DOWNSCALING SATELLITE MICROWAVE OBSERVATIONS TO FACILITATE HIGH RESOLUTION HYDROLOGICAL MODELLING

DOWNSCALING SATELLITE MICROWAVE OBSERVATIONS TO FACILITATE HIGH RESOLUTION HYDROLOGICAL MODELLING

By KURT CHRISTOPHER KORNELSEN, B.Sc., B.Ed.

A Thesis Submitted to the School of Graduate Studies
and the School of Geography and Earth Science
in Partial Fulfillment of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

DOCTOR OF PHILOSOPHY (2015) McMaster University
Geography Hamilton, Ontario

TITLE: DOWNSCALING OF SATELLITE OBSERVATIONS TO FACILITATE HIGH RESOLUTION HYDROLOGICAL MODELLING

AUTHOR: KURT CHRISTOPHER KORNELSEN, B.Sc., B.Ed. (Brock University)

SUPERVISOR: Professor Paulin Coulibaly

NUMBER OF PAGES: xix, pp. 260, A-34, B-88

ABSTRACT

Soil moisture is an essential climate variable and provides critical state information for hydrological applications. The state of soil moisture influences the exchange of water and energy between the earth surface and the atmosphere, partitions infiltration and runoff, can limit the net primary productivity of a region and govern the dynamics of geochemical processes. Satellite observations can be used to provide information about this important variable but are often available at a scale that is far greater than most hydrological processes. The scope of the research presented in this dissertation was to identify practical methods to facilitate the use of coarse scale satellite soil moisture information in higher resolution hydrological and land-surface modelling applications. Research was primarily conducted in the Hamilton-Halton watershed of Southern Ontario, Canada, although other watersheds and datasets were periodically used in some chapters.

A comprehensive review was conducted on the use of high resolution soil moisture information for hydrological applications, and data assimilation was identified as the most common method for integrating soil moisture information into a hydrological model. It was also identified that most watersheds displayed the property of temporal persistence and that root-zone soil moisture was of greater importance than surface soil moisture (Appendix B). In light of this information, the focus of this research was the downscaling of soil moisture and brightness temperature (TB) observations from the Soil Moisture and Ocean Salinity (SMOS) passive microwave satellite.

Satellite observations are sensitive to surface soil moisture, while rootzone soil moisture provides the greatest benefit to hydrological and land surface applications. To overcome this discrepancy, artificial neural networks (ANN) were evaluated as a method to estimate rootzone soil moisture from surface observations that accounted for the known non-linearities of soil moisture processes. The ANN model was trained with a numerical soil moisture physics model and validated using *in situ* observations from the McMaster Mesonet and USDA SCAN sites. The ANN was capable of accurately

depicting the rootzone soil moisture based on its training data at multiple sites, but was limited when the temporal distribution of soil moisture at a particular site was considerably different than the training data. Therefore, with the appropriate training data, ANNs are a viable method for predicting rootzone soil moisture from surface observations such as those available from satellites.

To provide high resolution soil moisture information from coarse resolution satellite data, bias correction was proposed and evaluated as a downscaling method for both soil moisture and TB. Using *in situ* data from two well instrumented USDA watersheds and a hydrological land-surface scheme (HLSS), it was found that temporal evolution of both soil moisture and TB at fine scale (~1 km) could be well characterized by the temporal evolution of the coarse scale (~20 km) soil moisture and TB. The fine scale spatial distribution of soil moisture could be predicted with a high degree of skill by correcting the bias between the coarse and fine scale soil moisture/TB.

In studying the correction of biases, it was found that naïve application of bias correction methods could result in the introduction of multiplicative biases in the bias corrected dataset. The theoretical implications of this for a data assimilation system were discussed although not yet evaluated. A bootstrap resampling approach was evaluated as a solution to this problem and it was found that resampled data could result in a robust bias correction that eliminated additive bias in most instances while limiting the induction of multiplicative bias. This new method was found to significantly outperform the standard bias correction techniques.

ACKNOWLEDGEMENT

I would like to express my appreciation to my supervisor Dr. Paulin Coulibaly for his guidance, support and mentoring over the past years. Without his patience and guidance the work accomplished herein would not be possible.

I would also like to thank the other members of my supervisory committee Dr. J. Michael Waddington and Dr. M. Altaf Arain for their support and challenges which have improved my research.

I am also thankful to the following individuals who were willing to share of their time and knowledge to guide me in the modelling and data processing that occurred herein: J. Samuel, B. Davison, E. Soulis, F. Seglenieks, B. Tolson, A. Haghnegahdar, A. Berg, Y. Kerr and his team at CESBIO.

I would like to acknowledge Environment Canada, the European Space Agency, the European Centre for Medium-Range Weather Forecasting, the Canadian Foundation for Innovation, the Ontario Innovation Trust, the Hamilton Conservation Authority and Conservation Halton who provided data and models that were used throughout this work. The research conducted herein was made possible by financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) and Ontario Graduate Scholarship (OGS).

Last and foremost I would like to express my sincere gratitude to my wife Tonya, son Callen, family and friends for supporting and being patient with me during this endeavor. Without your love, support and encouragement I would have never taken the leap necessary to begin this Ph.D. or made it to a successful conclusion.

Thank you all.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACF	Autocorrelation Function
AIEM	Advanced Integral Equation Model
AMSR-E	Advanced Microwave Scanning Radiometer - EOS
ARS	USDA Agricultural Research Service
ASAR	Advanced Synthetic Aperture Radar
ASCAT	Advanced Scatterometer
ANN	Artificial Neural Network
BFGS	Broyden-Fletcher-Goldfarb-Shanno Training Algorithm
BR	Bayesian Regularization Training Algorithm
BT	Bootstrap
CaPA	Canadian Precipitation Analysis
CAL/VAL	Calibration / Validation
CaLDAS	Canadian Land Data Assimilation System
CDF	Cumulative Distribution Function
CDFM	CDF Matching
CLASS	Canadian Land Surface Scheme
CMEM	Community Microwave Emissions Model
CS	Campbell Scientific
DDS	Dynamically Dimensioned Search
DGG	Discrete Global Grid
DOY	Day of Year
DQX	SMOS Data Quality Index
ECMWF	European Centre for Medium-Range Weather Forecasting
ECOCLIMAP	Land Surface Parameter Database
EKF	Extended Kalman Filter
EnKF	Ensemble Kalman Filter
ESA	European Space Agency
EOS	Earth Observing System
FEM	Finite Element Method
FDTD	Finite Difference Time Domain
GCOS	Global Climate Observing System
GEM	Global Environmental Multi-scale Weather Prediction Model
GOM	Geometric Optics Model
GR	Governor Road Site
GRMDM	Generalized Refractive Mixing Dielectric Model
GRU	Grouped Response Unit
H	Horizontal
HH	HH Polarization OR Hamilton-Halton Watershed
HLSS	Hydrological Land Surface Scheme
HYDAT	Hydrological Database
HYDRUS	Numerical Soil Moisture Model
IEM	Integral Equation Model

ISEA	Icosahedral Snyder Equal Area
ISMN	International Soil Moisture Network
K1	Kelso 1 Site
K2	Kelso 2 Site
L1C	SMOS Level 1C Data Product
L2	SMOS Level 2 Retrieved Soil Moisture product
LAI	Leaf Area Index
LDAS	Land Data Assimilation System
LM	Levenburg-Marquartd Algorithm
L-MEB	L-Band Microwave Emission of the Biosphere Model
LR	Little River, GA, Experimental Watershed OR Linear Rescaling
LS	Least Residual Slope
LSM	Land Surface Model
LUT	Look Up Table
LW	Little Washita, OK, Experimental Watershed
MAP	Maximum posterior probability estimator
MBSDM	Mineralogically Based Spectroscopic Dielectric Model
MEC	Modélisation Environmentale Communautaire
MESH	MEC – Surface Hydrology
MIRAS	Microwave Interferometric Radiometer with Aperture Synthesis
MKF	Multi-scale Kalman Filter
MLP	Multi-Layer Perceptron
MODIS	Moderate-resolution Imaging Spectroradiometer
MOD15A3	MODIS 1 km LAI Data Product
MoM	Method of Moments
MSD	Mean Squared Difference
MV	Minimum Variance Bayesian Technique
NASA	National Aeronautic and Space Administration
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NSE	Nash-Sutcliffe Efficiency
OR	Orchard Site
PDF	Probability Density Function
PTF	Pedotranfer Function
RADARSAT	Radar Satellite
RFI	Radio Frequency Interference
RMS	Root Mean Squared
RMSE	Root Mean Squared Error
RMSD	Root Mean Squared Difference
RT	Radiative Transfer
SAR	Synthetic Aperture Radar
SCAN	Soil Climate and Analysis Network
SGP	Southern Great Plains
SMAP	Soil Moisture Active Passive

SMAT	Soil Moisture Assessment Tool
SMEX	Soil Moisture Experiments
SML2UDP	SMOS Level 2 Soil Moisture User Data Product
SMOS	Soil Moisture and Ocean Salinity
SPM	Small Perturbation Method
TB	Brightness Temperature
TDR	Time Domain Reflectometry
UNSODA	Soil Characteristic Database
USDA	U.S. Department of Agriculture
V	Vertical
VWC	Vegetation Water Content
WCM	Water Cloud Model

PREFACE

This Ph.D. thesis is organized in a sandwich style based on the following published and submitted articles:

Chapter 2: McMaster Mesonet Soil Moisture Dataset: Description and Spatio-temporal Variability Analysis by K.C. Kornelsen and P. Coulibaly, Hydrology and Earth Systems Science, 17(1), 1-18, doi: 10.5194/hess-17-1-2013, 2013. (With permission from publisher).

Chapter 3: Root-zone Soil Moisture Estimation using Data-Driven Methods by K.C. Kornelsen and P. Coulibaly, Water Resources Research, 50, WR014127, doi: 10.1002/2013WR014127, 2014. (With permission from publisher).

Chapter 4: Potential of Bias Correction for Downscaling Satellite Observations by K.C. Kornelsen, M.H. Cosh and P. Coulibaly, Journal of Geophysical Research, Submitted, Manuscript Number 2014JD022683.

Chapter 5: Downscaling Satellite Soil Moisture and Brightness Temperature for a Land-Surface Model using Bias Correction by K.C. Kornelsen, B. Davison and P. Coulibaly, IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing, Submitted, Manuscript Number JSTARS-2014-01030.

Chapter 6: Improved Bias Correction Methods for Satellite Soil Moisture Retrievals by K.C. Kornelsen and P. Coulibaly, Remote Sensing of Environment, Submitted, Manuscript Number RSE-D-14-01088.

Appendix A: Comparison of Interpolation, Statistical and Data-Driven Methods for Imputation of Missing Values in a Distributed Soil Moisture Dataset by K.C. Kornelsen and P. Coulibaly, Journal of Hydrologic Engineering, 19(1), 26-43, doi: 10.1061/(ASCE)HE.1943-5584.0000767, 2014. (With permission from publisher).

Appendix B: Advances in Soil Moisture Retrieval from Synthetic Aperture Radar and Hydrological Applications by K.C. Kornelsen and P. Coulibaly, Journal of Hydrology, 476, 460-489, doi: 10.1016/j.jhydrol.2012.10.044, 2013. (With permission from publisher).

All of the articles were prepared by Kurt C. Kornelsen. Dr. Paulin Coulibaly provided guidance on research planning and helped in manuscript preparation. Dr. Michael Cosh provided the data for Chapter 5 as well as providing some insights into occurring processes at the study sites based on his personal experience with the study areas. Bruce Davison provided the model forcing data, support for the MESH modelling system and guidance on some model programming for Chapter 6, however all model updating was done by K.C. Kornelsen. The work reported in this dissertation was undertaken from September 2010 to October 2014.

Chapter 1: Introduction to Soil Moisture and Satellite Observations and their use in Hydrological Applications

1.1 Soil Moisture

Soil moisture is defined as the amount of water stored in the unsaturated soil zone. Numerous studies have shown that the state, spatial and temporal distribution of soil moisture are key to understanding Earth's hydrological and energy cycle's (Seneviratne et al. 2010). Due to the fundamental importance of soil moisture in geophysical applications it has been identified as an *essential climate variable* for the Global Climate Observing System (GCOS; GCOS-107, 2006).

In the hydrological cycle, the soil moisture state controls the partition between infiltration and runoff and influences the potential rate of soil water uptake by vegetation. Through its control of plant water availability, soil moisture also controls the vegetation distribution and is a critical factor controlling net primary productivity over much of the planet (Seneviratne et al. 2010). Also by partitioning infiltration and runoff, the state of soil moisture has a strong influence on the flood potential of a particular precipitation event or snow-melt (Seneviratne et al. 2010).

1.1.1 High Resolution Soil Moisture

High resolution soil moisture is a potentially powerful source of information that can benefit hydrological and meteorological applications. These benefits will be realized by implicitly accounting for watershed heterogeneity as modeling advances naturally tend towards high resolution distributed modeling (Wood et al. 2011). Many studies have demonstrated the benefits that can be derived from the integration of soil moisture, particularly at high resolution, into hydrological applications. For example, realistic initialization of the model state can have a significant impact on precipitation and temperature forecasting at lead times extending to 30 days (Koster et al. 2010), and can explain 10 to 60% of total runoff prediction skill (Mahanama et al. 2008). Loew et al. (2009) demonstrated that the assimilation of soil moisture can compensate for noisy precipitation data, while others show how soil moisture data can compensate for weak model state interactions/variability such as latent energy flux patterns (Merlin et al. 2006) and can provide time dependent adjustments of model parameters (Mohanty et al. 2000). Also, the uncertainty introduced by global climate change brings into question the temporal stability of model parameters (Peel and Blöschl, 2011) which can be better calibrated by the incorporation of soil moisture data (Koren et al. 2008).

The benefits of soil moisture data have already been proven for many of the above applications using *in situ* measurements with poor spatial representativeness (Loew et al. 2009; Koren et al. 2008) or at coarse spatial resolution (Bolten et al. 2010; Walker and Houser, 2001), although some examples do exist as to the benefits of high resolution soil moisture derived from Synthetic Aperture Radar (SAR) (Pauwels et al. 2002). In many cases the volumetric accuracy of soil moisture is less important than the spatial heterogeneity (Bronstert and Bardossy, 1999) as the spatial patterns allow for the better identification of important patterns such as runoff source areas. In this way the benefit of high resolution soil moisture lies less in the retrieval of soil moisture itself, but in the identification of spatial patterns across a catchment with respect to soil moisture state (Parada and Liang, 2008).

1.1.2 Temporal Persistence

The concept of temporal stability was introduced by Vachaud et al. (1985) and states that soil moisture at different locations, while variable in time, exhibits stability of rank when compared to the mean of other surrounding locations. That is, when several sites are ranked based on soil moisture, while the mean soil moisture state may change, the ranking of any individual site with respect to the others remains relatively constant in time. This concept has implications in terms of soil moisture downscaling and many authors have noted soil moisture measurements displayed temporal stability at both the field and catchment scale (Chapter 2 herein; Vivoni et al. 2008). Despite the propensity for temporal persistence shown in various studies, the strength of the temporal stability remains difficult to predict by external indicators. Topography influences temporal stability, where flat areas have been found to have poor stability of temporal variations and on mountainous terrain time stable locations were located at mid elevations whereas high and low elevations were relatively less stable (Vivoni et al. 2008). Additional factors influencing temporal stability include soil characteristics (Jacobs et al. 2004), which change only slowly, and vegetation (Vivoni et al. 2008) which plays only a minor role in temporal persistence. Unfortunately, temporal persistence may be interrupted at different spatial and temporal scales due to disturbances such as rainfall or vegetation growth at large scales and overland flow or land management at small scales (Mohanty et al. 2000) although these impacts only last for a limited period of time (Chapter 2 herein).

Parada and Liang (2008) were able to assimilate coarse soil moisture data disaggregated with temporal ratios and found recovery of information lost by coarse soil moisture data alone. However, this advantage was lost following precipitation which disturbed the rank stability relationship (Parada and Liang, 2008). Notwithstanding that limitation, the use of temporal stability has hydrological merit but requires parameterization without *in situ* measurement to be useful at the watershed scale or larger.

1.2 Satellite Soil Moisture Observations

The microwave portion of the electromagnetic spectrum is well suited for the remote sensing of Earth surface conditions because the longer wavelengths are known to penetrate clouds and rain (Fung, 1994; Ulaby et al. 1981). The most common frequencies for the remote sensing of soil moisture are 3.9-5.75 GHz (C band) and the longer wavelength 0.39-1.55 GHz (L band). Microwave energy at L band is particularly well suited for the remote sensing of soil moisture because its long wavelength (~21 cm) has been found to easily penetrate short and moderate vegetation cover and is internationally protected and therefore, theoretically, supposed to be free of interference (Kerr et al. 2010).

The land surface interacts with microwave energy, both active and passive in a number of ways. Primarily of interest for soil moisture remote sensing, the dielectric conductivity of the soil is a major contributor to the amount of active microwave energy that is backscattered and passive microwave energy emitted (Ulaby et al. 1981). Microwave energy also directly interacts with both solid vegetation structures, such as stalks and branches, and leaves (Ulaby et al. 1981). Some of the emission, backscatter or attenuation of the microwave signal is determined by the physical structure of the vegetation and some by the vegetation water content (Wigneron et al. 2007). Similarly the roughness of the surface, with respect to wavelength, can have a strong influence on the microwave signal observed by the satellite (Verhoest et al. 2008). In the latter two cases, the influence of vegetation and roughness on the signal for a given frequency is influenced by the incidence angle of the satellite with the surface. At high incidence angles the microwave signal must pass through greater amounts of vegetation causing greater attenuation of the soil moisture signal and larger influence of the vegetation signal (Kerr et al. 2012; Wigneron et al. 2007). The fact that the microwave signal is strongly influenced by various, potentially

unrelated, phenomena makes the retrieval of soil moisture for single satellite observations an ill-posed problem (Appendix B herein).

Active microwave remote sensing consists of the use of synthetic aperture radars (SAR), such as the Canadian Radarsat program, which illuminate the surface with a radar pulse, which is then backscattered to the satellite for observation. This remote sensing approach provides high resolution information about the Earth's surface, but signal interference and multiple radar reflections result in a noisy signal. SAR systems are also limited by long revisit periods and high energy requirements that limit the operational cycle for each orbit. Active scatterometers such as the Advanced Scatteromoter (ASCAT; Bartalis et al. 2008) use a low energy microwave pulse, have a wider swath and shorter revisit time than SAR, but can only collect information at relative coarse resolution (~ 25 km). Rather than emitting a signal and waiting for the response, passive radiometers observe the microwave energy that is naturally emitted or reflected by the Earth's surface. This allows for higher radiometric accuracy and less interference with the microwave signal, but requires a considerably larger antenna, which poses a physical limitation for satellites (Kerr et al. 2010). The result is also a coarse (~ 40 km) spatial resolution, although the satellite revisit time can be a little as three days at the equator (Kerr et al. 2010). Despite the low resolution, the short revisit times of radiometers and scatterometers makes them better suited for operational hydrological applications than the SAR alternative. Therefore, the work in this dissertation focused on satellite observations from the European Space Agency's (ESA) Soil Moisture and Ocean Salinity (SMOS) mission which was launched in November of 2009 (Kerr et al. 2010) and is directly relevant to the NASA Soil Moisture Active Passive (SMAP) mission (Entekhabi et al. 2010) which is yet to be launched at the time of writing.

1.2.1 Introduction to the Soil Moisture and Ocean Salinity (SMOS) Mission

SMOS is the second Earth Explorer Mission funded by the ESA, designed to provide information for climate and hydrological applications in the realm of land surface soil moisture, vegetation density and ocean salinity. For a complete overview of the SMOS mission and the SMOS soil moisture retrieval algorithms the interested reader is referred to Kerr et al. (2010), Kerr et al. (2012) and the references therein. A brief overview is provided here for context within this dissertation.

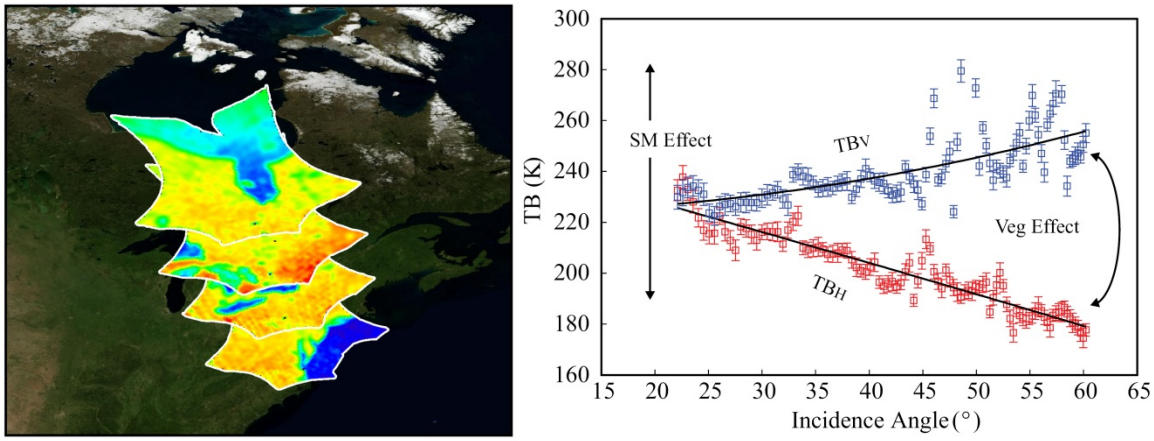

Figure 1-1: Series of overlapping SMOS observations (left) creating an angular profile of an individual SMOS DGG (right). The relative position and angle between the H and V polarization observations provides information for the soil moisture retrieval.

The SMOS satellite carries a single payload, the Microwave Interferometric Radiometer with Aperture Synthesis (MIRAS). MIRAS uses 69 small receivers positioned on the three SMOS arms to synthesize a large aperture radiometer that operates at a frequency of 1.4 GHz (L band) with an average ground resolution of 43 km (Kerr et al. 2010). Brightness temperature (TB) observations are made in both horizontal (TB_H) and vertical (TB_V) polarizations as well as the third and fourth stokes parameters. SMOS has a fixed tilt angle of 32.5° and an angular range of -10° to 60°. A two-dimensional 'image' of the Earth's surface is taken every 1.2 seconds resulting in multiple measurements at different incidence angles of the same area on the Earth. SMOS ascends across the equator at 0600h local solar time and descends across the equator at 1800h. Global coverage can be achieved every three days with shorter revisit times near the poles (Kerr et al. 2010).

SMOS Level 1 (L1C) multi-angular TB and Level 2 (L2) soil moisture retrievals are oversampled from the SMOS average 43 km resolution onto 15 km Icosahedral Snyder Equal Area (ISEA 4H9) projected grid. Soil moisture for the SMOS L2 data product is retrieved by minimizing a cost function that fits multi-angular SMOS L1C TB observations to match simulations from the L-band Microwave Emission of the Biosphere (L-MEB) model (Wigneron et al. 2007). The use of multi-angular snapshots results in solving a problem with more data than variables and thus eliminates the problem of ill-posed soil moisture retrievals. Figure 1-1 shows an example of concurrent SMOS snapshots forming a multi-angular observation. Generally, a change in soil moisture content would tend to shift both TB_H and TB_V curves vertically on the

figure and the vegetation effect is responsible for the divergence between TB_H and TB_V observations as incidence angle increases. Despite the relatively long wavelength of the L band frequency, penetration into the surface is shallow and the retrieved soil moisture is considered to be representative of the average conditions from the top 5 cm of the surface (Kerr et al. 2012).

Due to the difficulty of collecting distributed TB observations at the surface of the Earth, most SMOS calibration and validation (cal/val) efforts have focused on validation of L2 soil moisture products. SMOS retrievals have been validated in the United States (Al Bitar et al. 2012; Jackson et al. 2012; Collow et al. 2012), Canada (Gherboudj et al. 2012; Kornelsen and Coulibaly, 2014; Chapter 5 herein), Europe (Sanchez et al. 2012; Dall' Amico et al. 2012) and Australia (Su et al. 2013). SMOS soil moisture retrievals have been found to have good temporal correlation, but often contain persistent biases with respect to ground based measurements. Also, SMOS data have been found to have differing levels of performance between retrievals from ascending and descending half-orbits (Rowlandson et al. 2012; Collow et al. 2012; Chapter 5; Chapter 6). Possible reasons for the differences in retrievals between overpasses are discussed in Chapters 5 and 6.

