the Southern Great Plains 1999 Experiment. IEEE Transactions on Geoscience and Remote Sensing, 40(10): 2151-2161.

Jackson, T.J., Hawley, M.E. and O'Neill, P.E., 1987. Preplanting soil moisture using passive microwave sensors. JAWRA Journal of the American Water Resources Association, 23: 11-19.

Jackson, T.J. and Schmugge, T.J., 1989. Passive microwave remote sensing system for soil moisture: some supporting research. IEEE Transactions on Geoscience and Remote Sensing, 27(2): 225-235.

Jackson, T.J. and Vine, D.E.L., 1996. Mapping surface soil moisture using a aircraft-based passive microwave instrument: algorithm and example. Journal of Hydrology, 184: 85-99.

Jackson, T.J., Vine, D.M.L., Swifi, C.T., Schmugge, T.J. and Schiebe, F.R., 1995. Large area mapping of soil moisture using the ESTAR passive microwave radiometer in Washita'92. Remote Sensing of Environment, 53: 27-37.

Jacobs, J.M., Mohanty, B.P., Hsu, E.-C. and Miller, D., 2004. SMEX02: Field scale variability, time stability and similarity of soil moisture. Remote Sensing of Environment, 92(4): 436-446.

Jacobs, J.M., Myers, D.A. and Whitfield, B.M., 2003. Improved rainfall/runoff estimates using remotelyl sensed soil moisture. Journal of the American Water Resources Association, 39: 313-324.

Jeong, J. et al., 2010. Development and integration of sub-hourly rainfall–runoff modeling capability within a watershed model. Water Resources Management, 24(15): 4505-4527.

Joshi, C., Mohanty, B.P., Jacobs, J.M. and Ines, A.V.M., 2011. Spatiotemporal analyses of soil moisture from point to footprint scale in two different hydroclimatic regions. Water Resources Research, 47.

Kaheil, Y.H., Gill, M.K., McKee, M., Bastidas, L.A. and Rosero, E., 2008. Downscaling and Assimilation of Surface Soil Moisture Using Ground Truth Measurements. Geoscience and Remote Sensing, IEEE Transactions on, 46(5): 1375-1384.

Kannan, N., Santhi, C., Williams, J.R. and Arnold, J.G., 2008. Development of a continuous soil moisture accounting procedure for curve number methodology and its behaviour with different evapotranspiration methods. Hydrological Processes, 22(13): 2114-2121.

Kannan, N., White, S.M., Worrall, F. and Whelan, M.J., 2007. Sensitivity analysis and identification of the best evapotranspiration and runoff options for hydrological modelling in SWAT-2000. Journal of Hydrology, 332(3-4): 456-466.

Keppenne, C.L., 2000. Data assimilation into a primitive-equation model with a parallel Ensemble Kalman Filter. Monthly Weather Review, 128: 1971.

Kerr, Y.H. et al., 2010. The SMOS Mission: New Tool for Monitoring Key Elements of the Global Water Cycle. Proceedings of the Ieee, 98(5): 666-687.

Kerr, Y.H. et al., 2001. Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission. Geoscience and Remote Sensing, IEEE Transactions on, 39(8): 1729-1735.

Kim, N.W. and Lee, J., 2008. Temporally weighted average curve number method for daily runoff simulation. Hydrological Processes, 22(25): 4936-4948.

King, Kevin W. , Arnold, J.G. and Bingner, R.L., 1999. Comparison of Green-Ampt and Curve Number methods on Goodwin Creek watershed using SWAT Transactions of the ASABE, 42(4): 919-926

Komma, J., Blőschl, G. and Reszler, C., 2008. Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting. Journal of Hydrology, 357: 228-242.

Kozak, J.A., Ahuja, L.R., Green, T.R. and Ma, L.W., 2007. Modelling crop canopy and residue rainfall interception effects on soil hydrological components for semi-arid agriculture. Hydrological Processes, 21(2): 229-241.

Lakhankar, T., Karakauer, N. and Khanbilvardi, R., 2009. Applications of microwave remote sensing of soil moisture for agricultural applications. International Journal of Terraspace Science and Engineering, 2(1): 81-91.

Li, J. and Islam, S., 1999. On the estimation of soil moisture profile and surface fluxes partitioning from sequential assimilation of surface layer soil moisture. Journal of Hydrology, 220(1-2): 86-103.

Liu, Y.Y. et al., 2011. Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. Hydrology and Earth System Sciences, 15(2): 425-436.

Loew, A. and Schlenz, F., 2011. A dynamic approach for evaluating coarse scale satellite soil moisture products. Hydrology and Earth System Sciences, 15(1): 75-90.