1.2.2 Modelling Brightness Temperature

The problem of retrieving soil moisture from TB is complex as there are multiple models available for each component of the retrieval process and the proper selection of a model depends on frequency, surface conditions and information available. (See Appendix B for a similar and thorough discussion using synthetic aperture radar.) For this reason, the Community Microwave Emissions Model (CMEM; Drusch et al. 2009), which was used in Chapters 4 and 5 herein, was developed in a modular format. In order to constrain the discussion on the retrieval of soil moisture, the models presented herein are limited in scope to represent the state of the art retrieval model used by the SMOS L2 processor and the CMEM applications discussed in this thesis. It is based primarily on the application of L-MEB (Wigneron et al. 2007).

The general zeroth order radiative transfer equation for TB_p at the top of the atmosphere (TOA) for polarization p can be expressed as

$$TB_{TOA,tp} = TB_{au,p} + \exp\left(-\tau_{atm,p}\right) TB_{TOV,p} \tag{1-1}$$

where TB_{au} is the upwelling atmospheric TB, $-\tau_{atm}$ is the atmospheric optical depth calculated following Pellarin et al. (2003) and TB_{TOV} is the TB at the top of the vegetation layer (Drusch et

al. 2009). The TB observed by SMOS at the top of the atmosphere must also be corrected for Faraday rotation which results from the electron current in the ionosphere (Kerr et al. 2012). The microwave emissions at the top of the vegetation layer are the combined result of TB emissions from the soil TB_{soil}, which are attenuated by the vegetation layer, the vegetation emissions TB_{veg} which is attenuated by the canopy and reflected by the soil and the downwelling atmospheric contributions TB_{ad} which are reflected by the surface and attenuated by the vegetation canopy as

$$TB_{TOV} = TB_{soil,p}\exp\left(-\tau_{veg,p}\right) + TB_{veg,p}\left[1 + r_{r,p}\exp\left(-\tau_{veg,p}\right)\right] + TB_{ad,p}r_{r,p}\exp\left(-2\tau_{veg,p}\right) \quad (1\text{-}2).$$

The soil reflectivity term $r_{r,p}$ is determined as $r_{r,p} = 1 - e_{s,p}$ where $e_{s,p}$ is the surface emissivity and τ_{veg} is the vegetation optical depth which will be discussed later.

The microwave emissions from the soil TB_{soil} are governed by the effective temperature of the soil T_s and the soil emissivity and is generally expressed following Ulaby et al. (1981)

$$TB_{soil} = T_s \cdot e_{s,p} \quad (1\text{-}3).$$

Following the recommendation of de Rosnay and Wigneron (2006) the effective temperature is a function of the physical temperature of the soil at the surface and at depth (usully 0.5-1 m) as:

$$T_s = T_{soil\_depth} + C_t\left(T_{soil\_surf} - T_{soil\_depth}\right) \quad (1\text{-}4).$$

The parameter C_t depends on the frequency of observation and soil moisture content where

$$C_t = \min\left[\left(SM/\omega_0\right)^{b\omega 0}, 1\right] \quad (1\text{-}5)$$

where SM is volumetric soil moisture and ω_0 and $b\omega 0$ are parameters that depend on soil texture and structure (Kerr et al. 2012). The result is that effective temperature under dry conditions is represented by the maximum penetration depth and becomes more represented by surface soil temperature as moisture increases.

Soil reflectivity is modelled using the Fresnel Law for reflectivity of a specular surface

$$\left(r_p = r_p\left(\varepsilon_p, \theta\right)\right) \quad (1\text{-}6)$$

where ε_p is the dielectic constant of the soil based on the soil moisture content calculated using the formulation of Mironov et al. (2004) (See Appendix B for a discussion) and the incidence angle θ. Since most soil surfaces do not behave as specular surfaces the rough surface reflectivity is given based on a modification to Wang and Choudry's (1981) equation by Wigneron et al. (2007):

$$r_{r,p} = \left[Qr_{s,p} + (1-Q)r_{s,q}\right]\exp\left(-h(SM)\cos^N\theta\right) \quad (1\text{-}7)$$

where the rough surface reflectivity is determined by the specular reflectivity, the polarization *p* and the alternate polarization *q*, a polarization coupling factor *Q*, the roughness parameter *h* which is dependent on soil moisture and root mean square surface height (see Appendix B), and the incidence angle. At L-band the mixing factor *Q* is presumed to be equal to zero (Kerr et al. 2012). The result is that the reflectivity of the surface is determined by the theoretical reflection of a smooth surface with given dielectric properties, which is then attenuated by the satellite perceived roughness which includes the dielectric roughness caused by soil moisture.

The vegetation layer contribution is determined by the τ-ω approach where the vegetation has a direct contribution to TB and the vegetation canopy attenuates the signal from the soil based on its optical depth τ_{veg}. The vegetation contribution is modelled by

$$TB_{veg,p} = T_c\left(1-\omega_p\right)\left[1-\exp\left(-\tau_{veg,p}\right)\right] \tag{1-8}$$

where T_c is the canopy temperature and ω_p is the single scattering albedo at polarization p which is assumed constant for L-band at ω = 0.05 (Wigneron et al. 2007). For the purposes herein the canopy temperature is assumed to be approximately equal to the surface temperature. For ascending overpasses of SMOS in the early morning this assumption is not considered limiting because solar heating is not contributing to strong thermal gradients (Wigneron et al. 2007). In L-MEB the vegetation optical depth depends on the vegetation type and polarization

$$\tau_{veg,p} = \tau_{nadir}\left(\cos^2\theta + tt_p \sin^2\theta\right)\frac{1}{\cos\theta} \tag{1-9}$$

where tt_p is a parameter that represents the angular impact of optical depth. This parameter is set to one for horizontal polarization and is greater than one for vertical polarization to account for the vertical structures in vegetation which are important at L-band (Wigneron et al. 2007). The optical thickness at nadir is parameterized by

$$\tau_{nadir} = b'LAI + b'' \tag{1-10}$$

where LAI is the leaf area index and the parameters b' and b'' were calibrated based on vegetation structure and stored in look up tables (Wigneron et al. 2007).

1.3 Satellite Observations and Retrievals in Hydrological Applications

Soil moisture retrievals and satellite TB observations have been demonstrated to improve hydrological modelling applications by correcting errors in model forcing (Walker et al. 2001;

Crow and Wood, 2002), providing realistic initialization of weather and flood forecasts (Koster et al. 2010; Liu et al. 2012), and through updating the model state. Typically this information is used for model calibration (Pauwels et al. 2009), data assimilation (Reichle et al. 2008; Draper et al. 2012), or calibration and assimilation through dual-state-parameter assimilation techniques.

The assimilation of soil moisture information from both *in situ* sensors and satellite retrievals has been demonstrated to improve the accuracy of land surface models (Koster et al., 2010; Draper et al. 2012), numerical weather prediction (de Rosnay et al., 2013), hydrological models (Brocca et al. 2010; Liu et al. 2012) and flood forecasting (Liu et al. 2012). Unfortunately, most satellite observations of soil moisture have a spatial resolution of 25 - 40 km which is much larger than the resolution of many hydrological processes which occur at less than 1 km (Merlin et al., 2012). To overcome this impediment there has been a lot of attention placed on the downscaling of satellite observations.

1.3.1 Downscaling Satellite Observations and Soil Moisture Retrievals

Downscaling is a procedure where either a physical or statistical relationship is established between coarse scale soil moisture and fine scale predictor variables. At the fine scale, the spatial distribution of soil moisture can be considered as a function of an organized structure based on topography, vegetation and soil texture that is perturbed by stochastic influences such as precipitation, temperature, solar radiation, etc. (Wilson et al. 2005; Bronstert and Bardossy, 1999). Most practical downscaling methods presented seek to produce high resolution soil moisture fields by accounting for either the organized or stochastic component influencing soil moisture distribution.

Merlin et al. (2012) provide one of few examples in the literature of actual SMOS soil moisture data being disaggregated. Their method relies on MODIS derived surface skin temperature collected in the thermal infrared portion of the visible spectrum. The skin temperature data is related to NDVI values to derive the surface evaporative efficiency which is used as a fine scale proxy of surface soil moisture. A similar method is proposed by Kim and Hogue (2012) who use MODIS vegetation indices and surface temperature to disaggregate the AMSR-E soil moisture product. The method of Merlin et al. (2006; 2012) places more weight on the land surface temperature, whereas the UCLA method of Kim and Hogue (2012) weights both

vegetation and skin temperature evenly resulting in a model which represents a broader range of conditions, but is slightly less accurate (Kim and Hogue, 2012).

In order to account for the structured component of soil moisture, many proposed downscaling methodologies seek to relate topographic indices to coarse resolution soil moisture. Relying on the concept of temporal stability (Vachaud et al. 1985), Loew and Mauser (2008) developed a simple downscaling framework where fine scale soil moisture was related to spatially averaged soil moisture through simple linear regression using a physically based model to develop regression parameters. Using this technique each model grid cell was represented by a unique regression equation which maintained the grid cell's relative rank. This method, although simple, had an error of less than 4 % for almost every 1 km grid cell but was strongly influenced by bias and errors in the coarse resolution soil moisture (Loew and Mauser, 2008). Coleman and Niemann (2012) suggested that the soil moisture non-linearity could be overcome through the use of non-linear mixture models and a spatial artificial neural network. Both mixture models and spatial ANN are unsupervised machine learning methods that were used to relate topographic indices to spatially averaged soil moisture. They found these methods were able to predict spatial variability when soil moisture was highly organized and could capture some organization when soil moisture was less organized. In comparing these methods with multi-linear regression it was concluded that the spatial ANN was best suited to soil moisture downscaling as it was able to best account for the non-linearity of soil moisture scaling (Coleman and Niemann, 2012). In order to better represent the strength of the relationship between the soil moisture state and the topographic influences that affect soil moisture organization, Wilson et al. (2005) used a weighted combination of indices. Linking the weight factor of various topographic indices to the mean soil moisture state allowed the downscaling model to adapt as the relative dominance of various spatially different processes changed (Wilson et al. 2005). This method was enhanced by adding a map of average residuals to represent the soil moisture contribution which is persistent in time, but not related to terrain, where the residuals are interpreted to represent a map of soil properties (Wilson et al. 2005).

Mascaro et al. (2011) downscaled soil moisture using a multi-fractal statistical relationship with scale. They reduced soil type, topographic index, and land cover using principal components analysis (PCA) which was used to calculate the multi-fractal exponent of the scaling model for each statistical moment. The derived downscaling relationships were tested for four

different sites and were found to be site specific, reproducing the statistical properties of the observed soil moisture, but not the spatial properties (Mascaro et al. 2011).

In most of the downscaling examples above, the soil moisture or TB from the satellite source would maintain satellite biases with respect to a land surface model, and therefore bias correction would still be a necessary step (Merlin et al. 2012).

1.3.2 Data Assimilation

A land data assimilation system (LDAS) is designed to merge satellite observations or retrievals into a numerical model. Since both the model and observations are considered to contain errors, it is expected that the output from an LDAS is superior to either product alone (Reichle et al. 2014). In land surface applications this is most often accomplished with sequential assimilation techniques derived from the Kalman Filter such as the Extended Kalman Filter (EKF; de Rosnay et al. 2013) and the Ensemble Kalman Filter (EnKF; Evenson, 2003).

In the Kalman Filter the model is used to create an *a priori* estimate of the state known as the forecast using a Monte Carlo simulation. The observations are used to conduct an update of the model state based on Bayesian update of the *posterior* model states known as the analysis (Evenson, 2003). The update follows:

$$\hat{\mathbf{X}} = \mathbf{X} + \mathbf{K}(\mathbf{D} - \mathbf{H}\mathbf{X}) \tag{1-11}$$

where **X** is the forecast matrix, **D** contains the observation data plus a random vector of errors and **H** is the observation operator, which is used to transform the observed data in **D** to the model state space. The update is determined by the Kalman Gain **K** which is estimated by the sample error covariance matrices of model and observations. Because Kalman filter methods use model state covariance, the update from the assimilation of surface information can be propagated to deeper soil layers (Sabater et al. 2007). If soil moisture is being directly assimilated in the LDAS, then the observation operator often takes the form of a bias correction, which matches the statistical moments of the satellite retrieval to the model state. Currently, this is the most common form of LDAS (Lahoz and De Lannoy, 2014). Alternatively, a radiative transfer model, such as the Community Microwave Emissions Model (CMEM; Drusch et al. 2009) can be coupled to the land surface model and satellite observations indirectly assimilated to update the soil moisture state. Given simplifying assumptions necessary for the radiative transfer model, the correction of bias' between the satellite observations and model is still often necessary (Drusch

et al. 2009). The latter option has theoretical and practical advantages for LDAS because it avoids the use of potentially inconsistent ancillary data in soil moisture retrievals and TB observations are available in near real time (Lahoz and De Lannoy, 2014).

1.4 Rationale, Objectives and Outline

As previously discussed, soil moisture is a key component of the hydrological cycle. An improved soil moisture state is expected to improve many applications with particular emphasis on numerical weather prediction and streamflow forecasting. The goal of this work is to identify a practical method for the downscaling of satellite soil moisture retrievals and observations, with a particular focus on SMOS, for use with high resolution distributed modelling applications. This goal was accomplished through directed research on the role of spatial biases of soil moisture and brightness temperature using field data, a distributed hydrological land surface scheme and a radiative transfer model. The major assumptions used throughout much of the bias correction research were:

1) Watersheds can be considered temporally stable and therefore the spatial distribution of soil moisture at intermediate scales (~1 km) can be well described by bias and differences in variance from the mean watershed state;
2) At scales below that of a satellite grid cell, the temporal evolution of the land surface soil moisture is primarily driven by precipitation and evapotranspiration, which can be assumed to be spatially homogenous;
3) Most applications of satellite observations in hydrology involve sequential data assimilation and therefore bias correction is already a necessary step to integrate the satellite data.

The validity of assumptions (1) and (2) were verified in Chapters 4 and 5 and assumption (3) was established through precedent in the literature (Lahoz and De Lannoy, 2014; Reichle et al. 2014; Liu et al. 2012).

1.4.1 Thesis Outline

This thesis consists of seven chapters. Chapter 1 is an introduction to soil moisture in hydrology, satellite remote sensing of soil moisture, the correction of biases between satellite observations and retrievals, the modelling of surface soil moisture, and the role of bias correction in data assimilation. The introduction in Chapter 1 provides context and touches on research projects in Chapters 2-6.

Chapter 2 focuses on soil moisture hydrology and describing the data used in Chapter 5. Chapter 3 provides background to determine if a reliable system can be developed to predict root zone soil moisture based on information such as that used in Chapters 4-6. Chapters 4-6 are about the downscaling of soil moisture and TB with bias correction. Chapters 4 and 5 investigate bias correction using different types of data and Chapter 5 focuses on the bias correction method itself.

Chapter 2 presents the McMaster Mesonet distributed soil moisture dataset and a spatio-temporal analysis of the soil moisture data provided. A comparison was made between two different types of soil moisture probes and initial results from the meteorological stations of the McMaster Mesonet described. Temporal persistence in the Hamilton-Halton watershed was explored using the approach proposed by Vachaud et al. (1985) including a unique analysis of temporal persistence following a disturbance. An analysis was conducted of the spatial variability as it relates to the basin state, which is also related to results in Chapters 4-6. This work was published in the journal Hydrology and Earth Systems Science.

Chapter 3 explores the ability to predict rootzone soil moisture using artificial neural networks (ANN) and surface observations of soil moisture as may be available from satellite observations. The physically based model HYDRUS 1-D was forced with meteorological inputs from the McMaster Mesonet and database soil texture to generate surface and rootzone soil moisture. The HYDRUS 1-D results were used to train several ANN configurations and a sensitivity analysis conducted. These ANN output were validated using three stations from the McMaster Mesonet and three stations from the U.S. Department of Agriculture (USDA) Soil Climate Analysis Network (SCAN; Schaefer et al. 2007). This work was published in the journal Water Resources Research.

Chapter 4 investigates the feasibility of downscaling TB observations and soil moisture retrievals using a simple bias correction technique in two well instrumented USDA watersheds. CMEM was used to simulate TB; and the temporal persistence of TB and soil moisture was investigated. A simple bias correction technique was used to downscale watershed average TB and soil moisture to match point scale *in situ* measurements. This work has been submitted to Journal of Geophysical Research: Atmospheres.

Chapter 5 expands upon the work in Chapter 4. The physically based distributed Hydrological Land Surface Scheme (HLSS) Modélisation Environmentale Communautaire (MEC) – Surface Hydrology (MESH) was coupled to CMEM to provide distributed soil moisture and TB at high resolution. MESH was calibrated and validated in the Hamilton-Halton watershed using data from the McMaster Mesonet and Environment Canada. MESH-CMEM was independently validated using TB observations and soil moisture retrievals from SMOS. Bias correction was then assessed as a means to downscale SMOS observations to match a high resolution model as would be practiced in a LDAS. This work has been submitted to the IEEE Journal of Selected Topics on Applied Earth Observations and Remote Sensing.

Chapter 6 examines a shortcoming of the bias correction techniques described in Chapters 4 and 5 and suggested an improved method of bias correction. SMOS soil moisture retrievals were compared to several USDA SCAN sites distributed across the United States for two different years. Two bias correction techniques were examined and a Monte Carlo resampling technique proposed to provide robust bias correction. The benefits and shortcomings of this technique were analyzed. This work was submitted to the journal Remote Sensing of Environment.

Chapter 7 provides a summary of the conclusions of the previous chapters and provides recommendations for directions of future research.

Appendix A is a methodological paper that describes the infilling of soil moisture data. The methods described in the Appendix are applied in Chapters 2-4. This work was published in the Journal of Hydrologic Engineering.

Appendix B is a comprehensive review which can be broken into two main sections. The first section reviews the state of the art of soil moisture retrieval from SAR satellites. Based on the analysis of the first part of the chapter, the second part focuses on the hydrological applications of high resolution soil moisture, which provides the motivation for the remainder of

the thesis. The end of this chapter contains a thorough discussion of the merits and limitations of high resolution soil moisture retrievals, particularly from SAR satellites. This work was published in the Journal of Hydrology.

1.5 References

Al Bitar, A., Leroux, D., Kerr, Y. H., Merlin, O., Richaume, P., Sahoo, A., and Wood, E. (2012) Evaluation of SMOS soil moisture products over continental U.S. using the SCAN/SNOTEL network. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 1572–1586.

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J. and Anderson, C. (2008) Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT), *Geophysical Research Letters* 34, L20401.

Bolten, J. D., Crow, W. T., Xiwu Zhan, Jackson, T. J., and Reynolds, C. A. (2010) Evaluating the utility of remotely sensed soil moisture retrievals for operational agricultural drought monitoring. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3, 57-66.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., and Hasenauer, S. (2010) ASCAT soil wetness index validation through in situ and modeled soil moisture data in central Italy. *Remote Sensing of Environment*, 114, 2745–2755.

Bronstert, A., and Bardossy, A. (1999) The role of spatial variability of soil moisture for modelling surface runoff generation at the small catchment scale. *Hydrology and Earth System Sciences*, 3, 505-516.

Coleman, M. L., and Niemann, J. D. (2012) An evaluation of nonlinear methods for estimating catchment-scale soil moisture patterns based on topographic attributes. *Journal of Hydroinformatics*, 14, 800-814.

Collow, T.W., Robock, A., Basara, J.B., & Illsont, B.G. (2012). Evaluation of SMOS retrievals of soil moisture over the central United States with currently available in situ observations. *Journal of Geophysical Research*, 117, D09113.

Crow, W.T., and Wood, E.F. (2002) Impact of soil moisture aggregation on surface energy flux prediction during SGP'97. *Geophysical Research Letters*, 29, 1008.

Dall'Amico, J.T., Schlenz, F., Loew, A., and Mauser, W. (2012) First results of SMOS soil moisture validation in the upper Danube catchment. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 1507–1516.

de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., and Isaksen, L. (2013) A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, 139, 1199–1213.

Draper, C.S., Reichle, R.H., De Lannoy, G.J.M., and Liu, Q. (2012) Assimilation of passive and active microwave soil moisture retrievals. *Geophysical Research Letters*, 39, L04401.

Drusch, M., Holmes, T., de Rosnay, P. and Balsamo G. (2009) Comparing ERA-40 based L-band brightness temperatures with Skylab observations: A calibration / validation study using the Community Microwave Emission Model, *Journal of Hydrometeorology* 10, 213-226.

Entekhabi, D., Njoku, E.G. , O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, J.K, Moghaddam, Moran, M. S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L. and van Zyl, J. (2010) The Soil Moisture Active Passive (SMAP) Mission. *Proceedings of the IEEE*, 98, 704–716.

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

Fung, A.K., 1994. Microwave Scattering and Emission Models and Their Applications. Norwood, MA: Artech House Inc.

GCOS-107 (2006), Systematic observation requirements for satellite-based products for climate. Supplemental details to the satellite-based component of the "implementation plan for the global observing system for climate in support of the UNFCCC", GCOS-107, WMO/TD No. 1338, Sept 2006.

Gherboudj, I., Magagi, R., Goïta, K., Berg, A.A., Toth, B., and Walker, A. (2012) Validation of SMOS data over agricultural and Boreal forest areas in Canada. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 1623-1635.

Jackson, T. J., Bindlish, R., Cosh, M. H., Tianjie, Z., Starks, P. J., Bosch, D.D., Seyfried, M., Moran, M.S., Goodrich, D.C., Kerr, Y.H. and Leroux, D. (2012) Validation of Soil Moisture and Ocean Salinity (SMOS) soil moisture over watershed networks in the U.S. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 1530–1543.

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.-J., Font, J. , Reul, N., Gruhier, C., Juglea, S.E., Drinkwater, M.R., Hahne, A., Martin-Neira, M. and Mecklenburg, S. (2010) The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proceedings of the IEEE*, 98, 666–687.

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J.-P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S. E., Leroux, D., Mialon, A. and Delwart, S. (2012) The SMOS soil moisture retrieval algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 1384-1403.

Kim, J. and Hogue, T.S. (2012) Improving spatial soil moisture representation through integration of AMSR-E and MODIS products, *IEEE Transactions on Geoscience and Remote Sensing*, 50, 446-460.

Koren, V., Moreda, F., & Smith, M. (2008) Use of soil moisture observations to improve parameter consistency in watershed calibration. *Physics and Chemistry of the Earth,* 33, 1068-1080.

Kornelsen, K.C. and Coulibaly, P. (2014) Data-based disaggregation of SMOS soil moisture, *Proceedings of IGARSS, the IEEE Geoscience and Remote Sensing Symposium*, Quebec City, Canada, 3334-3337.

Koster, R.D., Mahanama, S.P.P., Yamada, T.J., Balsamo, G., Berg, A.A., Boisserie, M., Dirmeyer, P.A., Doblas-Reyes, F.J., Drewitt, G., Gordon, C.T., Guo, Z., Jeong, J.-H., Lawrence, D.M., Lee, W.-S., Li, Z., Luo, L., Malyshev, S., Merryfield, W.J., Seneviratne, S.I., Stanelle, T., van den Hurk, B.J.J.M., Vitart, F. and Wood, E.F. (2010). Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophysical Research Letters,* 37, L02402.

Lahoz, W.A., and De Lannoy, G.J.M. (2014) Closing the gaps in our knowledge of the hydrological cycle over land: Conceptual problems, *Surveys in Geophysics*, 35, 623-660.

Liu, Y., Weerts, A.H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., Seo, D.-J., Schwanberg, D., Smith, P., van Dijk, A.I.J.M., van Velzen, N., He, M., Lee, H., Noh, S.J., Rakovec, O., and Restrepo, P. (2012) Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. *Hydrology and Earth Systems Science,* 16, 3836-3887.

Loew, A., and Mauser, W. (2008) On the disaggregation of passive microwave soil moisture data using a priori knowledge of temporally persistent soil moisture fields. *IEEE Transactions on Geoscience and Remote Sensing,* 46, 819-834.

Loew, A., Schwank, M., and Schlenz, F. (2009) Assimilation of an L-band microwave soil moisture proxy to compensate for uncertainties in precipitation data. *IEEE Transactions on Geoscience and Remote Sensing,* 47, 2606-2616.

Mahanama, S.P.P., Koster, R.D., Reichle, R.H., and Zubair, L. (2008) The role of soil moisture initialization in subseasonal and seasonal streamflow prediction - A case study in Sri Lanka. *Advances in Water Resources*, 31, 1333-1343.

Mascaro, G., Vivoni, E.R., and Deidda, R. (2011) Soil moisture downscaling across climate regions and its emergent properties. *Journal of Geophysical Research-Atmospheres*, 116, D22114.