Ma, L. et al., 2001. Integrating system modeling with field research in agriculture: Applications of the root zone water quality model (RZWQM), Advances in Agronomy, Vol 71. Advances in Agronomy, pp. 233-292.

Ma, L., Ahuja, L.R. and Malone, R.W., 2007a. Systems modeling for soil and water research and management: Current status and needs for the 21st century. Transactions of the ASABE, 50(5): 1705-1713.

Ma, L. et al., 2007b. Sensitivity of tile drainage flow and crop yield on measured and calibrated soil hydraulic properties. Geoderma, 140(3): 284-296.

Ma, L. et al., 2007c. RZWQM simulation of long-term crop production, water and nitrogen balances in Northeast Iowa. Geoderma, 140(3): 247-259.

Ma, L. et al., 2003. Evaluation of RZWQM under varying irrigation levels in eastern Colorado. Transactions of the ASAE, 46(1): 39-49.

Margulis, S.A., McLaughlin, D., Entekhabi, D. and Dunne, S., 2002. Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment. Water Resources Research., 38.

Mattikalli, N.M., Engman, E.T., Ahuja, L.R. and Jackson, T.J., 1998. Microwave remote sensing of soil moisture for estimation of profile soil property. International Journal of Remote Sensing, 19(9): 1751-1767.

McLaughlin, D., 1995. Recent advances in hydrologic data assimilation. In U.S. National Report to the IUGG (1991-1994), Reviews of Geophysics, Supplement: 977-984.

Merlin, O., Al Bitar, A., Walker, J.P. and Kerr, Y., 2010. An improved algorithm for disaggregating microwave-derived soil moisture based on red, near-infrared and thermal-infrared data. Remote Sensing of Environment, 114(10): 2305-2316.

Michel, C., Andr?ssian, V. and Perrin, C., 2005. Soil Conservation Service Curve Number method: How to mend a wrong soil moisture accounting procedure? Water Resources Research, 41(2): W02011.

Miller, R.N., Ghil, M. and Gauthiez, F., 1994. Advanced data assimilation in strongly nonlinear dynamical systems. Journal of the Atmospheric Sciences, 51(8): 1037-1056.

Miralles, D.G., Crow, W.T. and Cosh, M.H., 2010. Estimating Spatial Sampling Errors in Coarse-Scale Soil Moisture Estimates Derived from Point-Scale Observations. Journal of Hydrometeorology, 11(6): 1423-1429.

Mishra, S.K. and Singh, V.P., 2006. A relook at NEH-4 curve number data and antecedent moisture condition criteria. Hydrological Processes, 20: 2755-2768.

Mohanty, B.P. and Skaggs, T.H., 2001. Spatio-temporal evolution and time-stable characteristics of soil moisture within remote sensing footprints with varying soil, slope, and vegetation. Advances in Water Resources, 24(9-10): 1051-1067.

Moradkhani, H., Sorooshian, S., Gupta, H.V. and Houser, P.R., 2005. Dual state-parameter estimation of hydrological models using ensemble Kalman filter, Advances in Water Resources, pp. 135-147.

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

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R. and King, K.W., 2002. Soil and water assessment tool. Theoretical documentation: Version 2005. TWRI TR-191. Texas Water Resources Institute, College Station, TX.

Ni-Meister, W., 2008. Recent Advances On Soil Moisture Data Assimilation. Physical Geography, 29(1): 19-37.

Ni-Meister, W., Houser, P.R. and Walker, J.P., 2006a. Soil moisture initialization for climate prediction: Assimilation of scanning multifrequency microwave radiometer soil moisture data into a land surface model. J. Geophys. Res., 111(D20): D20102.

Ni-Meister, W., Houser, P.R. and Walker, J.P., 2006b. Soil moisture initialization for climate prediction: Assimilation of scanning multifrequency microwave radiometer soil moisture data into a land surface model. Journal of Geophysical Research-Atmospheres, 111(D20): 15.

Njoku, E.G. and Entekhabi, D., 1996. Passive microwave remote sensing of soil moisture. Journal of Hydrology, 184: 101-129.

Njoku, E.G., Jackson, T.J., Lakshmi, V., Chan, T.K. and Nghiem, S.V., 2003. Soil moisture retrieval from AMSR-E. IEEE Transactions on Geoscience and Remote Sensing, 41(2): 215-229.

Pauwels, V.R.N., Hoeben, R., Verhoest, N.E.C. and De Troch, F.i.P., 2001. The importance of the spatial patterns of remotely sensed soil moisture in the improvement of discharge predictions for small-scale basins through data assimilation. Journal of Hydrology, 251(1-2): 88-102.

Pielke, R., Niyogi, D., Otto, J.-C. and Dikau, R., 2010. The Role of Landscape Processes within the Climate System Landform - Structure, Evolution, Process Control. Lecture Notes in Earth Sciences. Springer Berlin / Heidelberg, pp. 67-85.