Merlin, O., Chehbouni, A., Boulet, G., & Kerr, Y. (2006). Assimilation of disaggregated microwave soil moisture into a hydrologic model using coarse-scale meteorological data. *Journal of Hydrometeorology,* 7, 1308-1322.

Merlin, O., Rüdiger, C., Al Bitar, A., Richaume, P., Walker, J. P., and Kerr, Y. H. (2012). Disaggregation of SMOS soil moisture in southeastern Australia. *IEEE Transactions on Geoscience and Remote Sensing,* 50, 1556-1571.

Mohanty, B.P., Skaggs, T.H., and Famiglietti, J. S. (2000) Analysis and mapping of field-scale soil moisture variability using high-resolution, ground-based data during the southern great plains 1997 (SGP97) hydrology experiment. *Water Resources Research,* 36, 1023-1031.

Parada, L.M., and Liang, X. (2008) Impacts of spatial resolutions and data quality on soil moisture data assimilation. *Journal of Geophysical Research-Atmospheres,* 113, D10101.

Pauwels, V.R.N., Hoeben, R., Verhoest, N.E.C., De Troch, F.P., and Troch, P.A. (2002) Improvement of TOPLATS-based discharge predictions through assimilation of ERS-based remotely sensed soil moisture values. *Hydrological Processes,* 16, 995-1013.

Pauwels, V.R.N., Balenzano, A., Satalino, G., Skriver, H., Verhoest, N., and Mattia, F. (2009) Optimization of soil hydraulic model parameters using synthetic aperture radar data: An integrated multidisciplinary approach. *IEEE Transactions on Geoscience and Remote Sensing,* 47, 455-457.

Peel, M.C., and Blöschl, G. (2011). Hydrological modelling in a changing world. *Progress in Physical Geography,* 35, 249-261.

Pellarin, T., Wigneron, J.-P., Calvet, J.–C., Ferrazzoli, P., Douville, H., Lopez-Baeza, E. et al. (2003) Two-Year global simulation of L-band brightness temperatures over land. *IEEE Transactions on Geoscience and Remote Sensing,* 41, 2135-2139.

Reichle, R.H., Crow, W.T., Koster, R.D., Sharif, H.O., and Mahanama, S.P.P. (2008) Contribution of soil moisture retrievals to land data assimilation products. *Geophysical Research Letters,* 35, L01404.

Reichle, R.H., De Lannoy, G.J.M., Forman, B.A., Draper, C.S., and Liu, Q. (2014) Connecting satellite observations with water cycle variables through land data assimilation: Examples using the NASA GEOS-5 LDAS. *Surveys in Geophysics,* 35, 577-606.

Rowlandson, T.L., Hornbuckle, B.K., Bramer, L.M., Patton, J.C., and Logsdon, S.D. (2012). Comparison of evening and morning SMOS passes over the Midwest United States. *IEEE Transactions on Geoscience and Remote Sensing,* 50, 1544-1555.

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

Sánchez, N., Martínez-Fernández, J., Scaini, A., and Pérez-Gutiérrez, C. (2012) Validation of the SMOS L2 soil moisture data in the REMEDHUS network (Spain). *IEEE Transactions on Geoscience and Remote Sensing*, 50, 1602–1611.

Schaefer, G.L., Cosh, M.H., and Jackson, T.J. (2007) The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). *Journal of Atmospheric and Oceanic Technology*, 24, 2073-2077.

Seneviratne, S.I., Corti, T. , Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowski, B. and Teuling, A.J. (2010) Investigating soil moisture-climate interactions in a changing climate: A review, *Earth Science Reviews* 99, 125-161.
Su, C.-H., Ryu, D., Young, R.I., Western, A.W., and Wagner, W. (2013). Inter-comparison of microwave satellite soil moisture retrievals over the Murrumbidgee Basin, southeast Australia. *Remote Sensing of Environment*, 134, 1-11.

Ulaby, F.T., Moore, R.K., Fung, A.K., 1981. Microwave Remote Sensing: Active and Passive, Volume 1. Reading, Mass.: Addison-Wesley Publishers.

Vachaud, G., Silans, P. D., Balabanis, P., and Vauclin, M. (1985). Temporal stability of spatially measured soil water probability density function. *Soil Science Society of America Journal*, 49, 822-828.

Verhoest, N.E.C., Lievens, H., Wagner, W., Álvarez-Mozos, J., Moran, M.S., and Mattia, F. (2008) On the soil roughness parameterization problem in soil moisture retrieval of bare surfaces from synthetic aperture radar. *Sensors* 8, 4213-4248.

Vivoni, E.R., Gebremichael, M., Watts, C. J., Bindlish, R., and Jackson, T.J. (2008) Comparison of ground-based and remotely-sensed surface soil moisture estimates over complex terrain during SMEX04. *Remote Sensing of Environment*, 112, 314-325.

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, 11761-11774.

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

Wang, J. R. and Choudry, B. J. (1981) Remote sensing of soil moisture content over bare field at 1.4GHz frequency. *Journal of Geophysical Research*, 86, 5277-5282.

Wigneron, J.-P., Kerr, Y., Waldteufel, P., Saleh, K., Escorihuela, M.J., Richaume, P., Ferrazzoli, P., de Rosnay, P., Gurney, R., Calvet, J.-C., Grant, J.P., Guglielmetti, M., Hornbuckle, B., Mätzler, C., Pellarin, T. and Schwank, M. (2007) L-band microwave emission of the biosphere (L-MEB) model: Description and calibration against experimental data sets over crop fields. *Remote Sensing of Environment*, 107, 639–655.

Wilson, D.J., Western, A.W., and Grayson, R.B. (2005) A terrain and data-based method for generating the spatial distribution of soil moisture. *Advances in Water Resources,* 28, 43-54.

Wood, E.F., Roundy, J.K., Troy, T.J., van Beek, L.P.H., Bierkens, M.F.P., Blyth, E., De Roo, A., Doll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, M., Houser, P., Jaffe, P.R., Kollet, S., Lehner, B., Lettenmaier, D.P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A. and Whitehead, P. (2011). Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring earth's terrestrial water. *Water Resources Research,* 47, W05301.

Chapter 2: McMaster Mesonet Soil Moisture Dataset: Description and Spatio-temporal Variability Analysis

Summary of Paper 1: Kornelsen, K.C. and Coulibaly, P. (2013) McMaster Mesonet Soil Moisture Dataset: Description and Spatio-temporal Variability Analysis, Hydrology and Earth Systems Science, 17(1), 1-18, doi: 10.5194/hess-17-1-2013.

Summary:

This research sets the context for the study area used in much of the rest of this thesis. The McMaster Mesonet consists of four soil moisture monitoring sites with continunous hourly soil moisture information at multiple depths since 2006. A description is given of the watershed and the data and a spatial variability and temporal stability analysis was conducted. These are standard analysis techniques for soil moisture datasets and allow for direct comparison of the characteristics of the McMaster Mesonet to other datasets. This is the first such analysis to occur in Canada and considers the longest time period published to the author's knowledge. This work is also novel in that temporal stability is considered both over long periods of time and seasonally. Similarly, the hourly nature of the McMaster Mesonet data set allows for the first analysis of the persistence of the temporal pattern following rain events.

The results of this research demonstrate:

- The McMaster Mesonet sites were found to be temporally stable, particularly at the seasonal time step. This has implications for the bias correction work later in the thesis.
- The spatial variability of soil moisture, particularly at the Governor Road site, was sometimes greater than the target uncertainty of soil moisture retrievals from SMOS or SMAP.
- The temporal persistence of the spatial distribution of soil moisture was seldom interrupted and required either very wet conditions prior to a moderate size rain event or a large amount of water to be added to the soil through precipitation.
- Following large rain events, the spatial distribution of soil moisture returned to the temporally persistent pattern within 18-24 hours, even for the Kelso site which is flat and has poorly drained soils.

2.1 Abstract

This paper introduces and describes the hourly high resolution soil moisture dataset continuously recorded by the McMaster Mesonet located in the Hamilton-Halton Watershed in Southern Ontario, Canada. The McMaster Mesonet consists of a network of time domain reflectometer (TDR) probes collecting hourly soil moisture data at six depths between 10 cm and 100 cm at nine locations per site spread across four sites in the 1250 km^2 watershed. The sites for the soil moisture arrays are designed to further improve understanding of soil moisture dynamics in a seasonal climate and to capture soil moisture transitions in areas that have different topography, soil and land-cover. The McMaster Mesonet soil moisture constitutes a unique database in Canada because of its high spatio-temporal resolution. In order to provide some insight into the dominant processes at the McMaster Mesonet sites, a spatio-temporal and temporal stability analysis were conducted to identify spatio-temporal patterns in the data and to suggest some physical interpretation of soil moisture variability. It was found that the seasonal climate of the Great Lakes Basin causes a transition in soil moisture patterns at seasonal time scales. During winter and early spring months, and at the meadow sites, soil moisture distribution is governed by topographic redistribution, whereas following efflorescence in the spring and summer, soil moisture spatial distribution at the forested site was also controlled by vegetation canopy. Analysis of short-term temporal stability revealed that the relative difference between sites was maintained unless there was significant rainfall (>20 mm) or wet conditions *a priori*. Following a disturbance in the spatial soil moisture distribution due to wetting, the relative soil moisture pattern re-emerged in 18 to 24 hours. Access to the McMaster Mesonet data can be provided by visiting www.hydrology.mcmaster.ca/mesonet.

2.2 Introduction

The spatial and temporal variability of soil moisture both at the surface and in the root-zone is an important control in many hydrological and atmospheric fluxes. These fluxes play a critical role in water and energy balances, and have both a direct and indirect impact on water resources and local climate. Soil moisture is of great significance for scientific and operational applications such as flood prediction and forecasting (Komma et al. 2008; Mahanama et al. 2008; Brocca et

al. 2009), numerical weather prediction (Mohr et al. 2003; Loew et al. 2009; Alavi et al. 2010), climate modeling (Merlin et al. 2006; Seneviratne et al. 2010) and other disciplines, because it controls the partition of both mass and energy in hydro-meteorological processes. The potential of soil moisture data is being realized through recent technological advances, which have allowed for detailed *in situ* and remote soil moisture monitoring. As monitoring programs become more widespread and temporally consistent, they are providing a better understanding of the processes which determine the spatial and temporal distribution of soil moisture. The spatial distribution of soil moisture is determined by an organized structure that is perturbed by stochastic forcing (Bronstert and Bardossy, 1999), and analyses of soil moisture monitoring programs have revealed that the relative dominance of any organized or stochastic factor varies with basin, soil, topography, vegetation, meteorological and scale characteristics (Vanderlinden et al. 2012). However, no high resolution, in terms of spacing and periodicity, soil moisture data was available in Canada to carry out such analyses. The McMaster Mesonet was establised to fill that gap, and to allow for insight into the factors controlling soil moisture distribution. An area that requires further research (Vanderlinden et al. 2012).

Blöschl and Sivapalan (1995) discussed the scale of hydrological observations as being characterized by a scaling triplet consisting of extent, spacing and support. Extent refers to the areal size represented by the observations, spacing to the distance between observation points and support the area that is represented by each observation. They similarly disected the spatial (temporal) scale of hydrological processes into spatial extent (duration), space (time) period and integral scale (correlation length) (Blöschl and Sivapalan, 1995). Over the past decade(s) many soil moisture monitoring efforts have been undertaken to characterize the spatial-temporal distribution of soil moisture and its processes at a particular scale of interest. Intensive short term monitoring efforts for large areas such as the Southern Great Plains (SGP) (Familglietti et al. 1997; Mohanty et al. 2000b; Martinez-Fernandez and Ceballos, 2003; Ryu and Famiglietti, 2005; 2006) and Soil Moisture Experiments (SMEX) series (Cosh et al. 2004; Bosch et al. 2006; Choi and Jacobs, 2007; Das et al. 2008) were designed to monitor soil moisture at both a large spatial extent with moderate spacing for process description, but were limited by a moderate period and short duration. Large scale monitoring whether of short (SMEX, SGP) or long duration (Ceballos et al. 2005; Albergel et al. 2008; Lebel et al. 2009) for calibration/validation of radiometer scale (~40 km) soil moisture products, use multiple observations of wide spacing to increase the extent

of soil moisture observations. While coarse resolution (large spacing/extent) soil moisture is relatively abundant and has been shown to enhance hydrological and atmospheric modelling, the advantage of high resolution datasets is increasingly recognized (Wood et al. 2011) and the importance of small scale heterogeneity has been demonstrated (Merlin et al. 2006; Alavi et al. 2010; Minet et al. 2010). Efforts to characterize soil moisture data at high spatial resolution (hillslope scale) can be described as having a moderate period (daily) and duration (days-months) (Famiglietti et al. 1998; Mohanty et al. 2000a; Hupet and Vanclooster, 2002) or are sampled at low to moderate periodicity (Grayson and Western, 1998; Wilson et al. 2003; Western et al. 2004; Brocca et al. 2010). To monitor hydrological processes, soil moisture has also been monitored at very short periods (minutes) for short durations (hours) (Torres et al. 1998). The McMaster Mesonet was designed to provide appropriate soil moisture information needed for process undertanding and modeling, and developing soil moisture retrieval and extension algorithms by monitoring soil moisture with both a short (hourly) period and long duration (years) at a small extent and spacing.

With respect to spatial soil moisture distribution, some studies have found that soil moisture variability increases in wet conditions (Famiglietti et al. 1998; Vivoni et al. 2008), while others have found variability increases in dry conditions (Jacobs et al. 2004; Bosch et al. 2006; Choi and Jacobs, 2007; Brocca et al. 2010). Analyzing the results from many studies Brocca et al. (2007) found that in humid climates, spatial variability is greater when conditions are dry, whereas semi-arid environments have the highest variability in wet conditions. These relationships are also subject to considerations of scale and topography. For example, scale is important as homogenous rainfall tends to decrease soil moisture variability and heterogeneous rainfall increases it (Cosh et al. 2004), whereas Famiglietti et al. (1998) found that following a rainstorm upper portions of a hillslope were more variable than lower portions of the hillslope causing an overall increase in variability when the entire landscape was considered. At smaller extents, precipitation is generally homogenous and the redistribution of soil moisture by topography, soil texture and vegetation become important post precipitation (Wilson et al. 2003; Famiglietti et al. 2008). During wetting, soil moisture variability is dominated by the soils infiltration capacity and topographic redistribution (Famiglietti et al. 1998; Western and Blöschl, 1999; Western et al. 2004; Vivoni et al. 2008; Heathman et al. 2009) while, under dry conditions variability is maintained by the soil water holding capacity and concavity of the surface

(Famiglietti et al. 1998; Peters-Lidard et al. 2001; Vivoni et al. 2008). Vegetation is also a potentially important predictor of soil moisture distribution which can redistribute soil moisture affording a homogenizing effect (Ivanov et al. 2010) or partially explain soil moisture spatial variability in some landscapes (Hupet and Vanclooster, 2002; Bosch et al. 2006), whereas others have found that the role of vegetation in soil moisture distribution is only minor (Cosh et al. 2004). Vachaud et al. (1985) noticed that the relative rank of soil moisture at a particular location with respect to similar nearby locations was persistent in time, leading to the assertion of temporal stability, or more appropriately rank stability (Chen, 2006). The presence of temporally stable soil moisture patterns has been noted during several soil moisture campaigns (Martinez-Fernandez and Ceballos 2003; Cosh et al. 2004; Bosch et al. 2006; Vivoni et al. 2008) and has also been found to result from soil, topographic and vegetation influences (Vivoni et al. 2008). However, flat topography and soil moisture redistribution have also been observed to result in poor temporal stability (Mohanty et al. 2000b; Mohanty and Skaggs, 2001). None of those soil moisture dynamics analysis studies were conducted in Canada because of the lack of appropriate soil moisture data. The role of seasonal effects on soil moisture variability and stability in cool and snowy climates remains an open research area. The McMaster Mesonet database will help to fill that gap as well.

This paper introduces and describes the long-term high resolution McMaster Mesonet dataset located in the 1250 km^2 Hamilton-Halton Watershed in Southern Ontario, Canada. It also provides a spatio-temporal analysis of the hourly soil moisture data collected at four sites since autumn 2006. The experiment was designed specifically for application to high resolution remote sensing soil moisture validation, hydrological data assimilation, and process understanding. The unique aspect of this dataset is the multiple soil moisture profiles that are collected at each site for both an hourly period and long duration. This will allow for the characterization of hillslope scale soil moisture variability at event, seasonal and inter-annual temporal scales, and will provide insight into the influence of topography, vegetation and atmospheric conditions on small scale soil moisture dynamics. Also, most soil moisture experiments described in the literature are from the U.S.A., Europe and Australia and so, to the best knowledge of the authors, this paper represents the first attempt to describe long term soil moisture dynamics in Canada and specifically the Great Lakes Region. It is expected that the McMaster Mesonet will monitor hillslope scale processes that are representative of those occuring in the southern Great Lakes

Region of Canada in general and the Hamilton-Halton watershed in particular. Due to the limited spatial extent of the monitoring sites, it is surmised that the dataset is better representative of regional temporal trends than regional spatial soil moisture. The main dataset consists of a series of four high resolution soil moisture arrays collecting hourly distributed soil moisture profile information since 2006 with an expected lifespan of fifteen years. The soil moisture data is supplemented by six weather stations and nine rain gauges distributed throughout the watershed. This dataset can be made available to the broader research community by visiting the website www.hydrology.mcmaster.ca/mesonet and plans are underway to include the dataset in the International Soil Moisture Network (Dorigo et al. 2011). This paper would serve as an essential reference for the McMaster Mesonet data users.

2.3 McMaster Mesonet

2.3.1 Hamilton-Halton Watershed

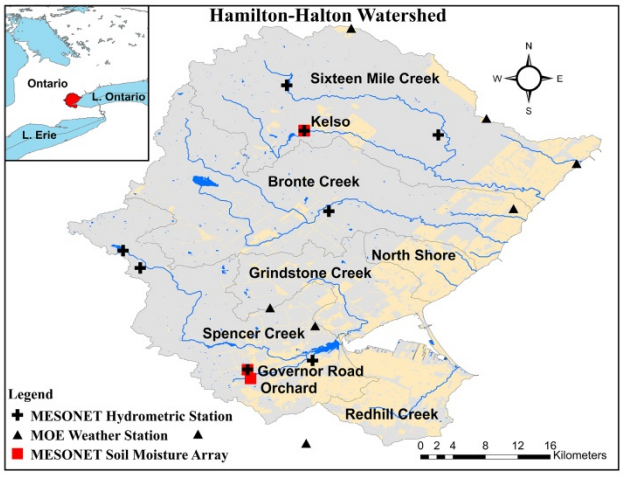

Figure 2-1: Location map of the Hamilton-Halton Watershed including sub-watersheds and the locations of the soil moisture arrays and hydro-meteorological stations.

The Hamilton-Halton Watershed (Fig. 2-1) is part of the Lake Ontario drainage basin in Ontario, Canada and has approximately 980 km^2 rural agricultural/forested land and 270 km^2 of urbanized/industrial land. The urbanized land in the watershed is concentrated within a band that extends approximately 7.5 km from the Lake Ontario shoreline, with the notable exception of the Town of Milton which covers an area of 25 km^2, in the northern-central part of the watershed. The major geographic features include the Lake Iroquois Plains, which are an ancient glacial extension of the current Lake Ontario shoreline and the Niagara Escarpment. The area surrounding the Niagara Escarpment is primarily mixed deciduous/coniferous woodland, and agriculture dominates the remainder of the rural area. The primary crops are maize, soy and grains with some tender fruit crops. The watershed is sub-divided into six primary sub-watersheds, Sixteen Mile Creek, Bronte Creek,

North Shore, Grindstone Creek, Spencer Creek and Red Hill Creek, each with their own network of tributaries. Sixteen Mile Creek is the northernmost sub-watershed and covers an area of 357 km^2 and is predominantly agricultural. The Bronte Creek Watershed encompasses an area of 304 km^2 and has the largest proportion of forested area around the Niagara Escarpment. Grindstone Creek has a catchment area of 99 km^2 and is largely rural agricultural with forest and some urban area in the south east. Both North Shore Creek (44 km^2) and Red Hill Creek (93 km^2) are predominantly urban areas and have been modified to accommodate urban storm water management. Spencer Creek encompasses an area of 260 km^2 and is predominantly rural/agricultural and includes part of the City of Hamilton and forested area around the Niagara Escarpment in the south eastern portion of the sub-watershed. It is important to note that areas reported herein are consistent with what is portrayed in Fig. 2-1, however, in reality the sub-watersheds as reported also contain small waterways which drain directly into Lake Ontario but are not distinguished herein.

The climate of the watershed can be classified as humid continental with average annual precipitation of 910 mm distributed evenly throughout the year. The watershed experiences four distinct seasons, with average summer temperatures of 21°C and average winter temperatures of -6°C (1971 to 2000 Canadian Climate Normals). A time series plot of climate variables and

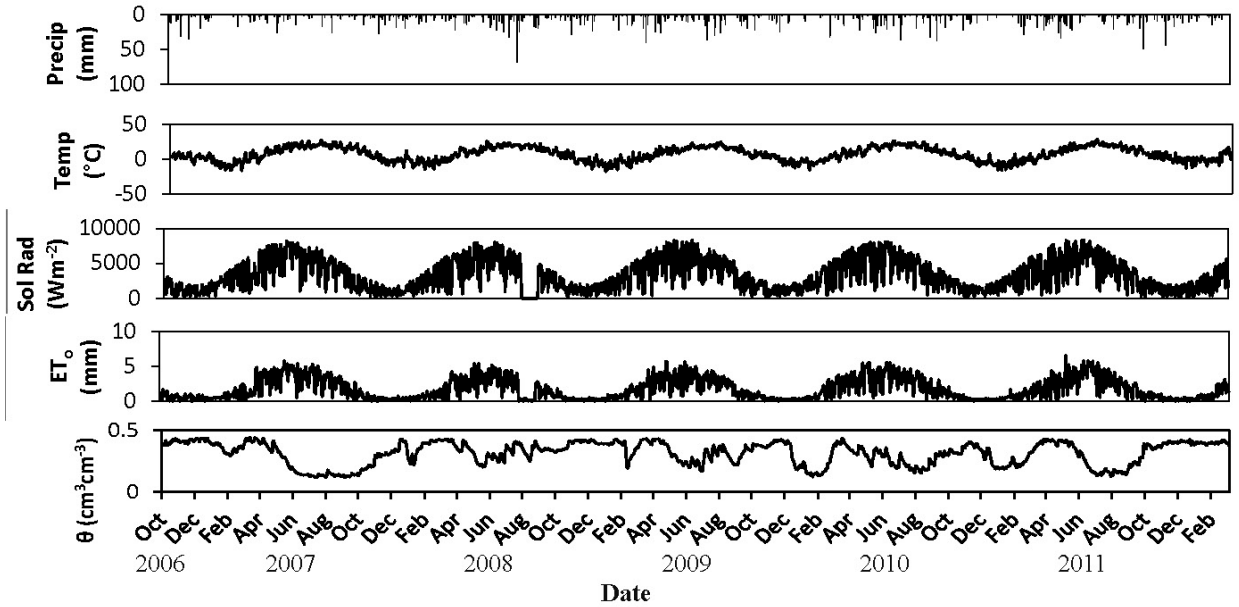

Figure 2-2: Daily time series plot of 10 cm site averaged soil moisture from K1 and daily time series of meteorological data collected at Kelso.

Table 2-1: Climatic conditions based on Brittania weather station (located at the centre of the watershed) and the Hamilton Airport weather station (Approx. 5km south of the watershed).