Rawls, W.J., Brakensiek, D.L. and Saxton, K.E., 1982. Estimation of soil-water properties. Transactions of the ASAE, 25(5): 1316-1320.

Reichle, R.H., 2008. Data assimilation methods in the Earth sciences. Advances in Water Resources, 31(11): 1411-1418.

Reichle, R.H. and Koster, R.D., 2002. Land data assimilation with the ensemble Kalman filter: assessing model error parameters using innovations. In: S.M. Hassanizadeh, R.J. Schotting, W.G. Gray and G.F. Pinder (Editors), Computational Methods in Water Resources, Vols 1 and 2, Proceedings. Developments in Water Science, pp. 1387-1394.

Reichle, R.H. and Koster, R.D., 2004a. Bias reduction in short records of satellite soil moisture. Geophysical Research Letters, 31(19).

Reichle, R.H. and Koster, R.D., 2004b. Bias reduction in short records of satellite soil moisture. Geophysical Research Letters, 31.

Reichle, R.H. and Koster, R.D., 2005. Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. Geophysical Research Letters, 32(2): 4.

Reichle, R.H., Koster, R.D., Dong, J.R. and Berg, A.A., 2004. Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation. Journal of Hydrometeorology, 5(3): 430-442.

Reichle, R.H. et al., 2007. Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). Journal of Geophysical Research - Atmospheres, 112.

Reichle, R.H., McLaughlin, D.B. and Entekhabi, D., 2002a. Hydrologic data assimilation with the Ensemble Kalman Filter. Monthly Weather Review, 130: 103-114.

Reichle, R.H., Walker, J.P., Koster, R.D. and Houser, P.R., 2002b. Extended versus Ensemble Kalman Filtering for land data assimilation. Journal of Hydrometeorology, 3: 728-740.

Robinson, D.A. et al., 2008. Soil moisture measurement for ecological and hydrological watershed-scale observatories: A review. Vadose Zone Journal, 7(1): 358-389.

Ryu, D., Crow, W.T., Zhan, X. and Jackson, T.J., 2009. Correcting unintended perturbation biases in hydrologic data assimilation. Journal of Hydrometeorology, 10(3): 734-750.

Sabater, J.M., Jarlan, L., Calvet, J.-C., Bouyssel, F. and De Rosnay, P., 2007. From near-surface to root-zone soil moisture using different assimilation techniques. Journal of Hydrometeorology, 8: 194-206.

Sahu, R.K., Mishra, S.K. and Eldho, T.I., 2010. An improved AMC-coupled runoff curve number model. Hydrological Processes, 24: 2834-2839.

Schmugge, T. and Jackson, T.J., 1994. Mapping surface soil moisture with microwave radiometers. Meteorology and Atmospheric Physics, 54: 213-223.

Schmugge, T.J., Kustas, W.P., Ritchie, J.C., Jackson, T.J. and Rango, A., 2002. Remote sensing in hydrology. Advances in water resources, 25: 1367-1385.

Scipal, K., Scheffler, C. and Wagner, W., 2005. Soil moisture-runoff relation at the catchment scale as observed with coarse resolution microwave remote sensing. Hydrology and Earth System Sciences, 9(3): 173-183.

Seuffert, G., Wilker, H., Viterbo, P., Drusch, M. and Mahfouf, J.F., 2004. The usage of screen-level parameters and microwave brightness temperature for soil moisture analysis. Journal of Hydrometeorology, 5(3): 516-531.

Seyfried, M.S., Grant, L.E., Du, E. and Humes, K., 2005. Dielectric loss and calibration of the Hydra Probe soil water sensor. Vadose Zone Journal, 4(4): 1070-1079.

Seyfried, M.S. and Murdock, M.D., 2004. Measurement of soil water content with a 50-MHz soil dielectric sensor. Soil Science Society of America Journal, 68(2): 394-403.

Starks, P.J., Heathman, G.C., Ahuja, L.R. and Ma, L., 2003. Use of limited soil property data and modeling to estimate root zone soil water content. Journal of Hydrology, 272(1-4): 131-147.

Starks, P.J., Heathman, G.C., Jackson, T.J. and Cosh, M.H., 2006. Temporal stability of soil moisture profile. Journal of Hydrology, 324(1-4): 400-411.

Topp, G.C. and Davis, J.L., 1985. Measurement of soil water content using time-domain reflectrometry (TDR): A field evaluation. Soil Science Society of America Journal, 49(1): 19-24.

Topp, G.C., Davis, J.L. and Annan, A.P., 1980. Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. Water Resources Research, 16(3): 574–582.

Troch, P.A., Paniconi, C. and McLaughlin, D., 2003. Catchment-scale hydrologic modeling and data assimilation. Advances in Water Resources, 26: 131-135.