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Ann
Hamilton Airport (1971-2000)													
Daily Max Temp (°C)	-2.2	-1.2	4.0	11.2	18.5	23.7	26.3	25.1	20.7	13.8	7.0	0.9	12.3
Daily Min Temp (°C)	-9.7	-9.1	-4.5	1.2	7.3	12.4	15.1	14.5	10.2	4.4	-0.4	-6.2	2.9
Daily Mean Temp (°C)	-6	-5.2	-0.3	6.3	12.9	18.0	20.8	19.8	15.5	9.1	3.3	-2.7	7.6
Rainfall (mm)	29.5	25.7	48.6	69.6	75.0	83.9	86.5	80.6	82.1	71.6	68.1	43.7	764.8
Precipitation (mm)	65.8	55.3	74.8	78.0	75.6	83.9	86.5	80.6	82.1	72.5	78.6	76.6	910.1
Mesonet – Brittania (2007)													
Daily Max Temp (°C)	-0.3	-4.9	4.2	9.2	20.3	26.0	26.1	26.5	23.4	17.8	5.5	-0.4	12.8
Daily Min Temp (°C)	-7.4	-13.1	-4.6	1.4	7.6	13.4	13.7	15.1	10.9	8.2	-1.9	-6.3	3.1
Daily Mean Temp (°C)	-3.8	-9.0	-0.2	5.3	14.0	19.7	19.9	20.8	17.2	13.0	1.8	-3.3	7.9
Precipitation (mm)	51.8	10.6	57.8	70.8	46.0	43.4	16.8	28.0	25.6	50.4	66.6	62.0	529.8
Mesonet – Brittania (2008)													
Daily Max Temp (°C)	0.5	-1.3	2.0	14.0	16.3	23.7	26.1	22.3	21.1	13.4	5.7	0.0	12.0
Daily Min Temp (°C)	-6.4	-9.9	-6.4	2.7	5.3	13.6	15.2	7.3	10.7	3.2	-1.4	-8.2	2.2
Daily Mean Temp (°C)	-2.9	-5.6	-2.2	8.3	10.8	18.7	20.7	14.8	15.9	8.3	2.2	-4.1	7.1
Precipitation (mm)	58.0	64.6	49.0	50.0	53.8	108.6	143.2	120.0	102.2	36.6	77.2	100.6	963.8
Mesonet – Brittania (2009)													
Daily Max Temp (°C)	-5.2	0.8	5.0	11.8	18.1	21.9	23.1	25.1	21.1	11.6	9.1	-0.2	11.9
Daily Min Temp (°C)	-13.5	-8.8	-4.5	1.8	6.5	11.4	13.2	14.7	10.5	3.7	1.0	-6.3	2.5
Daily Mean Temp (°C)	-9.4	-4.0	0.3	6.8	12.3	16.7	18.2	19.9	15.8	7.7	5.0	-3.3	7.2
Precipitation (mm)	17.8	71.2	63.0	137.6	50.0	52.8	97.4	120.8	31.4	82.9	32.5	91.6	849.0
Mesonet – Brittania (2010)													
Daily Max Temp (°C)	-2.6	-0.9	8.2	15.9	20.4	23.1	28.0	26.5	20.3	14.1	7.7	-1.8	13.3
Daily Min Temp (°C)	-8.8	-7.4	-1.1	3.8	8.9	13.6	16.4	15.7	10.2	4.6	-1.1	-7.2	4.0
Daily Mean Temp (°C)	-5.7	-4.1	3.6	9.8	14.7	18.4	22.2	21.1	15.2	9.4	3.3	-4.5	8.6
Precipitation (mm)	20.1	19.0	91.2	48.6	55.2	138.6	127.0	39.4	106.8	73.0	129.0	31.3	897.2
Mesonet – Brittania (2011)													
Daily Max Temp (°C)	-3.8	-1.2	3.0	10.6	17.7	23.3	29.6	26.5	21.7	14.5	10.3	3.5	13.0
Daily Min Temp (°C)	-11.9	-10.2	-5.8	1.2	8.4	12.5	16.7	15.0	11.3	5.5	2.0	-3.6	3.4
Daily Mean Temp (°C)	-7.8	-5.7	-1.4	5.9	13.0	17.9	23.1	20.8	16.5	10.0	6.2	0.0	8.2
Precipitation (mm)	24.4	34.2	87.0	100.6	142.2	54.0	12.0	86.4	74.2	125.2	82.6	68.0	890.8
Mesonet – Brittania (2007 - 2011)													
Daily Max Temp (°C)	-2.3	-1.5	4.5	12.3	18.6	23.6	26.6	25.4	21.5	14.3	7.7	0.2	12.6
Daily Min Temp (°C)	-9.6	-9.9	-4.5	2.2	7.3	12.9	15.0	13.6	10.7	5.0	-0.3	-6.3	3.0
Daily Mean Temp (°C)	-5.9	-5.7	0.0	7.2	13.0	18.3	20.8	19.5	16.1	9.7	3.7	-3.0	7.8
Precipitation (mm)	34.4	39.9	69.6	81.5	69.4	79.5	79.3	78.9	68.0	73.6	77.6	70.7	826.1

average soil moisture data from the McMaster Mesonet for 2006 to 2012 can be seen in Fig. 2-2 and a climate summary can be found in Table 2-1. The monthly and annual climate patterns at the Brittania weather station, located in the centre of the watershed, and the long term climate normals for the Hamilton Airport (5 km south of watershed) are presented for comparison. It should be noted that the Mesonet precipitation presented in Table 2-2 represents primarily rainfall as there is no active collection of data for snow water equivalent. The climate in the watershed is similar to the climate normals, with the exception of 2007, which had considerably lower than normal summer precipitation. Consequently, a prolonged period of low soil moisture can be seen throughout the summer of 2007 in Fig. 2-2 which is nearly two standard deviations below the mean soil moisture for the study period. A similar dry spell occured during the summer of 2011, but was compensated for annually by a wetter than normal spring and autumn in that same year. The climatic variability in the watershed during the course of study is ideal for studying soil moisture variability as a large variety of conditions have been observed within a relatively short period of study, by climatological standards.

2.3.2 Data Description

A soil moisture 'site' refers to the location of an entire soil moisture array, i.e. Kelso 1 (K1), Kelso 2 (K2), Governor Road (GR) or Orchard (OR). Each site contains nine stations which are numbered from 1 to 9, where a 'station' refers to a vertical soil moisture profile at a particular geographic location (Figs. 2-3 and 2-4). A station has six associated measurements for each sampling period for the six depths indicated in Table 2-2. Soil moisture characterized as 'daily' refers to the mean value, and its variance, of a discrete 24 hour period measured using Eastern Standard Time.

The McMaster Mesonet provides long term hourly soil moisture data at four sites, Kelso 1, Kelso 2, Governor Road and Orchard which have been collecting data continuously since 2006. Each site contains 54 Campbell Scientific CS616 multiplexed time domain reflectometery (TDR) probes attached to a CR10X datalogger through nine soil moisture profile stations with six TDR probes each. Each profile station collects measurements at six depths between 10 to 100 cm, with specific depths given in Table 3-2. At each of the four sites, the nine profile stations are distributed in a grid pattern, where a 100 cm pit was dug at each station and six TDR probes were inserted horizontally into the soil. The majority of probes were inserted at 10, 20, 30, 50, 70

Table 2-2: McMaster Mesonet site and station description.

Site Name & Description	Station	Probe Depths (cm)	Station Description
Orchard – Dundas Valley Soil Texture: silt loam Vegetation: short grass Terrain: gentle slopes, hill-slope Drainage: soil drains well Hydrology: moderate infiltration when wet	1	10,20,30,40,50,60	clear sky, side-slope
	2	10,20,30,50,70,100	partial canopy cover, downslope
	3	10,20,30,40,50,60	clear sky, downslope
	4	10,20,30,50,70,100	clear sky, downslope
	5	10,20,30,50,70,100	clear sky, side-slope
	6	10,20,30,40,50,60	Partial canopy cover, upslope
	7	10,20,30,40,50,60	clear sky, upslope
	8	10,20,30,50,70,100	under tree canopy, upslope
	9	10,20,30,40,50,60	clear sky, side slope
Governor Road – Dundas Valley Soil Texture: silt loam Vegetation: mixed forest; generally pine Terrain: gentle and steep slopes Drainage: soil drains well Hydrology: moderate infiltration when wet	1	10,20,30,50,70,100	clear sky, close to tree, mid-elevation
	2	10,20,30,40,50,60	canopy cover, mixed forest, side sloping
	3	10,20,30,40,50,60	partial canopy cover, in valley
	4	10,20,30,40,50,60	clear sky, close to tree, uphill
	5	10,20,30,40,50,60	dense canopy cover, mid-elevation
	6	10,20,30,40,50,60	partial canopy cover, down-hill
	7	10,20,30,40,50,60	clear sky, uphill, upslope
	8	10,20,30,50,70,100	canopy cover, mixed forest, upslope valley
	9	10,20,30,40,50,60	canopy cover, mixed forest, side valley
Kelso 1 Soil Texture: clay loam Vegetation: short grass (light vegetation) Terrain: generally flat Drainage: is imperfect and poor Hydrology: slow to very slow infiltration when wet	1	10,20,30,50,70,100	Clear sky, tall grass, mid elevation
	2	10,20,30,50,70,100	Clear sky, tall grass, low elevation
	3	10,20,30,50,70,100	Clear sky, tall grass, high elevation
	4	10,20,30,50,70,100	Clear sky, tall grass, high elevation
	5	10,20,30,50,70,100	Clear sky, short grass, mid elevation
	6	10,20,30,40,50,60	Clear sky, short grass, mid elevation
	7	10,20,30,40,50,60	Clear sky, short grass, low elevation
	8	10,20,30,40,50,60	Clear sky, short grass, mid elevation
	9	10,20,30,40,50,60	Clear sky, tall grass, high elevation
Kelso 2 Soil Texture: clay loam Vegetation: short grass (light-to-dense vegetation) Terrain: generally flat Drainage: is imperfect and poor Hydrology: slow to very slow infiltration when wet	1	10,20,30,40,50,60	Clear sky, short grass, low elevation
	2	10,20,30,50,70,100	Clear sky, tall grass, high elevation
	3	10,20,30,50,70,100	Clear sky, tall grass, high elevation
	4	10,20,30,50,70,100	Clear sky, tall grass, mid elevation
	5	10,20,30,40,50,60	Clear sky, short grass, mid elevation
	6	10,20,30,40,50,60	Clear sky, dense grass, low elevation
	7	10,20,30,50,70,100	Clear sky, dense grass, low elevation
	8	10,20,30,50,70,100	Clear sky, dense grass, mid elevation
	9	10,20,30,50,70,100	Clear sky, dense grass, high elevation

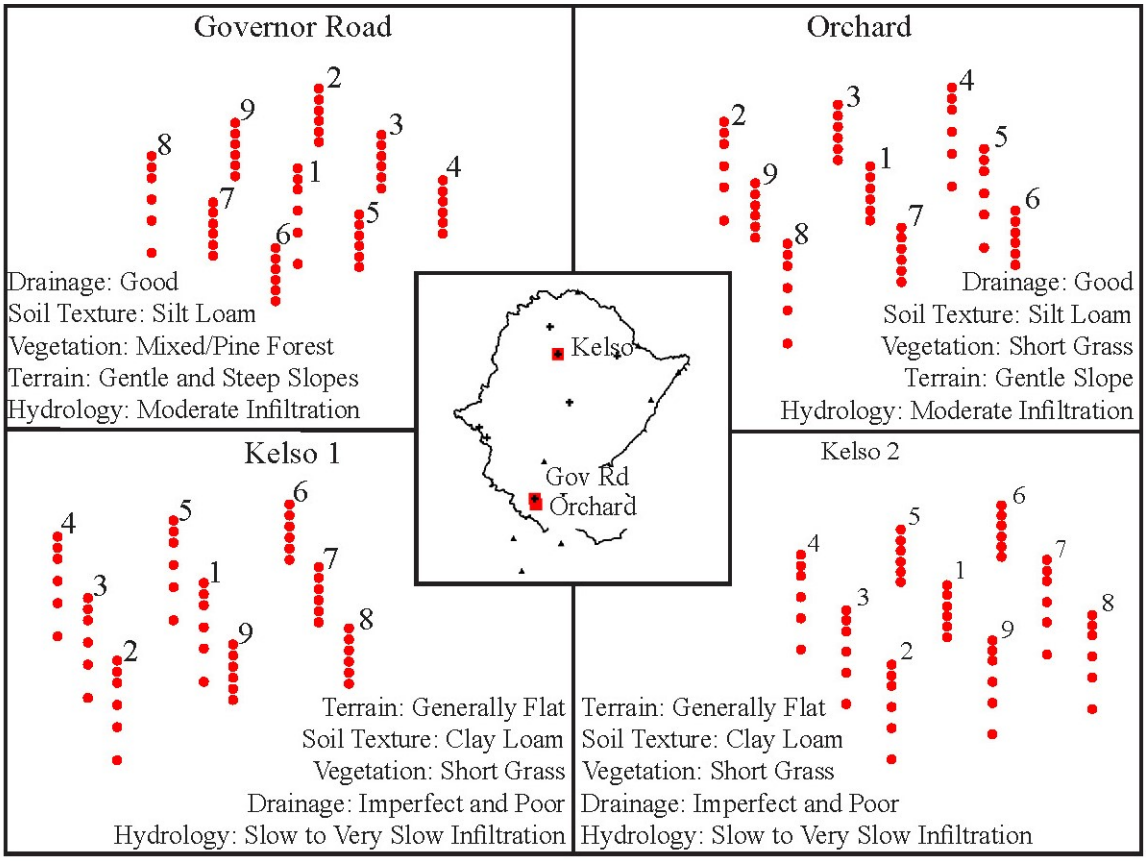

Figure 2-3: Three dimensional schematic representation of the McMaster Mesonet. Note: The vertical dimensions are to scale, whereas the spatial dimensions are not (See Fig. 2-4). The symbols in the centre of the map are consistent with Fig. 2-1.

and 100 cm depths in order to capture the change in soil moisture in the hydrologically dynamic layer. In some instances, high clay content and topographic conditions did not allow for the full 100 cm depth to be reached and the TDR probes were inserted at 10, 20, 30, 40, 50, 60 cm depths instead (see Table 2-2). All of the TDR profiles are connected to a data logger at the centre of the array at station 1. Each TDR array has an associated tipping bucket rain gauge and automated weather stations are associated with the TDR arrays located between K1 and K2 and at GR. The soil moisture data has been pre-processed to remove most erroneous measurements and when less than 72 temporally consecutive missing values were present, the data was infilled using linear interpolation (Kornelsen and Coulibaly, 2013) and/or the soil layer relative difference method if some values were missing between stations (Kornelsen and Coulibaly, 2013; Dumedah and Coulibaly, 2010). In addition to the Campbell Scientific datasets, a Stevens Water Hydra Probe array is operated in conjunction with the CS616 array at Kelso 1 providing

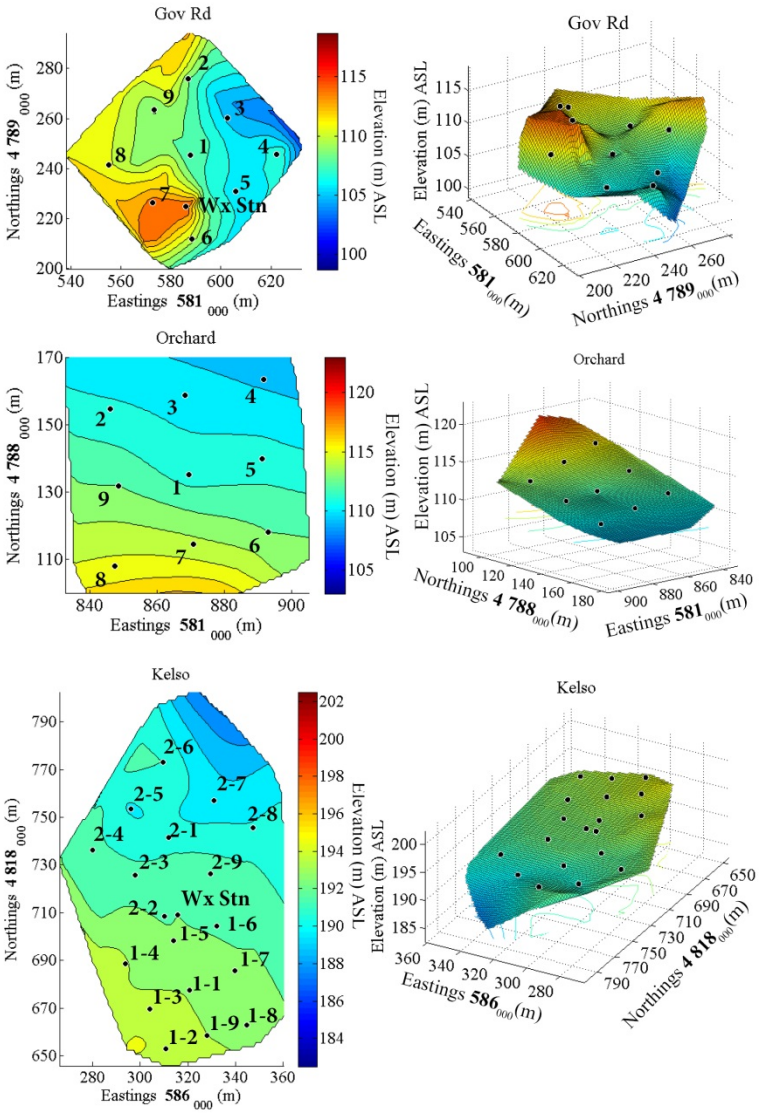

Figure 2-4: Topography and layout of the soil moisture arrarys at GR (top), OR (middle) and K1/K2 (bottom) as contour plots (left) and surface renderings (right). Contours were derived using Natural Neighbour Interpolatoin using topographic data collected with an Ashtech MM100 GPS and have an approximate horizontal RMSE of 20 cm and vertical RMSE of 50 cm. The surface plot vertical exaggeration is 3X. Note: Surface and contour plots have different orientations to enhance the visual interpretation of the surface plot.

an independent soil moisture dataset for comparison/validation, where Hydra Probes are co-located at each station and collect soil moisture information at depths of 5, 10, 20, 30, 40 and 50 cm at half-hourly intervals.

The sites were selected to represent three distinct landscapes common to the watershed and are located on protected conservation land to ensure the safety and longevity of the monitoring network. The spatial extent to which the soil moisture arrarys are representative is considered limited as hillslope scale soil moisture processes, rather than watershed scale dynamics, are emphasized by the distrubtion of TDR probes. However, because the sites are located in landscapes representative of those in the watershed, it is hoped that future studies can ascertain the support of the dataset and potentially extend the representativeness to the two sub-watersheds in which the arrays are located. The distribution of the additional hydro-meteorological stations of the McMaster Mesonet provides a relatively dense monitoring network for watershed scale applications which can be supported by the hillslope scale soil moisture arrays.

The Kelso sites (K1 & K2) are located in the northern portion of the watershed in an area which is predominantly agricultural land. The soil is a Gray-Brown Podzol (clay loam) that is a poorly drained with a higher clay content at depth (80-100 cm). The A horizon is light grey and mottled near the Bt horizon. These soils seat atop calcareous till at a depth of approximately 75 cm (Gillespie et al. 1971). The terrain is generally flat with some hummocks and a few small gentle sloping hills at the edge of the site. The land-cover is predominantly meadow with some recently planted coniferous and aspen trees scattered throughout the site. The K1 and K2 sites are located adjacent to each other to allow sampling of a larger area. However, both datasets are collected and stored independently and will herein often be considered as distinct data, inspite of their co-location, for the purposes of describing the dataset.

The GR and OR sites are located in the Dundas Valley Conservation Area. The Dundas Valley is part of the Niagara Escarpment and is predominantly covered by mixed Carolinian forests, and fields. Both sites have Ancaster Series Gray-Brown Podzolic soils with good drainage and moderate infiltration. Typical Ancaster soils have a thin Ap horizon of greyish-brown silt loam over a thicker Ae brown silt loam. The B horizons are well developed and contain clay and iron accumulations. The C horizon typically begins at 75 cm depth and has a higher clay content (Presant et al. 1965). OR is located in a reclaimed apple orchard which is covered by meadow vegetation and has sparse apple trees. The site covers the transition between a gentle north facing slope and a flat plateau. The GR site is located on mixed terrain and has steep and gentle slopes having mixed/pine forest covering most of the site with small open areas around the weather station and on the steep slopes. Figure 2-4 shows the terrain of each site and the locations of the soil moisture stations within each site.

The soil moisture arrays are supplemented with six weather stations and nine rain gauges distributed throughout the watershed. The weather stations are a mix of Campbell Scientific and HOBO stations, recording half-hourly observations of air temperature, relative humidity, vapour pressure, saturation vapour pressure, precipitation (rain only), incoming solar radiation, wind speed and wind direction. Additionally, grass reference evapotranspiration is calculated online from the collected data within the weather station using the Penman-Monteith equation (Campbell Scientific, 1999). The weather/hydrometric stations were distributed to provide good characterization of the watershed with considerations given to accessibility and security. There are also 8 weather stations within or near the watershed which are operated by the Ontario

Ministry of Natural Resources or Environment Canada, which are also shown on Fig. 2-1. While not continously recorded by the McMaster Mesonet, limited snowfall and snow depth measurements in and near the watershed are available from Environment Canada weather stations. In addition to meteorological data, water level, discharge and ground water data are available independently from Environment Canada‘s HYDAT database and the Ontario Ministry of the Environment.

2.3.3 Campbell Scientific TDR and Stevens Hydra Probe Comparison

At the K1 site, Stevens Hydra Probes were installed alongside the CS616 TDR probes at depths of 10-50 cm in 10 cm increments in order to provide a comparison between the two data products and Hydra Probes were additionally installed at each station at a depth of 5 cm. Gravimetric sampling was not conducted regularly at this site in conjunction with automated measurements, therefore, only a comparison between measurements can be made, which is summarized in Fig. 2-5. In general, the soil moisture data from the CS TDR and Hydra Probes follow similar temporal trends at all sites in terms of the rate of wetting and drying. The positive differences in Fig. 2-5 indicate that the CS TDR records

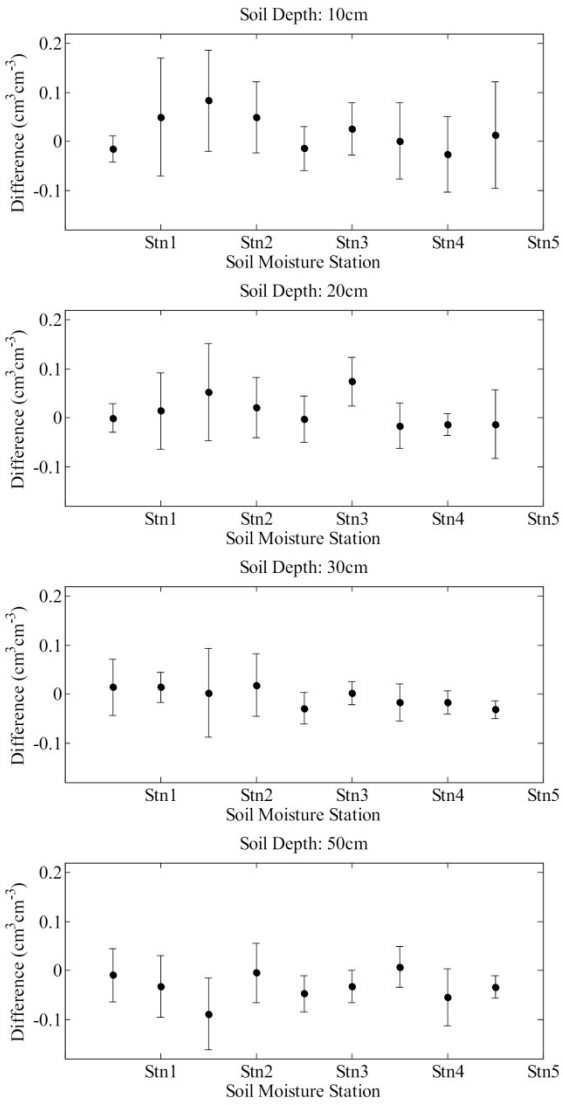

Figure 2-5: Mean difference between Campbell Scientific CS616 TDR and Stevens Water Hydra Probe hourly soil moisture values at K1 from 2007 to 2011. The error bars represent one standard deviation.

higher soil moisture values than the Hydra Probe, which is common amongst all depths. Overall, small differences (less than 5% on average) are observed between the two records. The smallest differences between the two data sets occur at the 30 and 50 cm depths, whereas larger differences are observed between 10 and 20 cm depths. This is mostly due to the large variability of soil moisture in the top layers compared to deeper layers. Some larger differences between the two measurements occur at stations 3 and 4 at all depths (Fig. 2-5). This is in part due to the specific locations of the two stations (3 and 4). At those two locations the CS TDR measurements show higher variability compared to other sites. However in general, the differences are mostly due to the noise in hourly data.

2.4 Methodology

2.4.1 Statistical Methods

For both the analysis and manipulation of soil moisture data, and for ease of use/presentation the data is characterized using standard sample statistics. Herein, both spatial and temporal soil moisture statistics will be presented and so we will differentiate the statistics in terms of variability in both space and time.

Let θ_{ijk} be the soil moisture θ at station i and time k for the sampling depth j. The spatial mean soil moisture for a site $\overline{\theta_{jk}}$ can be calculated as:

$$\overline{\theta_{jk}} = \frac{1}{Ni}\sum_{i=1}^{Ni}\theta_{ijk} \tag{2-1}$$

where Ni is the number of stations i (usually 9) at which soil moisture is sampled for a given depth j. Similarly the daily mean soil moisture for a station $\overline{\theta_{ij}}$ is given as:

$$\overline{\theta_{ij}} = \frac{1}{M}\sum_{k=1}^{M}\theta_{ijk} \tag{2-2}$$

where M is the number of hours (usually 24) over which the mean value is taken. The daily soil moisture for a site at a specific depth is given by:

$$\overline{\theta_{j}} = \frac{1}{NiM}\sum_{i=1}^{Ni}\sum_{k=1}^{M}\theta_{ijk} = \frac{1}{Ni}\sum_{i=1}^{Ni}\overline{\theta_{ij}} \tag{2-3}$$

where the mean daily soil moisture for each station is first derived and the spatial mean is then calculated. For ease of consideration, the standard deviation and variance of soil moisture is only considered based on the second summation. Therefore the daily variance of a station is given by:

$$\sigma_{ij}^2 = \frac{1}{M-1}\sum_{k=1}^{M}\left(\theta_{ijk} - \overline{\theta_{jk}}\right)^2 \tag{2-4}$$

whereas the variance of the daily soil moisture at the entire site for depth *j* is given by:

$$\sigma_{j}^2 = \frac{1}{Ni}\sum_{i=1}^{Ni}\left(\theta_{ij} - \overline{\theta_{j}}\right)^2 \tag{2-5}$$

where the temporal variance determined in Eq. (2-4) is ignored when calculating the daily site variance in Eq. (2-5). Because hourly soil moisture is highly autocorrelated, the daily temporal variability of soil moisture at each station is minimal in comparison to the spatial variability. Therefore, the loss of information resulting from the considered averaging scheme is assumed negligible. The above equations can also be extended to apply to monthly soil moisture where daily soil moisture at a site/station is first calculated and Eq. (2-2/2-3) and (2-4/2-5) are reapplied using *M* as days instead of hours.

2.4.2 Temporal Stability

The concept of temporal stability was first proposed by Vachaud et al. (1985) and is used to determine the temporal persistence of the spatial soil moisture pattern. Analysis of the temporal persistence leads to some understanding of the processes that influence the organized portion of the spatial soil moisture pattern. Temporal stability analysis is conducted using the parametric test of the relative differences, where the relative difference δ_{ijk} at station *i*, depth *j* and time *k* is given by:

$$\delta_{ijk} = \frac{\theta_{ijk} - \overline{\theta_{jk}}}{\overline{\theta_{jk}}} \tag{2-6}$$

In order to mitigate the effects of stochastic influences on the soil moisture pattern and the relative difference statistic, the mean and standard deviation can be calculated where the mean relative difference $\overline{\delta_{ij}}$ and its standard deviation $\sigma(\delta_{ij})$ are given by:

$$\overline{\delta_{ij}} = \frac{1}{M}\sum_{k=1}^{M}\delta_{ijk} \tag{2-7}$$

$$\sigma(\delta_{ij}) = \sqrt{\frac{1}{M-1}\sum_{k=1}^{M}\left(\delta_{ijk} - \overline{\delta_{ij}}\right)^2} \tag{2-8}$$

where the variable *M* can be taken as a daily, monthly, annual or other value. It is common to express the mean relative difference based on the entire study period. A relative difference of zero refers to a station that is representative of the mean soil moisture value, where high and low values represent sites that are consistently wet or dry respectively compared to the mean. In order to better select sampling locations that were both representative of the mean and had low variability Jacobs et al. (2004) proposed the combination of $\sigma(\delta_{ij})$ and δ_{ij} into the root mean square error measure:

$$RMSE_i = \sqrt{\overline{\delta_{ij}}^2 + \sigma(\delta_{ij})^2} \tag{2-9}$$

2.5 Results and Discussion

The soil moisture in the Hamilton-Halton watershed follows a seasonal cycle, with temperature and precipitation patterns seen in Fig. 2-2 and seasonal mean soil moisture presented in Table 2-3. While the precipitation is relatively evenly distributed throughout the year, during the winter months there is little evapotranspiration and precipitation tends to accumulate on the surface as snow. The spring thaw results in strong wetting of the soil, until the late-spring and summer temperature increase causes a net dry-down of the soil. Thus, there are seasonal periods of wetting from the late fall to early spring months followed by a drying period from late spring to early fall. The notable exception to this pattern in Table 2-3 is the soil moisture at depths greater than 50 and 70 cm at GR/OR and Kelso respectively, where the autumn experiences a slight net loss from the previous season. Close analysis (results not shown) reveal this secondary dry-down is the result of dry autumns in 2007 and 2009 (Table 2-1) skewing the temporal statistics. The variability in the surface soil moisture dampened the impact of these dry years in the statistics of the upper soil layers. Since the soil moisture data of the McMaster Mesonet have been collected onward from 2006, the dataset covers several unseasonably wet and dry periods and numerous wetting and drying events, which makes the dataset especially useful for hydrological analysis. Despite the advantages of a continuous dataset, a potential limitation must be noted with respect to freezing soil conditions. The Hamilton-Halton Watershed experiences mild winters in

Table 2-3: Mean soil moisture for Winter (DJF), Spring (MAM), Summer (JJA), and Fall (SON) and the average change between seasons at each site and depth.

Orchard								
	Winter		Spring		Summer		Autumn	
Depth (cm)	Mean	Change	Mean	Change	Mean	Change	Mean	Change
10	0.30	0.00	0.33	0.03	0.27	-0.06	0.30	0.02
20	0.33	0.03	0.34	0.01	0.29	-0.06	0.30	0.01
30	0.35	0.04	0.36	0.01	0.31	-0.05	0.31	0.00
40	0.36	0.04	0.37	0.01	0.33	-0.04	0.33	0.00
50	0.37	0.05	0.37	0.01	0.33	-0.04	0.32	-0.01
60	0.39	0.04	0.39	0.01	0.36	-0.03	0.35	-0.02
70	0.37	0.05	0.38	0.01	0.34	-0.04	0.32	-0.02
100	0.38	0.05	0.39	0.01	0.36	-0.03	0.33	-0.03
Governor Road								
	Winter		Spring		Summer		Autumn	
Depth (cm)	Mean	Change	Mean	Change	Mean	Change	Mean	Change
10	0.28	0.02	0.30	0.02	0.24	-0.06	0.26	0.02
20	0.31	0.04	0.32	0.01	0.26	-0.07	0.26	0.01
30	0.32	0.05	0.33	0.01	0.27	-0.07	0.27	0.00
40	0.33	0.06	0.34	0.01	0.27	-0.07	0.27	0.00
50	0.34	0.06	0.35	0.01	0.29	-0.07	0.28	-0.01
60	0.35	0.06	0.35	0.01	0.30	-0.05	0.29	-0.01
70	0.39	0.05	0.40	0.01	0.35	-0.05	0.33	-0.02
100	0.39	0.06	0.39	0.00	0.35	-0.04	0.33	-0.02
Kelso 1								
	Winter		Spring		Summer		Autumn	
Depth (cm)	Mean	Change	Mean	Change	Mean	Change	Mean	Change
10	0.32	0.03	0.35	0.03	0.24	-0.11	0.29	0.05
20	0.37	0.06	0.38	0.01	0.28	-0.10	0.30	0.03
30	0.38	0.05	0.39	0.01	0.31	-0.08	0.32	0.01
40	0.38	0.05	0.39	0.01	0.32	-0.07	0.32	0.01
50	0.38	0.05	0.39	0.01	0.32	-0.07	0.33	0.01
60	0.38	0.05	0.39	0.01	0.34	-0.05	0.33	0.00
70	0.37	0.04	0.38	0.01	0.34	-0.05	0.33	-0.01
100	0.35	0.02	0.37	0.01	0.35	-0.02	0.34	-0.01
Kelso 2								
	Winter		Spring		Summer		Autumn	
Depth (cm)	Mean	Change	Mean	Change	Mean	Change	Mean	Change
10	0.36	0.07	0.38	0.02	0.25	-0.14	0.29	0.05
20	0.38	0.07	0.39	0.00	0.29	-0.10	0.31	0.02
30	0.39	0.06	0.40	0.00	0.33	-0.07	0.34	0.01
40	0.39	0.04	0.40	0.00	0.35	-0.05	0.35	0.01
50	0.39	0.05	0.39	0.00	0.34	-0.05	0.34	0.00
60	0.37	0.04	0.38	0.01	0.34	-0.04	0.34	0.00
70	0.37	0.04	0.38	0.01	0.34	-0.04	0.33	0.00
100	0.37	0.03	0.37	0.00	0.34	-0.03	0.33	-0.01

comparison to much of Canada, and periodic freezing and thawing are experienced throughout the winter months, especially in the upper soil layers. Freezing conditions are known to impact TDR measurements (Flerchinger et al 2006; Kahimba et al. 2007), an effect which is not accounted for directly herein, or in the native McMaster Mesonet dataset. Due to the low

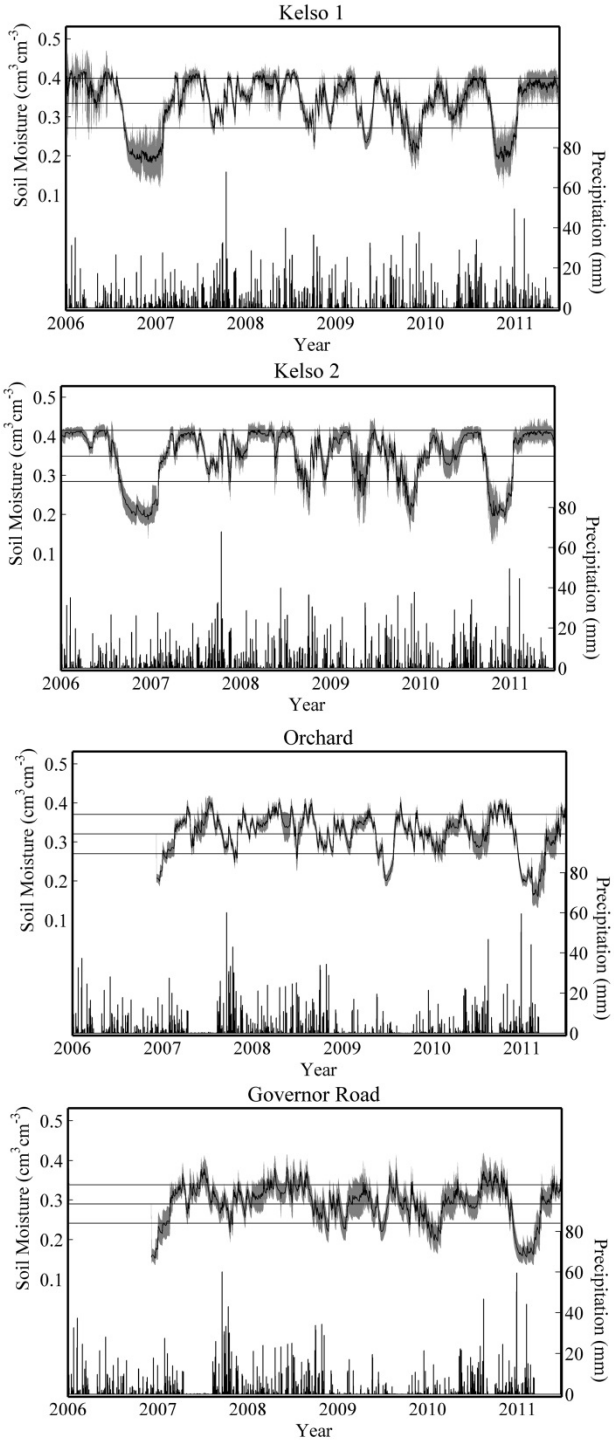

Figure 2-6: Daily timeseries of mean soil moisture in the top 50cm (black line) and precipitation (bars) for the McMaster Mesonet. The middle horizontal lines represent the mean of all observations at each site and one standard deviation, where the gray shaded area represents the spatial standard deviation of each sampling day.

variability of soil moisture, near saturated conditions and comparatively mild conditions throughout the winter season, freezing is not expected to substantially impact the data or analysis. However, further research and data collection is required to confirm this assumption.

The further soil moisture data analyses conducted herein are only for the 10, 20 and 50 cm depths using the data from the CS616 TDR probes, as soil moisture observations are available for every site and station at these depths. The following sections will present the results of the spatio-temporal and temporal stability analysis. First, the monthly mean and standard deviation will be discussed with links made to the local soil moisture state and variability as a result of local ground cover and topography. The temporal stability analysis will follow, focusing on seasonal trends, temporal consistency of the mean relative difference ranks and the time required for spatial organization to return following a disturbance.

2.5.1 Spatio-temporal Analysis

A time series of the mean daily soil moisture in the top 50 cm is shown in Fig. 2-6, along with the precipitation at each site. As previously noted, the soil moisture exhibits seasonality, with dry periods during summers

and recharge of the soil moisture during the winter and spring. The distribution of the soil moisture about the temporal mean shows the temporal soil moisture to be negatively skewed (GR = -0.89; OR = -0.97; K1 = -0.90; K2 = -1.02), largely resulting from summer dry-down, particularly in 2007 and 2011, and the upper boundary imposed by saturation. The poor drainage and low infiltration of K1 and K2 results in longer periods near the saturation boundary, whereas GR and OR have higher infiltration and topographic runoff, resulting in a more rapid transition of the soil moisture state. The spatial variability tends to be relatively uniform through time at all sites, although this is somewhat visually masked by the vertical transitions in Fig. 2-6. The greatest spatial variability, as represented by the shaded area, tends to occur during periods with little transitions, or the plateaus in Fig. 2-6. The greater standard deviation during the non-transitional periods is the result of the spatial heterogeneity induced by geographic factors, such as depressions between hummucks at K1 and K2 storing water, or the redistribution of water to valleys at GR.

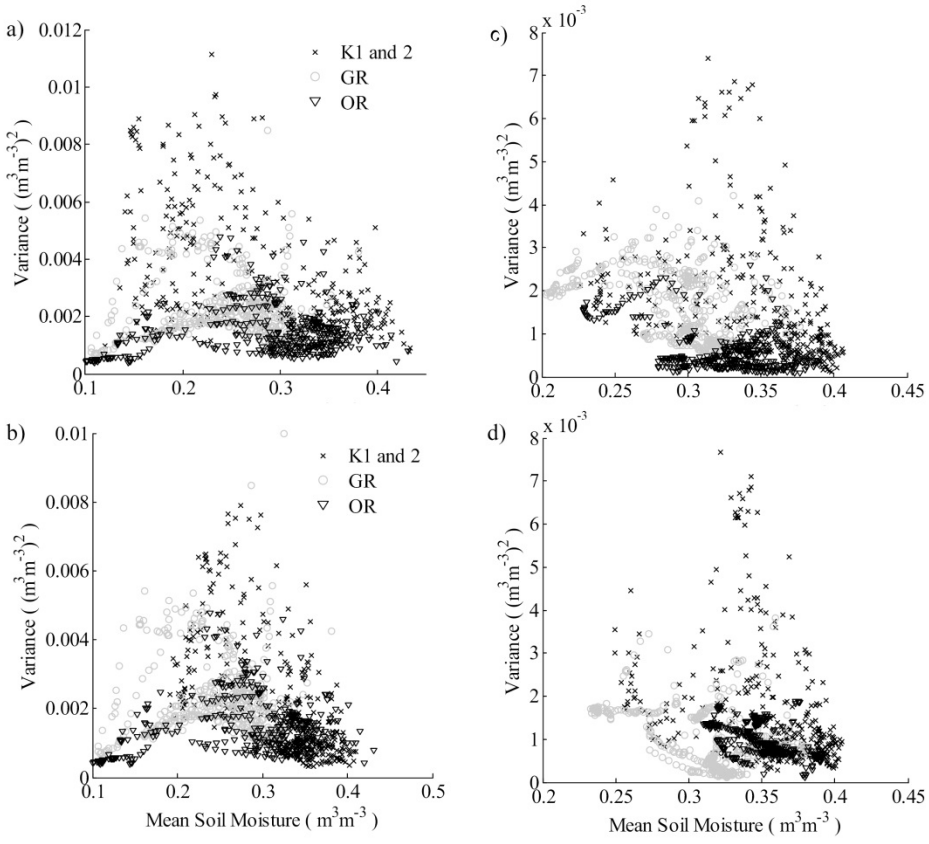

Figure 2-7: Relationship between mean soil moisture and variance for the McMaster Mesonet sites at depths of (a) 10cm, (b) 20cm, (c) 30cm and (d) 50cm.

Figure 2-7 shows the relationship between the spatial mean and variance of soil moisture for each site and depth. At the 10 and 20 cm depths variability is highest at intermediate quantities of soil moisture at GR and OR, resulting in a slight concave down parabolic relationship. At K1 and K2 variability generally decreases with increasing soil moisture. The concave down relationship was also found by Famiglietti et al. (2008) at extents greater than 800 m, whereas the decreasing mean-variance relationship of Kelso has been more often found at the field scale (Brocca et al. 2010; Brocca et al. 2007; Western et al. 2004; Hupet and Vanclooster, 2002), particularly in humid climates (Brocca et al. 2007). The concave down relationship at GR and OR likely results from topographic redistribution, an influence under-represented in hillslope scale studies, and suggests that topography is a primary control and climate a secondary control of the change in variability with mean soil moisture state. Soil moisture at the 30 and 50 cm depths became less variable as soil moisture increased, however, the relatively high soil moisture content at these depths resulted in few observations where the mean moisture content was less than 0.25 cm^3cm^{-3}, the range where variability decreased in the surface measurements at GR and OR.

Since the soil moisture in the watershed exhibits seasonality, the data have been aggregated on a monthly basis for the years 2007 to 2011 and analyzed as spatial soil moisture patterns. The seasonal (monthly) mean and seasonal (monthly) standard devation of soil moisture for each site can be seen in Fig. 2-8 and Supp. 1 respectively as contour plots, which were created by interpolating the soil moisture values between points. For the purposes of making physical interpretation of the soil moisture patterns, the results will be presented as first an inter-site comparison making broad generalizations, and then intra-site comparisons making note of specific anomalies and patterns.

Both K1 and K2 have uniformly distributed soil moisture at all depths and are persistently wetter than the other sites. This results from the lower infiltration capacity and flat topography at this site. Also, the hummocky terrain provides many small depressions in which surface water is stored. OR also has a relatively uniform soil moisture pattern at all depths, although there is a marked spatial pattern at the 10 and 20 cm depths resulting from the sloping terrain. GR has the most variability in soil moisture pattern during all months resulting from the complex topography and vegetation. All sites experience the highest variability (standard

deviation) during the month of January, which is believed to be the result of freeze-thaw

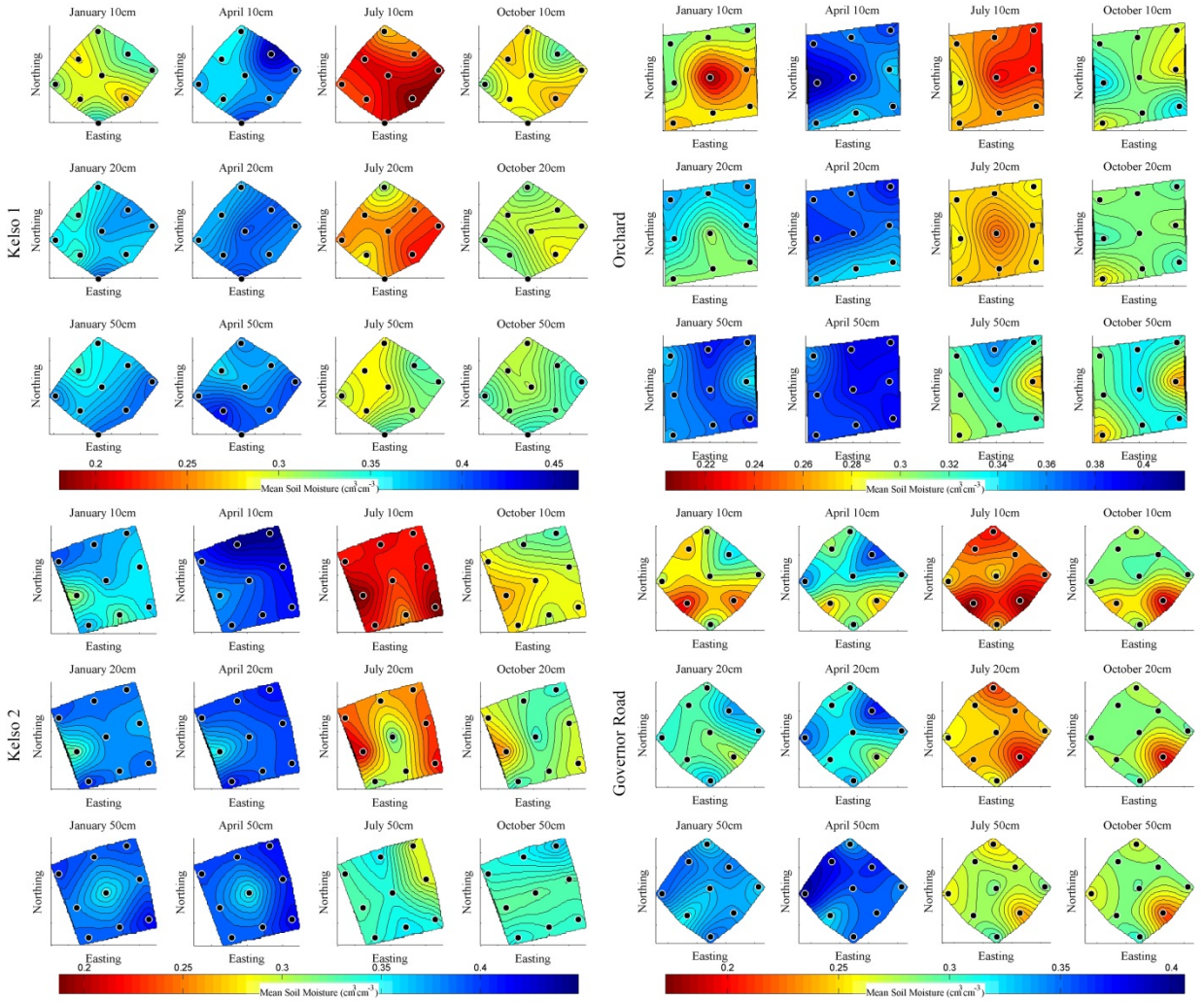

Figure 2-8: Seasonal (monthly) mean soil moisture of McMaster Mesonet sites. The contour plots are created by a Natural Neighbour Interpolation between observed soil moisture points. A complimentary contour plot of seasonal (monthly) standard deviation can be found in the supplementary material.

processes with temperature fluctuations around 0°C causing sporadic redistribution of soil moisture, and inducing errors in the TDR measurements. The lowest variability is during the spring season, when the soil is consistently wet due to snowmelt and low evapotranspiration. During the summer months, there is moderate variability, brought on by wetting and drying cycles.

In general, the results of the soil moisture analysis are consistent with previous studies (Famiglietti et al. 1998), which found that there tends to be low soil moisture values and high variability at the top of slopes and high soil moisture and low variability at the bottom of slopes,

with moderate values mid-slope. This pattern was also present at K1 and K2, where there are few prominent terrain features, with the notable exceptions of a small embankment at the south-east of K1 and a hill in the north-east of K2. These features result in lower mean soil moisture, which is particularly prominent during the summer and fall months and result in the greatest contrast in standard deviations during the winter months. The slope pattern in Kelso is also consistent, although somewhat muted, to a depth of 50 cm, however it declines at 70 to 100 cm (results not shown). This pattern is also generally present at OR, with the exceptions of OR-1 which has low soil moisture and high variability but is located mid-slope. This anomaly results from a combination of a slight depression in the surface at OR-1 temporarily storing water, and decreased vegetation density due to human disturbance and a protecting fence. Similarly, OR-5 and 8 are located on the upper and midslope respectively and have lower mean and more highly variable soil moisture than expected. At OR-8 this results from the sensor being located near a tree, whereas OR-5 is on a side-slope facing a roadway, where drainage to the road side is expected. While still present, the impact of the minor surface features are not as prevalent at the 20 and 50 cm depth suggesting that, at lower depths, soil moisture is more representative of general patterns in the landscape. At GR, stations 7 and 4 have low mean soil moisture and moderate to high variability due to their location near the crown of a hill. Similarly, the higher moisture content present in the valleys is consistent with previous findings (Famiglietti et al. 1999) and demonstrates the importance of topographic redistribution at small scales and the long correlation length of variability induced by topography.