USDA-SCS, 1972. SCS National Engineering Handbook, Section 4, Hydrology. Chapter 10, Estimation of Direct Runoff from Storm Rainfall. Hydrology. U.S. Department of Agriculture, Soil Conservation Service, Washington, D.C. pp. 10.1-10.24.

Vachaud, G., Desilans, A.P., Balabanis, P. and Vauclin, M., 1985. Temporal stability of spatially measured siol water probability density function. Soil Science Society of America Journal, 49(4): 822-828.

van Delft, G., El Serafy, G.Y. and Heemink, A.W., 2009. The ensemble particle filter (EnPF) in rainfall-runoff models. Stochastic Environmental Research and Risk Assessment, 23(8): 1203-1211.

Verstraeten, W.W., Veroustraete, F. and Feyen, J., 2008. Assessment of evapotranspiration and soil moisture content across different scales of observation. Sensors, 8(1): 70-117.

Verstraeten, W.W., Veroustraete, F., van der Sande, C.J., Grootaers, I. and Feyen, J., 2006. Soil moisture retrieval using thermal inertia, determined with visible and thermal spaceborne data, validated for European forests. Remote Sensing of Environment, 101(3): 299-314.

Wagner, W., Lemoine, G. and Rott, H., 1999. A method for estimating soil moisture from ERS scatterometer and soil data. Remote Sensing of Environment, 70(2): 191-207.

Walker, J.P. and Houser, P.R., 2001. A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations. Journal of Geophysical Research-Atmospheres, 106(D11): 11761-11774.

Walker, J.P. and Houser, P.R., 2005. Hydrologic data assimilation. In: U. Aswathanarayana (Editor), Advances in Water Science Methodologies Taylor & Francis, A.A. Balkema, The Netherlands, pp. 230.

Walker, J.P., Willgoose, G.R. and Kalma, J.D., 2001a. One-dimensional soil moisture profile retrieval by assimilation of near-surface measurements: A simplified soil moisture model and field application. Journal of Hydrometeorology, 2(4): 356-373.

Walker, J.P., Willgoose, G.R. and Kalma, J.D., 2001b. One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: a comparison of retrieval algorithms. Advances in Water Resources, 24: 631-650.

Walker, J.P., Willgoose, G.R. and Kalma, J.D., 2002. Three-dimensional soil moisture profile retrieval by assimilation of near-surface measurements: Simplified Kalman filter covariance forecasting and field application. Water Resources Research, 38(12).

Wang, X., Shang, S., Yang, W. and Melesse, A.M., 2008. Simulation of an agricultural watershed using an improved Curve Number method in SWAT. Transactions of the ASABE, 51(4): 1323-1339.

Weerts, A.H. and El Serafy, G.Y.H., 2006. Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models. Water Resources Research, 42(9): 17.

Weissling, B.P., Xie, H. and Murray, K.E., 2007. A multitemporal remote sensing approach to parsimonious streamflow modeling in a southcentral Texas watershed, USA. Hydrology and Earth System Sciences Discussions 4(1): 1-33.

Western, A.W., Grayson, R.B. and Bloschl, G., 2002. Scaling of soil moisture: A hydrologic perspective. Annual Review of Earth and Planetary Sciences, 30: 149-180.

Whitaker, J.S. and Hamill, T.M., 2002. Ensemble data assimilation without perturbed observations. Monthly Weather Review, 130(7): 1913-1924.

White, E.D. et al., 2010. Development and application of a physically based landscape water balance in the SWAT model. Hydrological Processes, 25(6): 915-925.

Wood, A.W., Maurer, E.P., Kumar, A. and Lettenmaier, D.P., 2002. Long-range experimental hydrologic forecasting for the eastern United States. Journal of Geophysical Research-Atmospheres, 107(D20).

Xie, X.H. and Zhang, D.X., 2010. Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. Advances in Water Resources, 33(6): 678-690.

Zhang, S., Li, H., Zhang, W., Qiu, C. and Li, X., 2006. Estimating the soil moisture profile by assimilating near-surface observations with the ensemble Kalman filter (EnKF). Advances in Atmospheric Sciences, 22(6): 936-945.

VITA

VITA

Eunjin Han was born in Youngdong, Republic of Korea (South Korea). She graduated with a B.S in Earth Environment and Construction Engineering in 2001 from Hanyang University, Seoul, Korea. She received her M.S. degree in Civil and Environmental Engineering from University of Illinois at Urbana-Champaign in 2004. She joined the graduate program in Civil Engineering at Purdue University in 2007 and received Doctor of Philosophy degree in December, 2011. She worked for Korea Environment Institute and Mine Reclamation Corporation in South Korea before joining the graduate program at Purdue University.