2.5.2 Temporal Stability Analysis

The temporal stability of a station refers to the rank of a station relative to the mean soil moisture state at a particular site. It has often been noted that while absolute soil moisture changes, the ranking of soil moisture at a particular location with respect to the mean value is relatively constant (Vachaud et al. 1985; Martinez-Fernandez and Ceballos, 2003; Cosh et al. 2004; Vivoni et al. 2008). Here, the temporal stability is analyzed for the entire length of record, amalgamating the Kelso stations as well as GR and OR due to proximity, and each site has been analyzed on a seasonal (monthly) basis for the 10, 20 and 50 cm depths.

2.5.3 Long-term Temporal Stability

Figure 2-9 shows the results of the temporal stability analysis at each depth for the entire duration of the dataset. Stations that are consistently positive are wet with respect to the spatial mean soil moisture, whereas negative values correspond to dry stations. RMSE, represented as the dotted line, gives an indication of stations which are both stable and representative of the mean. There is generally greater variability in the relative difference of the 10 and 20 cm depths than the 50 cm depth showing that variability is greatest near the surface (Choi and Jacobs, 2007) and that rank stability is less important with depth (Heathman et al. 2009). The relative difference is generally consistent across depths, where only minor changes in rank position occur. Exceptions to this are the result of small scale features, where vegetation roots, canopy or depression storage impacts the relative difference pattern more strongly at certain depths (Famiglietti et al. 1998). OR-1, GR-9 and K-19 are examples of this, where the position of the surface and sub-surface soil moisture relative to the mean changes with depth. The distribution of both OR and GR stations throughout the relative difference plot suggests that topography is more important in determining the spatial distribution of soil moisture than vegetation which was also noted by Cosh et al. (2004) in the Walnut Creek Experimental Watershed. Generally, the stations with the highest

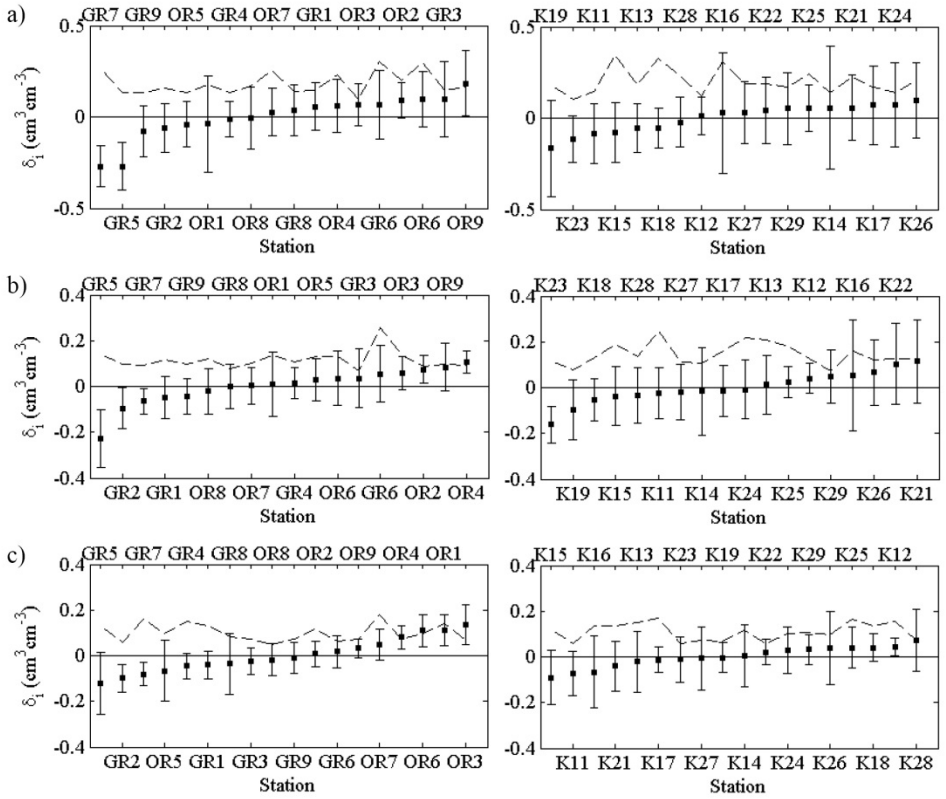

Figure 2-9:Mean relative difference for the entire study period at the McMaster Mesonet sites at depths of (a) 10cm, (b) 20cm, (c) 50cm. GR and OR are grouped due to their close spatial proximity, as are K1 and K2. The bars represent one standard deviation of the relative difference and the dotted line shows the RMSE from Eqn (3-9).

rank are located at the bottom of the slopes, whereas stations with the lowest rank are located in upslope areas. This is also true of the Kelso sites, where the change in elevation is minor. However, the relatively flat topography at Kelso results in less stability in the spatial distribution of soil moisture than is observed at OR and GR, as indicated by the standard deviation of the relative difference.

2.5.4 Seasonal Temporal Stability

The seasonal spatial patterns of the mean relative difference are similar to the mean soil moisture patterns found in Fig. 2-8, and so will not be re-produced here. For the flat/gently sloping terrain at K1, K2 and OR the relative rank of each station remains generally consistent throughout time. Unlike the other sites, there is a marked change in temporal rankings at GR throughout the year as seen in Fig. 2-10. At the monthly scale, a similar $\overline{\delta_{ij}}$ may result in some stations changing

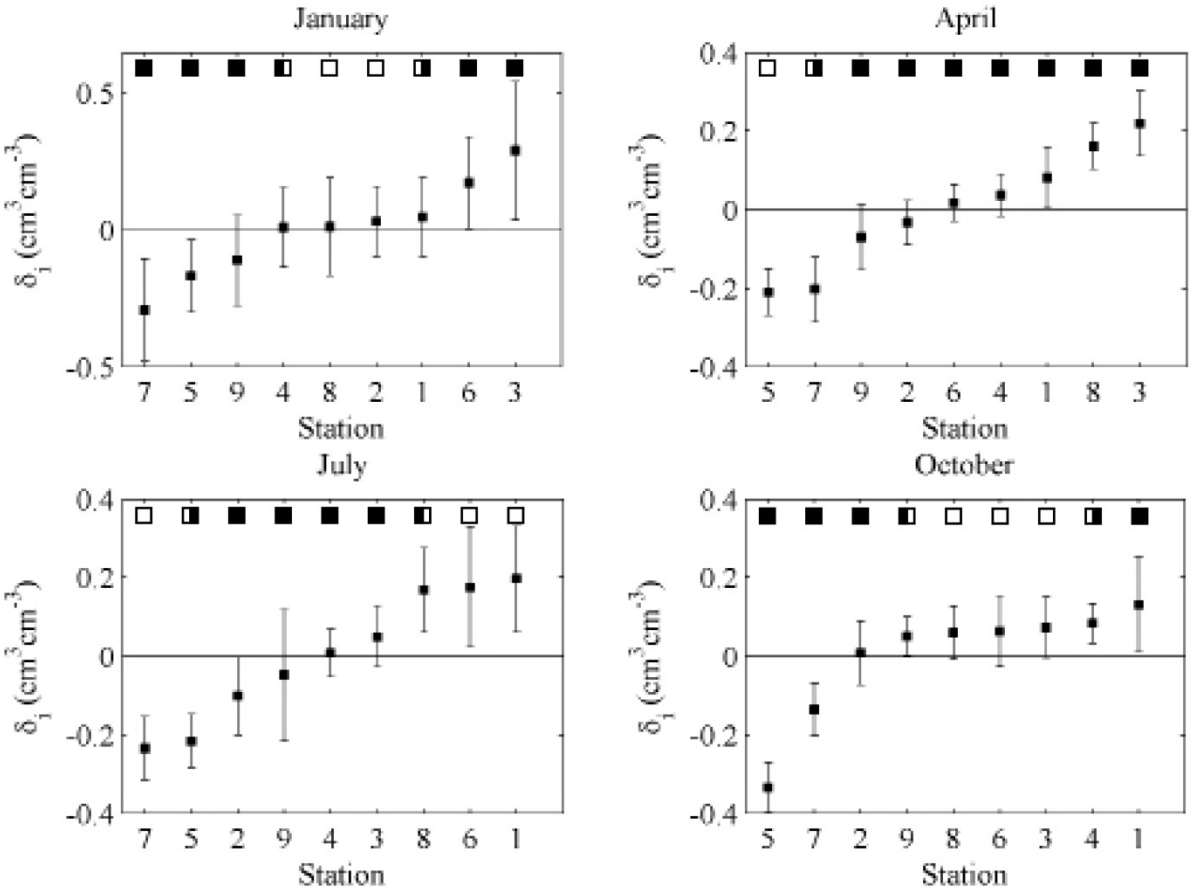

Figure 2-10:Seasonal temporal stability at GR at a depth of 10cm. The boxes show the mean relative difference and the bars one standard deviation of the relative difference. The white half boxes at the top of the figure show where a station is similar (α = 0.05) to its neighbor on the respective side.

ranked position. For example, in Fig. 2-10 stations 7 and 5 both represent dry stations but have changed rank order as the driest station. Often due to the overlap in variance of $\overline{\delta_{ij}}$, especially for mean representative sites, the change in station rank order is not meaningful (α = 0.05). An example is Station 4 in Fig. 2-10, where the station is always representative of the mean or median value but changes rank order.The relative difference for the stations at GR have generally low variability, but shift at the 10 and 20 cm depths. The greatest change in rank occurs at GR-3 and 4. During the winter and spring the downslope station (GR-3) is consistently the highest ranked station (wet), however, during the summer and fall months GR-3 becomes mid-ranked and is more representative of the site mean soil moisture. A noticeable feature in Fig. 2-10, is the presence of a definitive slope or stronger ranking during the spring and summer periods. In comparison, the October mean relative difference plot shows most stations, except for GR-5 and 7, have a relatively flat trend.

It has been suggested that temporal stability analysis can be used to select monitoring sites representative of the field/catchment mean (Brocca et al. 2010; Jacobs et al. 2004). The results show the need for caution when using this approach as sites representative of the mean under a particular set of soil conditions (seasons) were not representative during all time periods, and the suggestion to use multiple mid-ranked sites is prudent. The change in rank of the wet stations (GR-3) is unlike other findings that wet and dry sites are stable (Cosh et al. 2008; Jacobs et al. 2004; Martinez-Fernandez and Ceballos, 2003) and indicates a transition in influences controlling spatial distribution, where topography dominates in wet conditions (Heathman et al. 2009; Western and Bloschl, 1999; Famiglietti et al. 1998) and vegetation and soil texture dominate in dry conditions (Vivoni et al. 2008; Hupet and Vanclooster, 2002). The demonstrated 'seasonality' should be taken into account when using temporal stability at the hillslope to catchment scale in complex terrain, especially where vegetation changes may influence soil moisture variability (Hupet and Vanclooster, 2002).

Since the majority of soil moisture monitoring programs have been carried out in meadows and agricultural fields (Famiglietti et al. 1998; 1999; Brocca et al. 2007; Heathman et al. 2009; and others), GR offers unique insight into the importance of higher density vegetation in determining the soil moisture spatial pattern. It has been shown that at small scales precipitation increases homogeneity with wetting, which is subsequently redistributed by

topography, soil texture and evapotranspiration following wetting in areas with low tree density (Wilson et al. 2003; Famiglietti et al. 2008), and is consistent with results at K1, K2 and OR. However, GR is covered by a mix of forested areas and open meadows, resulting in vegetation causing a seasonal organization in soil moisture during wetting. The seasonality of the mean relative difference patterns shows the possibility of temporal instability in seasonal climates such as those present in mid-latitudes.

2.5.5 Persistence of Ranks

The fundamental concept of temporal stability is the persistence of the relative spatial pattern of soil moisture, which has already been demonstrated to be more distinct during the spring and summer months. However, it is known that precipitation has a homogenizing effect on the soil moisture pattern and can disrupt the temporal stability pattern at all scales (Wilson et al. 2003; Famiglietti et al. 2008), which may hamper the use of temporal stability in hydrological applications (Loew and Mauser, 2008). Due to the collection of hourly data at the McMaster Mesonet a preliminary assessment of the potential of precipitation to impact temporal stability can be performed. Four rain storms were selected to analyze the length of interruption of the temporal stability pattern and to provide insight as to the duration required for the dominant soil moisture pattern to re-emerge. The four rain events were selected based on two primary criteria: (i) significant rainfall being present at both Kelso (K1) and Dundas Valley (GR & OR) and (ii) a sufficient period without rain is present both prior to and following the rain event so as to allow

Table 2-4: Characteristics of analyzed rain events. If a rain event is broken into two distinct rainfalls, the amount of rain in each is separated, and the duration is separated as first rainfall duration, gap length, and second rainfall duration.

Rain Event	Date and Time	Amount of Rain (mm)	Duration (hours)
	Kelso (K1)		
1	11 Jul 2009 – 0900-1100 h	6.2	3
2	28-29 Aug 2009 – 2200-1100h	22.1 / 0.1	8 / 5 / 1
3	15-16 Mar 2011 – 2300-1000h	6.1	12
4	3 Aug 2011 – 0300-2200h	8.6 / 1.6	4 / 10 / 6
	Dundas Valley (GR & OR)		
1	11 Jul 2009 – 0900-1400h	9.0 / 0.1	2 / 1 / 1
2	28-29 Aug 2009 – 2200-2000h	32.2 / 3.2	11 / 11 / 1
3	15-16 Mar 2011 – 2200-1000h	4.8	13
4	03 Aug 2011 – 1400-2400h	1.6 / 1.0	2 / 3 / 6

for the presence of a stable soil moisture pattern before and after the rain. The details of the four selected storms can be found in Table 2-4, and the relative difference for storm 2 is shown in Fig. 2-11. Only storm 2 at 10 cm is presented herein for brevity and K2 is ignored because of the close proximity and similar topography to K1.

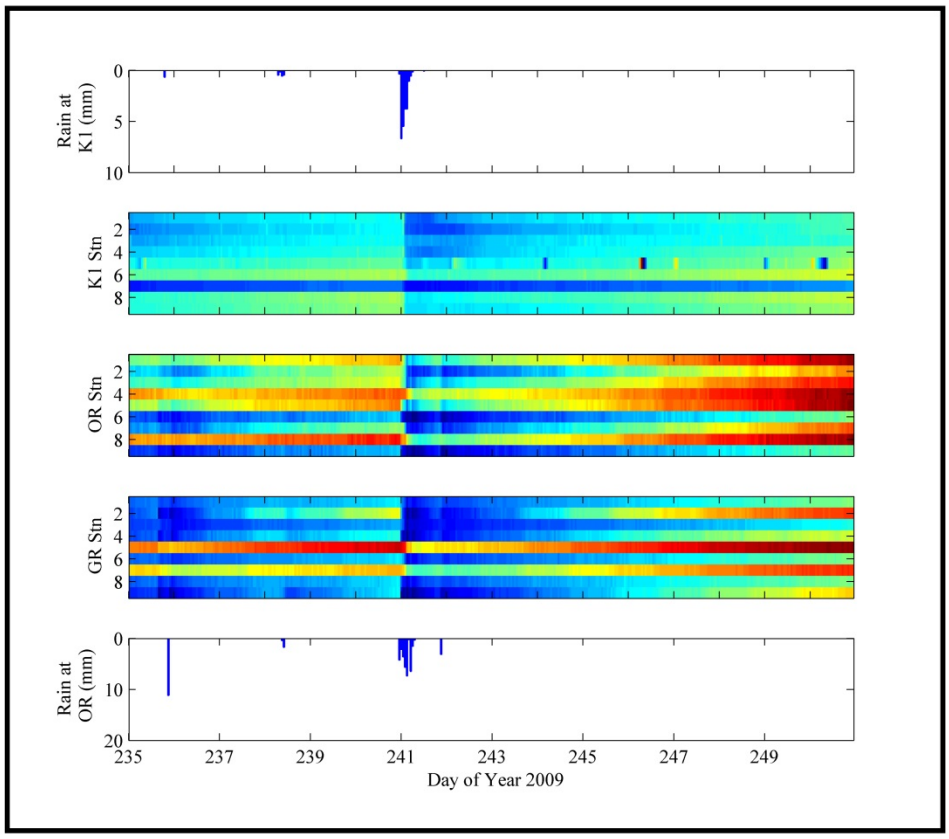

Figure 2-11: Analysis of the mean relative difference before and after rain event 2 at a depth of 10 cm at K1, OR and GR. The colour scale changes from red (dry) to blue (wet) where mean soil moisture values (δ_{jk} = 0) is cyan. For visual interpretation the colourmap is stretched for each image and colour values are relative.

Due to the dry conditions prior to rainfall and the sloping topography, rain event 1 did not have a distinguishable impact on the soil moisture pattern at any site, however it did cause noticeable wetting at K1. While some change occurred immediately following the precipitation, the peak of the impact at K1 occurred 24 to 40 hours following the rain event as ponded water slowly infiltrated the poorly drained soil. Rain storm 2 added significantly more water to the soil than the other storms and had a strong wetting effect at all sites, with the greatest impact at the 20 cm depth. At K1, the storm caused wetting of the already moist soil, but did not disrupt the relative ranks of the spatial soil moisture pattern. Due to the relatively higher clay content and

flat topography the impact of the rain also persisted longer at K1 compared to the other sites, with the impact at some stations lingering for as many as 6 days after the rainfall. Similarly, at OR the rain caused a large increase in the amount of soil moisture but did not lead to the homogenization of soil moisture, and thus the persistent pattern was still present. At GR, stations 5 and 7, which are under dense vegetation and on the hill-top respectively, experienced some wetting but were less impacted than other stations and were stable as relatively dry sites during that time period. The remaining GR sites, which were still wet from a previous rainfall, became homogenous disrupting the temporal stability pattern. The relative soil moisture pattern began to re-emerge within 18 to 24 hours following the peak rainfall. At all sites the spatial pattern had the greatest disruption which was sustained for the longest period of time at the 20 cm depth. The lack of surface evaporation and already greater homogeneity at this depth caused the effect of the precipitation to linger for approximately 2 days longer than at the 10 cm depth. While there was some wetting, the rainfall did not affect the spatial soil moisture pattern at 50 cm at any site. The rainfall from events 3 and 4 did result in an increase in soil moisture but was not strong enough to impact the relative spatial pattern at any site or depth with the exception of OR at 20 cm. This site was already wet from snow melt and spring rain and so additional water did disrupt the persistence of the spatial pattern.

These results lead to several insights about the nature of temporal stability following rain. It should first be noted, that without the soil being wet prior to the addition of water, the temporal stability of the soil moisture was not impacted without the addition of a considerable amount of rain (>20 mm). The normal rainfall amounts for the watershed (2 to 10 mm) did not cause enough homogenization to affect the temporal (rank) stability pattern, even at the flat and poorly drained Kelso site. Also important is the impact of topography and vegetation as was best noted at the GR site. Dense vegetation resulted in high interception leaving GR-5 persistently dry, whereas GR-4 is located on a uphill area and so while its rank was impacted by rainfall, the higher potential for soil water redistribution caused that station to recover its relative rank with only a short delay.

2.6 Conclusions

The McMaster Mesonet was introduced and spatio-temporal and temporal stability analysis carried out in order to characterize the soil moisture patterns representative of the Mesonet datasets. The McMaster Mesonet consists primarily of the infrastructure necessary to monitor hourly soil moisture profiles using a high spatial resolution TDR array at four sites, as well as associated hydro-meteorological stations. The sites represent different topographies and vegetation covers as well as providing some insight into the seasonal patterns of soil moisture experienced in the climatic context of the southern Great Lakes Basin.

Analysis of the data reveals a moderately strong organized soil moisture pattern which is temporally persistent on a seasonal basis. Topography represents a dominant control of the soil moisture pattern in wet to intermediate soil moisture conditions, whereas vegetation, soil texture and micro-topography have greater influence in intermediate-dry conditions. The spatial pattern at sub-seasonal scales was persistent in most conditions, unless a substantial rainfall/snowmelt resulted in homogenization of the soil moisture. Following disturbance by precipitation, the relative soil moisture pattern re-emerged after 18 to 24 hours of drying and the moisture added by the precipitation was removed within 4 to 7 days. At the sub-seasonal time scale, soil moisture variability was the greatest at a depth of 10 cm, and was lowest at a depth of 50 cm. The spatial pattern of soil moisture at a depth of 20 cm was more representative of mean topography, where the influence of hummocks and small depressions at the surface was not as prevalent.

These types of analysis of soil moisture patterns at multiple temporal scales are only possible where monitoring is continuous and of a long duration and high temporal resolution such as the McMaster Mesonet. Unlike studies with shorter duration or random sampling periods, these results lead to the conclusion that hydrological applications of the temporal (rank) stability concept should account for temporal (seasonal) changes, especially in complex terrain. Where topography was simple, as at Kelso, spatial variability was greatest in dry conditions due to storage in hummocks and differences in vegetation cover. At GR and OR, variability was highest at intermediate soil moisture conditions as topography and vegetation both provided a strong control on the soil moisture spatial pattern.

The long term high-spatial resolution hourly soil moisture profiles recorded by the McMaster Mesonet can provide insights into the nature of this important hydrological state variable. Future work will use data from the McMaster Mesonet for data assimilation to improve hydrological forecasts, downscaling and validation of remote sensing soil moisture products and to study the complex interactions between climate, soil, topography and vegetation.

2.7 References

Alavi, N., Berg, A. A., Warland, J. S., Parkin, G., Verseghy, D. and Bartlett, P.: Evaluating the Impact of Assimilating Soil Moisture Variability Data on Latent Heat Flux Estimation in a Land Surface Model, Can. Water Resour. J., 35, 157-172, 2010.

Albergel, C., R¨udiger, C., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., and Martin, E.: From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations, Hydrol. Earth Syst. Sci., 12, 1323–1337, doi:10.5194/hess-12-1323-2008, 2008.

Blöschl, G. and Sivapalan, M.: Scale issues in hydrological modelling: A review. Hydrol. Process., 9, 251-290. 1995.

Bosch, D. D., Lakshmi, V., Jackson, T. J., Choi, M. and Jacobs, J. M.: Large scale measurements of soil moisture for validation of remotely sensed data: Georgia soil moisture experiment of 2003, J. Hydrol., 323, 120-137, 10.1016/j.jhydrol.2005.08.024, 2006.

Brocca, L., Morbidelli, R., Melone, F. and Moramarco, T.: Soil moisture spatial variability in experimental areas of central Italy, J. Hydrol., 333, 356-373, 10.1016/j.jhydrol.2006.09.004, 2007.

Brocca, L., Melone, F., Moramarco, T. and Singh, V. P.: Assimilation of Observed Soil Moisture Data in Storm Rainfall-Runoff Modeling, J. Hydrol. Eng., 14, 153-165, 10.1061/(ASCE)1084-0699(2009)14:2(153), 2009.

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

Bronstert, A. and Bardossy, A.: The role of spatial variability of soil moisture for modelling surface runoff generation at the small catchment scale, Hydrol. Earth Sys. Sci., 3, 505-516, 1999.

Campbell Scientific: On-Line Estimation of Grass Reference Evapotranspiration with the Campbell Scientific Automated Weather Station: App Note Code: 4-D. Logan, UT. 1999.

Ceballos, A., Scipal, K., Wagner, W. and Martinez-Fernandez, J.: Validation of ERS scatterometer-derived soil moisture data in the central part of the Duero Basin, Spain, Hydrol. Process., 19, 1549-1566, 10.1002/hyp.5585, 2005.

Chen, Y.: Letter to the editor on "Rank stability or temporal stability", Soil Sci. Soc. Am. J., 70, 306-306, 10.2136/sssaj2005.0290l, 2006.

Choi, M. and Jacobs, J. M.: Soil moisture variability of root zone profiles within SMEX02 remote sensing footprints, Adv. Water Resour., 30, 883-896, 10.1016/j.advwatres.2006.07.007, 2007.

Cosh, M. H., Jackson, T. J., Bindlish, R. and Prueger, J. H.: Watershed scale temporal and spatial stability of soil moisture and its role in validating satellite estimates, Remote Sens. Environ., 92, 427-435, 10.1016/j.rse.2004.02.016, 2004.

Cosh, M. H., Jackson, T. J., Moran, S. and Bindlish, R.: Temporal persistence and stability of surface soil moisture in a semi-arid watershed. Remote Sens. Environ., 112(2), 304-313, 10.1016/j.rse.2007.07.001, 2008.

Das, N. N., Mohanty, B. P., Cosh, M. H. and Jackson, T. J.: Modeling and assimilation of root zone soil moisture using remote sensing observations in Walnut Gulch Watershed during SMEX04, Remote Sens. Environ., 112, 415-429, 10.1016/j.rse.2006.10.027, 2008.

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A. and Jackson, T.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, Hydrol. Earth Syst. Sci., 15, 1675-1698, 10.5194/hess-15-1675-2011, 2011.

Dumedah, G. and Coulibaly, P.: Evaluation of statistical methods for infilling missing values in high-resolution soil moisture data, J. Hydrol., 400, 95-102, 10.1016/j.jhydrol.2011.01.028, 2011.

Famiglietti, J. S., Rudnicki, J. W. and Rodell, M.: Variability in surface moisture content along a hillslope transect: Rattlesnake Hill, Texas, J. Hydrol., 210, 259-281, 10.1016/S0022-1694(98)00187-5, 1998.

Famiglietti, J. S., Devereaux, J. A., Laymon, C. A., Tsegaye, T., Houser, P. R., Jackson, T. J., Graham, S. T., Rodell, M. and van Oevelen, P. J.: Ground-based investigation of soil moisture variability within remote sensing footprints during the Southern Great Plains 1997 (SGP97) Hydrology Experiment, Water Resour. Res., 35, 1839-1851, 10.1029/1999WR900047, 1999.

Famiglietti, J. S., Ryu, D., Berg, A. A., Rodell, M. and Jackson, T. J.: Field observations of soil moisture variability across scales, Water Resour. Res., 44, W01423, 10.1029/2006WR005804, 2008.

Flerchinger, G. N., Seyfried, M. S. and Hardegree, S. P.: Using soil freezing characteristics to model multi-season soil water dynamics, Vad. Zone J., 5, 1143-1153, 10.2136/vzj2006.0025, 2006.

Gillespie, J. E., Wicklund, R. E. and Miller, M. H.: The Soils of Halton County: Report No. 43 of the Ontario Soil Survey. Ontario Department of Agriculture. 80 pp. 1971.

Heathman, G. C., Larose, M., Cosh, M. H. and Bindlish, R.: Surface and profile soil moisture spatio-temporal analysis during an excessive rainfall period in the Southern Great Plains, USA, Catena, 78, 159-169, 10.1016/j.catena.2009.04.002, 2009.

Hupet, F. and Vanclooster, M.: Intraseasonal dynamics of soil moisture variability within a small agricultural maize cropped field, J. Hydrol., 261, 86-101, 10.1016/S0022-1694(02)00016-1, 2002.

Ivanov, V. Y., Fatichi, S., Jenerette, G. D., Espeleta, J. F., Troch, P. A. and Huxman, T. E.: Hysteresis of soil moisture spatial heterogeneity and the "homogenizing" effect of vegetation, Water Resour. Res., 46, W09521, 10.1029/2009WR008611, 2010.

Jacobs, J. M., Mohanty, B. P., Hsu, E. C. and Miller, D.: SMEX02: Field scale variability, time stability and similarity of soil moisture, Remote Sens. Environ., 92, 436-446, 10.1016/j.rse.2004.02.017, 2004.

Kahimba, F. C. and Sri Ranjan, R.: Soil temperature correction of field TDR readings obtained under near freezing conditions, Can. Biosys. Eng., 49, 1.19-1.26, 2007.

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

Kornelsen, K. and Coulibaly, P.: Comparison of interpolation, statistical and data driven methods for imputation of missing values in a distributed soil moisture dataset, 2013, J. Hydrol. Eng. Accepted

Lebel, T., Cappelaere, B., Galle, S., Hanan, N., Kergoat, L., Levis, S., Vieux, B., Descroix, L., Gosset, M., Mougin, E., Peugeot, C. and Seguis, L.: AMMA-CATCH studies in the Sahelian region of West-Africa: An overview, J. Hydrol., 375, 10.1016/j.jhydrol.2009.03.020, 2009.

Loew, A. and Mauser, W.: On the disaggregation of passive microwave soil moisture data using a priori knowledge of temporally persistent soil moisture fields, IEEE Trans. Geosci. Remote Sens., 43, 819-834, 10.1109/TGRS.2007.914800, 2008.

Loew, A., Schwank, M. and Schlenz, F.: Assimilation of an L-Band Microwave Soil Moisture Proxy to Compensate for Uncertainties in Precipitation Data, IEEE Trans. Geosci. Remote Sens., 47, 2606-2616, 10.1109/TGRS.2009.2014846, 2009.

Mahanama, S. P. P., Koster, R. D., Reichle, R. H. and Zubair, L.: The role of soil moisture initialization in subseasonal and seasonal streamflow prediction - A case study in Sri Lanka, Adv. Water Resour., 31, 1333-1343, 10.1016/j.advwatres.2008.06.004, 2008.

Martinez-Fernandez, J. and Ceballos, A.: Temporal stability of soil moisture in a large-field experiment in Spain, Soil Sci. Soc. Am. J., 67, 2003.

Merlin, O., Chehbouni, A., Boulet, G. and Kerr, Y.: Assimilation of disaggregated microwave soil moisture into a hydrologic model using coarse-scale meteorological data, J. Hydrometeorol., 7, 1308-1322, 10.1175/JHM552.1, 2006.

Minet, J., Laloy, E., Lambot, S. and Vanclooster, M.: Effect of high-resolution spatial soil moisture variability on simulated runoff response using a distributed hydrologic model, Hydrol. Earth Sys. Sci., 15, 1323-1338, 10.5194/hess-15-1323-2011, 2011.

Mohanty, B. P., Famiglietti, J. S. and Skaggs, T. H.: Evolution of soil moisture spatial structure in a mixed vegetation pixel during the Southern Great Plains 1997 (SGP97) Hydrology Experiment, Water Resour. Res., 36, 3675-3686, 10.1029/2000WR900258, 2000a.

Mohanty, B. P., Skaggs, T. H. and Famiglietti, J. S.: Analysis and mapping of field-scale soil moisture variability using high-resolution, ground-based data during the Southern Great Plains 1997 (SGP97) Hydrology Experiment, Water Resour. Res., 36, 1023-1031, 10.1029/1999WR900360, 2000b.

Mohanty, B. P. and Skaggs, T. H.: Spatio-temporal evolution and time-stable characteristics of soil moisture within remote sensing footprints with varying soil, slope, and vegetation, Adv. Water Resour., 24, 1051-1067, 10.1016/S0309-1708(01)00034-3, 2001.

Mohr, K. I., Baker, R. D., Tao, W. K. and Famiglietti, J. S.: The sensitivity of West African convective line water budgets to land cover, J. Hydrometeorol., 4, 62-76, 10.1175/1525-7541(2003)004<0062:TSOWAC>2.0.CO;2, 2003.

Peters-Lidard, C. D., Pan, F. and Wood, E. F.: A re-examination of modeled and measured soil moisture spatial variability and its implications for land surface modeling, Adv. Water Resour., 24, 1069-1083, 10.1016/S0309-1708(01)00035-5, 2001.

Presant, E. W., Wicklund, R. E. and Matthews, B. C.: The Soils of Wentworth County: Report No. 32 of the Ontario Soil Survey. Ontario Department of Agriculture. 72 pp. 1965.

Ryu, D. and Famiglietti, J. S.: Characterization of footprint-scale surface soil moisture variability using Gaussian and beta distribution functions during the Southern Great Plains 1997 (SGP97) hydrology experiment, Water Resour. Res., 41, W12433, 10.1029/2004WR003835, 2005.

Ryu, D. and Famiglietti, J. S.: Multi-scale spatial correlation and scaling behavior of surface soil moisture, Geophys. Res. Lett., 33, L08404, 10.1029/2006GL025831, 2006.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B. and Teuling, A. J.: Investigating soil moisture-climate interactions in a changing climate: A review, Earth-Sci. Rev., 99, 125-161, 10.1016/j.earscirev.2010.02.004 ER, 2010.

Torres, R., Dietrich, W. E., Montgomery, D. R., Anderson, S. P. and Loague, K.: Unsaturated zone processes and the hydrologic response of a steep, unchanneled catchment, Water Resour. Res., 34, 1865-1879, 1998.

Vachaud, G., Silans, P. D., Balabanis, P. and Vauclin, M.: Temporal stability of spatially measured soil water probability density function, Soil Sci. Soc. Am. J., 49, 822-828, 1985.

Vanderlinden, K., Vereecken, H., Hardelauf, H., Herbst, M., Martínez, G., Cosh, M. and Pachepsky, Y.: Temporal stability of soil water contents: A review of data and analyses, Vadose Zone J., 11, 10.2136/vzj2011.0178, 2012.

Vivoni, E. R., Gebremichael, M., Watts, C. J., Bindlish, R. and Jackson, T. J.: Comparison of ground-based and remotely-sensed surface soil moisture estimates over complex terrain during SMEX04, Remote Sens. Environ., 112, 314-325, 10.1016/j.rse.2006.10.028, 2008.

Western, A. W. and Bloschl, G.: On the spatial scaling of soil moisture, J. Hydrol., 217, 10.1016/S0022-1694(98)00232-7, 1999.

Western, A. W., Zhou, S. L., Grayson, R. B., McMahon, T. A., Bloschl, G. and Wilson, D. J.: Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes, J. Hydrol., 286, 113-134, 10.1016/j.jhydrol.2003.09.014, 2004.

Wilson, D. J., Western, A. W., Grayson, R. B., Berg, A. A., Lear, M. S., Rodell, M., Famiglietti, J. S., Woods, R. A. and McMahon, T. A.: Spatial distribution of soil moisture over 6 and 30 cm depth, Mahurangi river catchment, New Zealand, J. Hydrol., 276, 254-274, 10.1016/S0022-1694(03)00060-X, 2003.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Doell, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffe, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A. and Whitehead, P.: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, Water Resour. Res., 47, W05301, 10.1029/2010WR010090, 2011.

Chapter 3: Root-zone Soil Moisture Estimation using Data-Driven Methods

Summary of Paper 2: Kornelsen, K.C. and Coulibaly, P. (2014) Root-zone Soil Moisture Estimation using Data-Driven Methods, Water Resources Research, 50, WR014127, doi: 10.1002/2013WR014127.

Summary:

The goal of this research was to determine if artificial neural networks (ANN) could be used to reliably estimate the soil moisture state in the root-zone from observations of surface soil moisture. In most ANN applications the trained network is site specific. It was sought to remove that limitation by training the ANN using simulations from the HYDRUS-1D model. A bootstrap aggregating technique was used to derive confidence limits of the ANN output and to conduct a sensitivity analysis. By conducting the sensitivity analysis on the 'bagged' ANNs, this paper uniquely explored the sensitivity of the ANN to model inputs and the consistency with which the ANN were sensitive to those inputs.

The results of this research demonstrate:

- ANNs were capable of representing the non-linear dynamics of shallow (≤ 30 cm) soil moisture using surface observations and meteorological information.
- The sensitivity analysis showed that surface soil moisture, evapotranspiration and solar radiation were consistently the most important for simulating root-zone soil moisture.
- The ANN could be applied at a regional scale, which in this case was validated using several independent sites in the Great Lakes Basin.
- The use of HYDRUS-1D to train the ANN provided a limitation, where the ANN modelled root-zone soil moisture with high performance at sites where HYDRUS-1D modelled soil moisture with high performance, but suffered when HYDRUS-1D did not represent local soil moisture well.
- An ensemble approach using multiple training runs and multiple ANN configurations resulted in better estimation of root-zone soil moisture and was robust when individual neural networks encountered conditions for which they were less accurate.

3.1 Abstract

The soil moisture state partitions both mass and energy fluxes and is important for many hydrogeochemical cycles, but is often only measured within the surface layer. Estimating the amount of soil moisture in the root-zone from this information is difficult due to the non-linear and heterogeneous nature of the various processes which alter the soil moisture state. Data-driven methods, such as artificial neural networks (ANN), mine data for non-linear interdependencies and have potential for estimating root-zone soil moisture from surface soil moisture observations. To create an ANN root-zone model that was non-site specific and physically constrained, a training set was generated by forcing HYDRUS-1D with meteorological observations for different soil profiles from the UNSODA database. Ensemble ANNs were trained to provide soil moisture at depths of 10, 20 and 50 cm below the surface using surface soil moisture observations and local meteorological information. Insights into the processes represented by the ANNs were derived from a clamping sensitivity analysis and by changing the ANNs input data. Further model testing based on synthetic soil moisture profiles from three McMaster Mesonet and three USDA SCAN sites suggests that ANNs are a flexible tool capable of predicting root-zone soil moisture with good accuracy. It was found that ANNs could well represent soil moisture as estimated by HYDRUS-1D, but performance was reduced in comparison to *in situ* soil moisture observations outside the training conditions. The transferability of the model appears limited to the same geographic region.

3.2 Introduction

Soil moisture is an important hydrological state variable which controls the partition between infiltration and runoff as well as latent and sensible heat flux. The last decade has seen an increased focus placed on both active and passive microwave remote sensing to provide spatially distributed soil moisture information [Kornelsen and Coulibaly, 2013a; Kerr et al. 2010; Entekhabi et al. 2010]. Particular focus has been placed on the dedicated Soil Moisture and Ocean Salinity (SMOS) and Soil Moisture Active Passive (SMAP) missions, which both have a goal of retrieving soil moisture with an accuracy of 0.04 cm^3cm^{-3} [Kerr et al. 2010; Entekhabi et al. 2010]. Unfortunately, these and other remote sensing instruments are only able to collect soil

moisture information to an estimated depth of approximately 5 cm [Kerr et al. 2010]. Despite these advances, a gap still exists with respect to soil moisture in the root-zone, which is relevant for hydro-meteorological processes such as plant transpiration and hydraulic redistribution [Seneviratne et al. 2010]. Since surface soil moisture is coupled to root-zone soil moisture through diffusion processes, it is possible to implement retrieval algorithms to return profile soil moisture from measurements of the surface state [Albergel et al. 2008; Calvet and Noilhan 2000; Kumar et al. 2010; Singh 2010; Wagner et al. 1999]. The spatial and temporal patterns of surface soil moisture tend to have greater variability related to meteorological forcing, whereas the temporal dynamics of the subsoil are diminished due to soil water redistribution [Rosenbaum et al. 2012].

In recent years, the integration of surface soil moisture measurements into a hydrological model through data assimilation has proven a promising approach to predict root-zone soil moisture [Das and Mohanty, 2006; Walker et al. 2001a; b; Lu et al. 2010; Dumedah and Coulibaly, 2012; Reichle et al. 2012]. The data assimilation approach estimates profile soil moisture by updating the soil moisture state in a hydrological model through merging the observed and modeled surface soil moisture, accounting for the error in each, and propagating the state updates into the soil profile either solely through model physics [Walker et al. 2001a; Heathman et al. 2003] or a measure of the covariance between the profile and surface state [Walker et al. 2001b; Lu et al. 2010; Das and Mohanty, 2006]. The Ensemble Kalman Filter (EnKF) is a widely used land surface data assimilation algorithm and has also been coupled with HYDRUS-ET in the Soil Moisture Assessment Tool (SMAT) [Das and Mohanty, 2006]. The EnKF based approach is not only computationally demanding, but also has strong limitations for nonlinear systems [Clark et al. 2008]. Similarly, the evolutionary data assimilation technique proposed by Dumedah and Coulibaly [2012] also exhibits a high computational cost due to the multi-objective evolutionary search strategy. Thus, the appeal of effective and computationally efficient alternatives is obvious. A simple alternative is the regression approach, where a fitted curve is used to find the relationship between surface and subsurface soil moisture [Srivastava et al. 1997]. Singh [2010] proposed entropy theory for the prediction of root-zone soil moisture, where the maximization of Tsallis entropy characterizes the diffusion of water through the soil for a given time period and Wagner et al. [1999] proposed the use of an exponential filter to capture the dynamics of sub-surface soil moisture from a time series of surface measurements.

This study proposes using an ensemble of artificial neural networks (ANN) for the estimation of root-zone soil moisture from surface measurements. Neural networks are a strong candidate for this task as the relationship they model is determined from data, such that they can be presumed to account for soil water transmission characteristics as well as the impact of surface fluxes in a way which is representative of the known non-linear soil moisture interactions. In hydrology, ANNs have been found to be equally or more effective than various type of models for predicting streamflow [Khan and Coulibaly, 2006; Abrahart et al. 2012] and have been used in soil moisture hydrology as pedotransfer functions (PTF) [Schaap et al. 2001] and for calculating soil water retention curves [Jain and Srinivasulu 2004]. Kornelsen and Coulibaly [2014] found that feed-forward neural networks were effective at infilling missing soil moisture data and Elshorbagy and Parasuraman [2008] tested neural networks for estimating soil moisture at depths up to 1 m using soil temperature and meteorological forcing.

The purpose of this study is to determine the feasibility of using ANNs as a computationally efficient and non-linear alternative to linear methods or physically based land data assimilation schemes for mapping root-zone soil moisture. The main objective is to obtain a flexible ANN model that captures the main spatio-temporal patterns at different sites to estimate root-zone soil moisture. Given that many regions of the world have limited soil moisture profile measurements, HYDRUS-1D has been used to generate artificial soil moisture profiles with which to train and validate the ANN models. Additional spatial testing is carried out by assimilating soil moisture measurements from the McMaster Mesonet and select USDA SCAN sites into HYDRUS-1D. The resulting ANN model is different from conventional site-specific ANN applications as it is expected to be capable of estimating point scale sub-surface soil moisture from surface observations and commonly available meteorological information within a geographic region. If trained with a representative data set, such a tool could be useful at large scale to retrieve root-zone soil moisture at any location.

3.3 Study Area and Data

The ANN soil moisture model is expected to be valid for conditions present in the Lower Great Lakes region of Canada and the U.S.A. Forcing data for HYDRUS and the ANN, as well as soil

moisture measurements to generate testing data, were collected by the McMaster Mesonet and three USDA SCAN sites located around Lake's Ontario and Erie (Fig 3-1).

The McMaster Mesonet time domain reflectometry (TDR) arrays and the associated hydro-meteorological stations are located in the Hamilton-Halton Watershed in southern Ontario, Canada [Kornelsen and Coulibaly, 2013b]. The 1250 km^2 watershed is predominantly rural agricultural land on mixed loam soils with urban areas concentrated around the shoreline of Lake Ontario. The Canadian Climate Normals (1971-2000) at the Hamilton International Airport (43°10'N, 79°55'W) show the region to have a humid continental climate with mean annual precipitation of 910 mm evenly distributed throughout the year, a mean summer temperature of 21°C and winter temperature of -6°C. A more comprehensive description of the watershed climate can be found in Kornelsen and Coulibaly [2013b].

Figure 3-1: Study sites in the Lower Great Lakes Basin.

The McMaster Mesonet consists of four TDR arrays referred to as 'sites'. Two sites are located adjacent to each other at Kelso (K1 and K2) in the northern part of the watershed, of which data from K2 were used herein, and two sites are located in close proximity in the south of the watershed at Governor Rd (GR) and Orchard (OR). At each site soil moisture is automatically sampled every hour at nine stations arranged in a grid with a spacing of 25 m at depths from 10 to 100 cm [Kornelsen and Coulibaly, 2013b]. Collocated with each TDR array is a tipping bucket rain gauge and nine hydro-meteorological stations are distributed throughout the watershed to provide necessary ancillary data. This study area was considered, because the high spatial density of soil moisture samples results in high confidence in the observed soil moisture values. In this study, we focused on meteorological data between day of year (DOY) 91-304 of 2009 to generate training/validation data and DOY 121-273 of 2011 for testing. This time frame equates to the frost free season in the watershed where the large date range in the training data provides a greater range of conditions than the testing data, which is necessary to ensure that ANN training captured all of the potential conditions.

Independent testing data from three USDA SCAN sites [Schaefer et al. 2007] in the north eastern United States covering a period of DOY 121-273 of 2011 were used to compliment the McMaster Mesonet data. The Geneva, NY (GN), Rock Springs, PA (RS), and Sunleaf Nursery, OH (SL) sites were selected based on their similar geography to that of the trained neural networks. This will test the use of neural networks as a point scale root-zone soil moisture estimation method valid within a specific region. Soil moisture data were collected at each SCAN site at 5, 10, 20 and 50 cm depths using Stevens Hydraprobes. Soil moisture data for the sites were downloaded from the International Soil Moisture Network [Dorigo et al. 2011]. The associated meteorological data and soil descriptive information were downloaded from the USDA SCAN website (http://www.wcc.nrcs.usda.gov/scan/). ANN testing was carried out for each SCAN site using meteorological data from the individual site assuming constant soil texture for the entire soil profile. A summary of both the McMaster Mesonet and SCAN sites can be found in Table 3-1.

Table 3-1: Site Description

Site	USDA Stn #	Sand (%)	Silt (%)	Clay (%)	Soil Texture	Drainage	Land-cover	LAI
K2		30	35	35	Clay Loam	Imperfect and Poor	Meadow Grass	1.71
GR		20	65	15	Silt Loam	Good	Mixed Forest	3.92
OR		20	65	15	Silt Loam	Good	Meadow Grass	1.71
GN	2011	48	37.8	14.2	Loam	Good	Grass/ Herbaceous	1.71
RS	2036	17.9	63.6	18.5	Silt Loam	Well	Grass/ Herbaceous	1.71
SL	2073	75.5	18.7	5.8	Sandy Loam	Moderately Well	Grass/ Herbaceous	1.71

3.4 Methods

3.4.1 Artificial Neural Networks

The ANN discussed here are the conventional feed-forward neural networks or multilayer perceptron (MLP), which are the most commonly used neural network type and have proven effective in a variety of applications [Khan and Coulibaly, 2006; Jain and Srinivasulu 2004; Abrahart et al. 2012]. The differences in the MLP configurations will be outlined and the reader

is referred to Haykin [1999] and other appropriate references for relevant details on neural networks. In a conventional MLP, a vector of inputs $\mathbf{x}^n$ is passed through a series of non-linear hidden neuron activation functions $G(\cdot)$ to an output layer y_k via a series of optimized weight matrices $\mathbf{w}_j$. For a MLP the output of the network is given by:

$$y_n = f\left\{\sum_{j=1}^{h} \mathbf{w}_j \mathbf{G}(\mathbf{s}_i) + b_k\right\} \tag{3-1}$$

where f is the activation function of the output neuron k, is the output neurons bias and $\mathbf{s}_i$ is the weighted sum of the input data to the hidden neuron activation functions for each layer j. A neural network is trained using a representative dataset $D = \{\mathbf{x}^n, t^n\}_{n=1}^{N}$ by adjusting the network parameters **w** (weights and biases) to minimize an error function such as the mean squared error:

$$E_D = \frac{1}{N}\sum_{n=1}^{N}\{y_n(\mathbf{x}^n;\mathbf{w}) - t_n\}^2 \tag{3-2}$$

For each input/target pair {**x**, t}, the network output y(**x**; **w**) is calculated for the entire series to measure the error between the target t_n and the network output y_n. The error is minimized by repeated evaluation of the gradient of E_D using variants of the back propagation algorithm. In order to obtain a parsimonious network structure, to ensure that output is driven by inputs and to avoid the tendency of the MLP to 'overfit' the data, it is sometimes beneficial to additionally minimize the values in the weight matrix. Regularization modifies the network error function to:

$$E(\mathbf{w}) = \beta E_D + \alpha E_W \tag{3-3}$$

where $E_W = \frac{1}{N}\sum_{n=1}^{N}\mathbf{w}_j^2$ and the parameters β and α are hyper-parameters where $\alpha = 1 - \beta$. The hyper-parameter β can either be selected manually or can be derived using the Bayesian Regularization (BR) training algorithm [MacKay, 1992]. The MLPs were trained using gradient descent known as backpropagation.

The Broyden-Fletcher-Goldfarb-Shanno quasi-Newton (BFGS) training algorithm determines the gradient of the cost function using an approximation of Newton's method by calculating an approximate Hessian matrix. The Levenberg-Marquardt (LM) algorithm shifts from a gradient descent towards a quasi-Newton's method near an error minimum and therefore runs faster than BFGS and can converge to a more stable minimum. Both of these training algorithms allow for a regularized cost function that balances the error and weight matrix values.

This can be accomplished manually or automatically using Bayesian regularization (BR) which is the LM algorithm with a Bayesian update of the regularization parameter. Specific characteristics of the ANN models used are described latter herein. When training the neural networks, 70% of the data were used for network training, 15% for cross-validation and early stopping and 15% were reserved for the model validation.

3.4.2 Estimating Uncertainty in Neural Networks

Since ANN's are inherently deterministic models, the evaluation of confidence bounds is a difficult but important task [Abrahart et al. 2012]. One approach is to use the method of bootstrap aggregating (bagging) [Breiman, 1996]. Using a bagging approach, several neural networks (often 25) with the same architecture are trained with a subset $\{D_k\}$ of the training set $\{D\}$. The training subset for each network was selected by a statistical bootstrap method, where individual input vectors are randomly selected from the entire dataset with replacement and D_k had the same number of elements as D. The result was an ensemble of MLPs each trained with a unique training subset that maintained the statistical properties of the original sample while decreasing the possibility of model over-fitting. The final output and uncertainty are derived from the mean and standard deviation of the MLP ensemble. The result of bagging is greater stability in the final MLP prediction and lower tendency to over-fit the training data [Breiman, 1996].

3.4.3 HYDRUS-1D Description

HYDRUS-1D (referred to herein as HYDRUS) was used in this study to establish a training set for the MLP models. HYDRUS uses a Galerkin-type linear finite element scheme to numerically solve the governing modified Richards equation:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left\{ K\left[\frac{\partial h}{\partial z} + \cos\alpha \right] \right\} - S \tag{3-4}$$

where h is the water pressure head {cm}, θ is the volumetric water content {cm^3cm^{-3}}, t is time {T}, z is the depth {cm} (positive upward), S is the sink term {$cm^3cm^{-3}cm^{-1}$}, α is the angle between the flow directions and the vertical axis, and K is the unsaturated hydraulic conductivity function {cm T^{-1}}. This form of the Richards equation assumes air plays an insignificant role in liquid flow and the thermal gradients are neglected in the water flow. In this application $\alpha = 0$, and therefore only vertical soil water flow is assumed. After Das and Mohanty [2006], each soil profile was considered as an individual stream-tube where lateral flows both in the subsurface and across the surface were ignored for reasons of simplicity. The dependence between the unsaturated hydraulic conductivity K and volumetric soil moisture θ is determined by the relationships of van Genuchten [1980].

The sink term S is defined as the volume of water removed due to plant water uptake. It is represented by the formulation of Feddes et al. [1978] given by:

$$S(h) = \alpha(h) S_p \tag{3-5}$$

where $\alpha(h)$ is the dimensionless root-water uptake stress response function of the soil pressure head and S_p is the water uptake rate during periods of no water stress. The roots are presumed to be non-uniformly distributed, increasing linearly to a distribution peak at 10 cm and decreasing linearly from 10 cm to 100 cm. The potential water uptake from this root-zone is given as:

$$S_p = b(z) T_p \tag{3-6}$$

where $b(z)$ is the described root distribution and T_p is the potential transpiration rate. When root-water uptake does not proceed at its potential, the S_p is modified by the root-water uptake stress response function given by van Genuchten [1987]:

$$\alpha(h) = \left\{1 + (h / h_{50})^p\right\}^{-1} \tag{3-7}$$

in which h is the soil water pressure head and h_{50} is the soil water pressure head at which $a(h)$ is reduced by 0.50. In this study we assumed h_{50} = -1200 cm and p = 2, which are the values used by Lu et al. [2010, 2011] and Homaee et al. [2002]. Because the purpose herein was to develop a generalized tool, this root distribution was considered independent of vegetation type and did not experience growth throughout the season.

HYDRUS was forced using hourly meteorological observations and the output used in this study is a profile of volumetric soil moisture generated at the end of each day (midnight local). The end of day soil moisture was selected because it could be practically implemented in

the model and because it allowed for the integration of daily forcing in the soil column. For ANN training, validation and testing, soil moisture is extracted from the 5, 10, 20 and 50 cm nodes of the HYDRUS profile.

3.4.4 Initial and Boundary Conditions

Due to the presumed unknown root-zone soil moisture state, the initial condition was assumed to follow a uniform gradient, where the first 20 days of simulation are dropped from analysis as model spin-up. The bottom boundary assumes a free-drainage condition at a depth of 100 cm, where the bottom flux depends on the hydraulic conductivity.

The upper boundary condition was determined by the soil moisture state of the surface node and the atmospheric conditions at the soil-air interface at time *t*. Following Ritchie [1972], potential evaporation and transpiration fluxes are calculated from potential evapotranspiration using Beer's law as:

$$\begin{aligned} T_p &= ET_p \left\{1 - \exp\left[-k \cdot LAI\right]\right\} \\ E_p &= ET_p \exp\left\{-k \cdot LAI\right\} \end{aligned} \tag{3-8}$$

where E_p and ET_p are potential evaporation and potential evapotranspiration respectively, LAI is the leaf area index and the constant k governs the radiation extinction by the canopy due to sun angle, the distribution of plants and the arrangement of leaves (assuming short grass k = 0.463 [after Nouvellon et al. 2000]). Potential evapotranspiration was calculated from the Penman-Monteith equation, and due to the unavailability of measurements, LAI is determined from literature values [Scurlock et al. 2001] and is presumed to be constant throughout the study period. Thus, the upper boundary condition was generally governed by the surface flux. In the event of rainfall, if the rainfall rate exceeds the soil infiltration rate, surface ponding was assumed to take place and excess water was removed at the end of each day. This condition allowed for small ponding on what were assumed relatively flat surfaces, where large amounts of precipitation were assumed to runoff at the end of each day. A similar approach was proposed by Das and Monhanty [2006], and the potential error of not allowing perpetual ponding for estimating soil moisture is assumed to be negligible. This is because infiltration excess events lasting for more than short periods of time are extremely rare and it is presumed soil conditions in which ponding would carry on for periods of greater than a few hours in the lower Great Lakes basin would be near saturation conditions.

3.3 Data Used

MLPs are often employed in hydrology as a site specific empirical model, however, in this application it was sought to obtain a flexible model that can be used at different sites within a target geographic region to estimate root-zone soil moisture. As a data driven model, it was important that the MLPs have a large and representative training set, from which a statistical relationship can be derived. HYDRUS was used as a forward model to generate the training set, and to provide large and varied training samples. In order to provide consistent evaluation of MLP performance, HYDRUS-EnKF was also used to derive a synthetic 'truth' dataset for independent model testing. A flowchart of the methodology applied can be found in Fig 3-2.

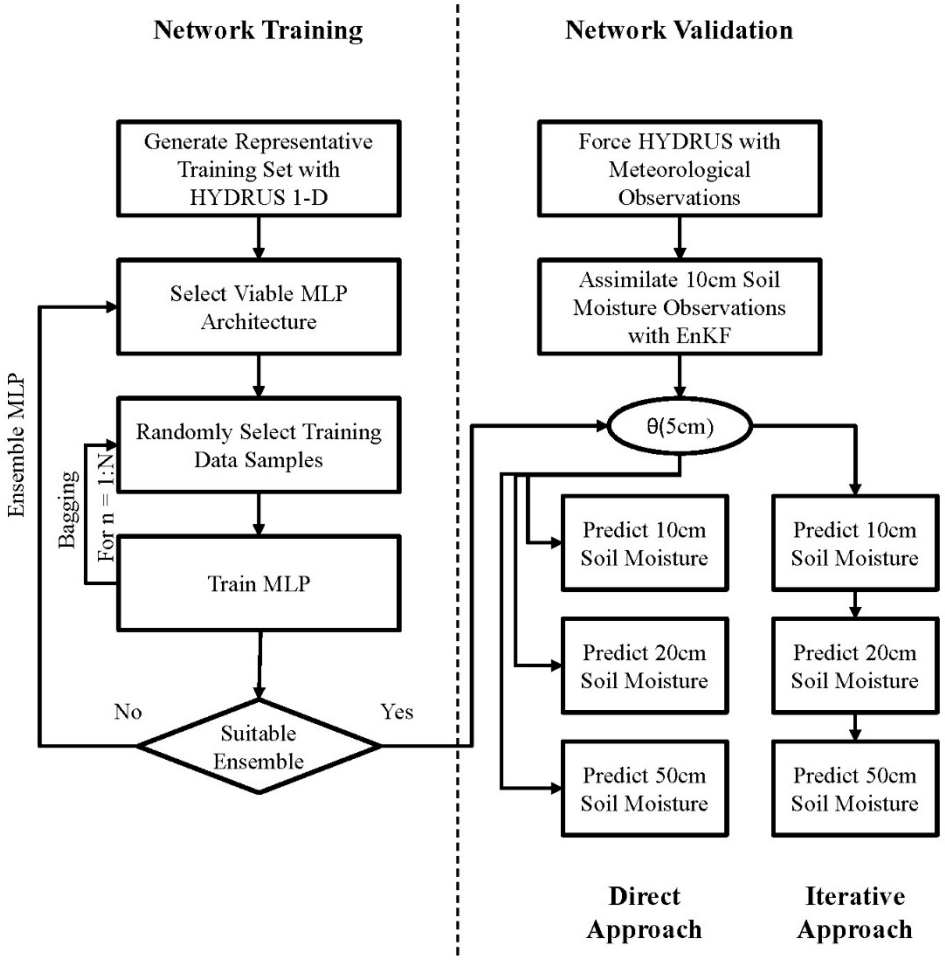

Figure 3-2: Flowchart of methods used. MLPs were trained using general UNSODA-HYDRUS derived data and cross-validated with a subset of the UNSODA-HYDRUS training set until a suitable ensemble was found. The performance of the selected networks was then tested by simulation of independent data representative of specific site conditions.

3.4.5 ANN Training Data and Target Data

The HYDRUS soil column was discretized into 101 nodes from 0 to 100 cm in depth distributed uniformly across the soil column. Each profile was assumed to contain one soil layer, where the soil hydraulic properties (van Genuchten parameters) were selected from the UNSODA database [Leij et al. 1996]. A restriction was put on the soil database, where only soils with a sand content of < 40 % were selected, such that the MLPs were only trained with soil profiles that could potentially be found in the southern Ontario/ north-eastern United States region, resulting in 554 soil profiles. For simplicity, it was assumed that the amount of soil water uptake and interception would be governed only by LAI, which was randomly selected for each HYDRUS profile. Since the HYDRUS profiles were used as "random" samples, LAI was held constant throughout the simulation year of each profile. HYDRUS was forced on an hourly time step using meteorological data collected from the Kelso weather station between DOY 91-304 of 2009. The end of day soil moisture values were extracted from the 5, 10, 20 and 50 cm nodes of the HYDRUS soil profile to serve as daily soil moisture ('targets') for MLP training. Forcing data used as MLP input were sampled to daily values as either mean (i.e. air temperature) or total (i.e. ET_o) as appropriate. Due to computational cost and practical limitations of computer memory, samples from 40 of the 554 soil profiles that had similar texture to other profiles, were assumed redundant profiles and were removed from the original training set. Finally, the actual total number of soil profiles was 514, from which samples were taken every other day. Therefore, the total training set consisted of 54 998 samples (input/target pairs) of which approximately 38 500 (70%) were selected and ordered randomly for network training, 8 250 (15%) were reserved for cross-validation, and another 15% for the validation of each network. The 15% cross-validation data were used for early stopping of the network training to prevent over-fitting and enhance model generalization ability. This cross-validation data was part of the UNSODA profile data, and was different from the independent test data described in Section 3.4.7.

In order to ensure the MLP did not over-represent days without rain and to provide some memory from the rainfall process, the Antecedent Precipitation Index (API) was used in lieu of observed rainfall. The API is given by:

$$API_n = \{API_{n-1} + N_{n-1}\}e^{\alpha t} \tag{4-9}$$

where the value of the index before rain API_n at time n is the sum of the previous index value and the previous rainfall N subject to the decay factor αt, taken as 0.9 [Benkhaled et al., 2004].

Two approaches were used for the soil moisture input to the MLPs. The first assumed that surface observations provided sufficient information for the estimation of sub-surface soil moisture and is known as the 'direct' approach. Therefore, soil moisture observations at 5 cm were used as input for the MLPs at all depths. The second approach removed this assumption by 'iterating' the network soil moisture predictions to the subsequent depth, where the predicted value for depth z was used as input for the model predicting soil moisture at depth z+1. For MLP training this was accomplished using idealized data extracted from HYDRUS, whereas the model validation propagated MLP predictions to the subsequent model. The MLP output was volumetric soil moisture at a particular depth and the inputs for each MLP model are summarized in Table 3-2. The standard (baseline) input for the MLP's consisted of soil moisture, temperature, relative humidity, solar radiation, wind speed, potential evapotranspiration, API, LAI, silt and clay content. The baseline input set (B) was used for most MLP architectures and deviations from this norm are listed in Table 3-2.

Table 3-2: Neural Network Inputs and Architecture.

MLP Number	Network	Inputs	Number of Hidden Neurons	Training Algorithm	Reg Ratio
	Baseline	B = θ, T, RH, SR, WS, ET, API, Silt, Clay, LAI	20	LM	0.0
1	LM-MLP(B,20,α0)	B	20	LM	0.0
2	BFGS-MLP(B,20,α0.2)	B	20	BFGS	0.2
3	BR-MLP(B,20,αA)	B	20	BR	Auto
4	LM-MLP(B,[15,15],α0)	B	[15 15]	LM	0.0
5	LM-MLP(B,20,α0.3)	B	20	LM	0.3
6	LM-MLP(B,20,α0.5)	B	20	LM	0.5
7	LM-MLP(B,20,α0.7)	B	20	LM	0.7
8	LM-MLP-T(B,20,α0)	B	20	LM	0.0
9	LM-MLP(M,15,α0)	M = θ, SR, Silt	15	LM	0.0
10	LM-MLP(S,15,α0)	S = θ, Silt, Clay	15	LM	0.0

Note: θ – soil moisture (vol.), T – temperature (°C), RH – relative humidity (%), SR – solar radiation (Wm-2), WS – windspeed (ms-1), ET – evapotranspiration (mm), API – antecedent precipitation index (-), LAI – leaf area index (-), LM – Levenberg-Marquardt Backpropagation, BFGS – BFGS quasi-Newton Backpropagation, BR – Bayesian Regularization

3.4.6 Independent Testing Data

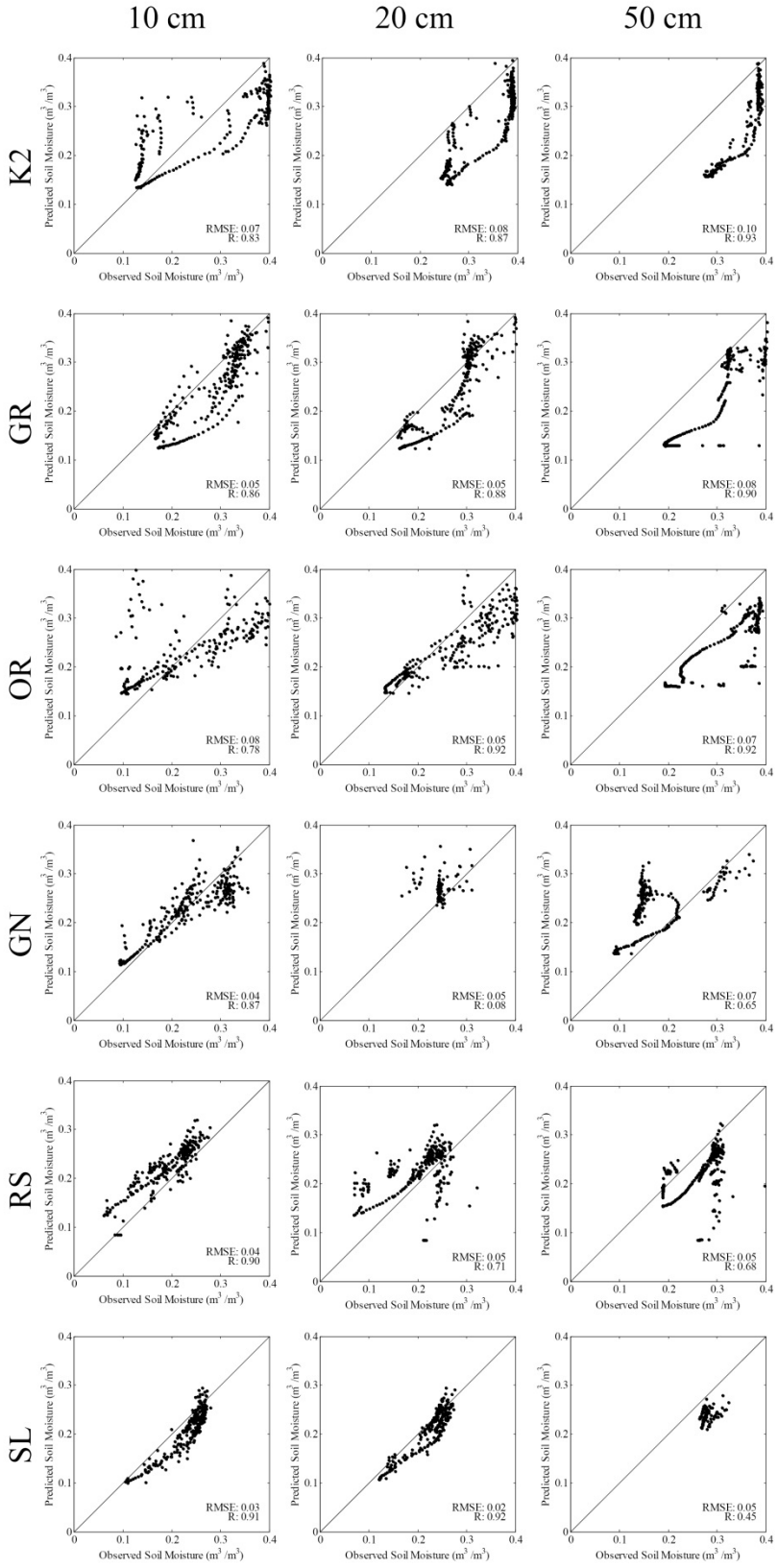

Figure 3-3: Evaluation of HYDRUS as an open loop model in comparison to in situ soil moisture observations from the McMaster Mesonet and SCAN sites.

Since soil moisture data were not available at a depth of 5 cm at the McMaster Mesonet and many missing values were sometimes present in the SCAN data, synthetic soil moisture was generated by assimilating soil moisture from the 10 cm observations into HYDRUS using the EnKF [Moradkhani et al. 2005]. The purpose of the data assimilation step, as opposed to an open loop model, was primarily to force the HYDRUS data to be more representative of actual site conditions. The MLP soil moisture estimation method was tested against the HYDRUS-EnKF synthetic soil moisture for the McMaster Mesonet and the SCAN network sites. This method was selected to ensure common soil moisture climatology between models and datasets, as well as eliminating missing values. Also, by comparing MLP

results to HYDRUS-EnKF derived soil moisture, errors in soil moisture predictions are representative of errors associated with the MLP and are less likely to be compounded by errors associated with HYDRUS. The *in situ* observed soil moisture is used only to evaluate the HYDRUS-EnKF results (Fig. 3-3) and as a time-series plot for comparison. Soil hydraulic properties for the test sites were derived from the Rosetta PTF [Schaap et al. 2001] using the soil texture information [Kornelsen and Coulibaly, 2013b; USDA Website: http://www.wcc.nrcs.usda.gov/scan/] and LAI for each vegetation cover type was estimated from literature values [Scurlock et al. 2001].

For the McMaster Mesonet, hourly soil moisture data from nine stations were aggregated first spatially and then to daily soil moisture. This aggregation was done using a simple arithmetic mean to provide the equivalent of a point data source. This method was chosen since the spatial coverage of each McMaster Mesonet site is relatively small (~50 m × 50 m) [Kornelsen and Coulibaly, 2013b] and the MLP methodology presented does not account for spatial heterogeneity at this scale. The SCAN sites served as further independent test data as neither the soil moisture nor forcing data from these sites were used in the development of the ANN training set. SCAN data were downloaded at the daily time step and are only available for a single profile at each site.

3.5 Results

3.5.1 Evaluation of HYDRUS

In order to justify the use of HYDRUS for the generation of the ANN training and validation datasets, HYDRUS was first evaluated at the McMaster Mesonet sites and the SCAN sites for comparison. HYDRUS was initialized by interpolating the observed soil moisture over the 101 nodes of the soil profile and forcing the model with hourly forcing data from each site. Model parameters for each site were selected using a PTF [Schaap et al. 2001] and literature values [Scurlock et al. 2001]. HYDRUS was run with all of the assumptions used in the generation of the ANN training data and no attempt was made to calibrate the model. Figure 3-3 shows that HYDRUS has an acceptable level of performance, particularly in the upper soil layers. Many of the discrepancies between the simulated and observed soil moisture can be explained by the simplifying assumptions during model development. For example, at K2 at all depths HYDRUS

tends to have a dry bias which results from the assumption that ponded water is removed daily, where the field site itself typically experiences ponding throughout most of the spring and fall months. Similarly, the assumptions regarding LAI, root distribution and homogeneity of the soil profile result in degraded simulation of soil moisture at the 50 cm depth, a limitation also found in other studies [Lu et al. 2010]. The limited plots at GN-20 cm and SL-50 cm are the result of many missing values in the observations from these sites.

3.5.2 Selection of Network Configuration and Sensitivity Analysis

In order to define the best neural network architecture, inputs, and training algorithms for root-zone soil moisture estimation, several network configurations were evaluated and a sensitivity analysis conducted using the training set. The best performing networks are presented in Table 3-2. The MLP notation *F-MLP*(x,h,α) indicates the network type (MLP), the training function F, the selected input set **x**, the number of neurons h in the hidden layer(s) **G** and the regularization ratio α.

Table 3-3: van Genuchten Parameter & Soil Texture Comparison.

	van Genuchten		Texture	
Depth	**RMSE**	**R**	**RMSE**	**R**
10 cm	0.02	0.92	0.02	0.92
20 cm	0.03	0.82	0.03	0.86
50 cm	0.03	0.81	0.03	0.86

Several potential forcing variables are provided by the weather stations and the selected inputs for each network can be found in Table 3-2. Some variables, such as saturation vapour pressure and wind direction, were excluded because the network was provided complimentary information (i.e. temperature and relative humidity) or the parameters were not considered to have a strong effect on soil moisture dynamics. An important aspect of the multilayer perceptron's ability to estimate root-zone soil moisture from surface measurements is the ability of the network to generalize the soils hydraulic conductivity and soil water retention given the surface soil moisture state. Using the baseline inputs of the training set, as described in Table 3-2, separate MLPs were trained replacing percent silt and clay with all of van Genuchten's soil water retention parameters $(\theta_r, \theta_s, \alpha, \eta, K_s, l)$ [van Genuchten, 1980]. The results in Table 3-3 demonstrate the ability of the network to infer soil moisture dynamics given only the soil texture, where the added details of the van Genuchten parameter set provided no benefit to RMSE and decreased the network performance in terms of correlation at all depths. An advantage of this result is the increased

flexibility and lower data demand, as soil texture information is more widely available than soil water retention parameters.

To gain insight into the influence of the input variables on network performance, a sensitivity analysis was conducted by a simple clamping method [Wang et al. 2000] using the training data set. After the MLP ensembles were trained using the bagging approach, they were used to simulate the soil moisture to produce a standard control run. Each input was then individually held constant (clamped) at its mean value and the network was used to predict soil moisture with the remaining inputs. The impact ratio between the control run $\left(y(\mathbf{x})\right)$ and the clamped run $\left(y\left(\mathbf{x} \mid x_i = \overline{x_i}\right)\right)$ for each input gives an indication of the networks sensitivity to the various inputs using Eq. 3-10 as seen in Fig 3-4.

$$\xi(x_i) = 1 - \frac{y\left(\mathbf{x} \mid x_i = \overline{x_i}\right)}{y(\mathbf{x})} \tag{3-10}$$

Since the sensitivity analysis is conducted on an ensemble of networks with the same architecture, the bars in Fig 3-4 provide an overall indication of the sensitivity of network architecture to a particular input, whereas the deviation bars (error bars) give an indication of the consistency of the network sensitivity. All network configurations were sensitive to the soil moisture input, with decreased sensitivity to surface soil moisture as depth increases. Of the meteorological inputs, evapotranspiration and solar radiation are the most important respectively. In some cases, ET_o is of similar or greater importance than soil moisture, particularly for the two standard LM networks, LM-MLP(B,20,α0), which have moderate sensitivity to soil moisture. This resulted in relatively poor performance of these two networks resulting in their being removed from the MLP ensembles presented in Section 3.3. This reflects the importance of water vapour flux in determining the soil moisture state. In contrast, these networks display relatively little sensitivity to API, likely because the addition of moisture through precipitation is accounted for in the soil moisture measurement and because there are fewer days in the watershed where there was a net gain of soil water from precipitation. When applied to the 20 and 50 cm layers, the MLP models trained with only surface measurements tended to be more sensitive to meteorological variables with increasing depth, whereas the models using the iterative approach tended to be decreasingly sensitive to meteorological variables in favor of soil moisture measurements. This is particularly evident when comparing Fig 3-4. a) and b): LM-

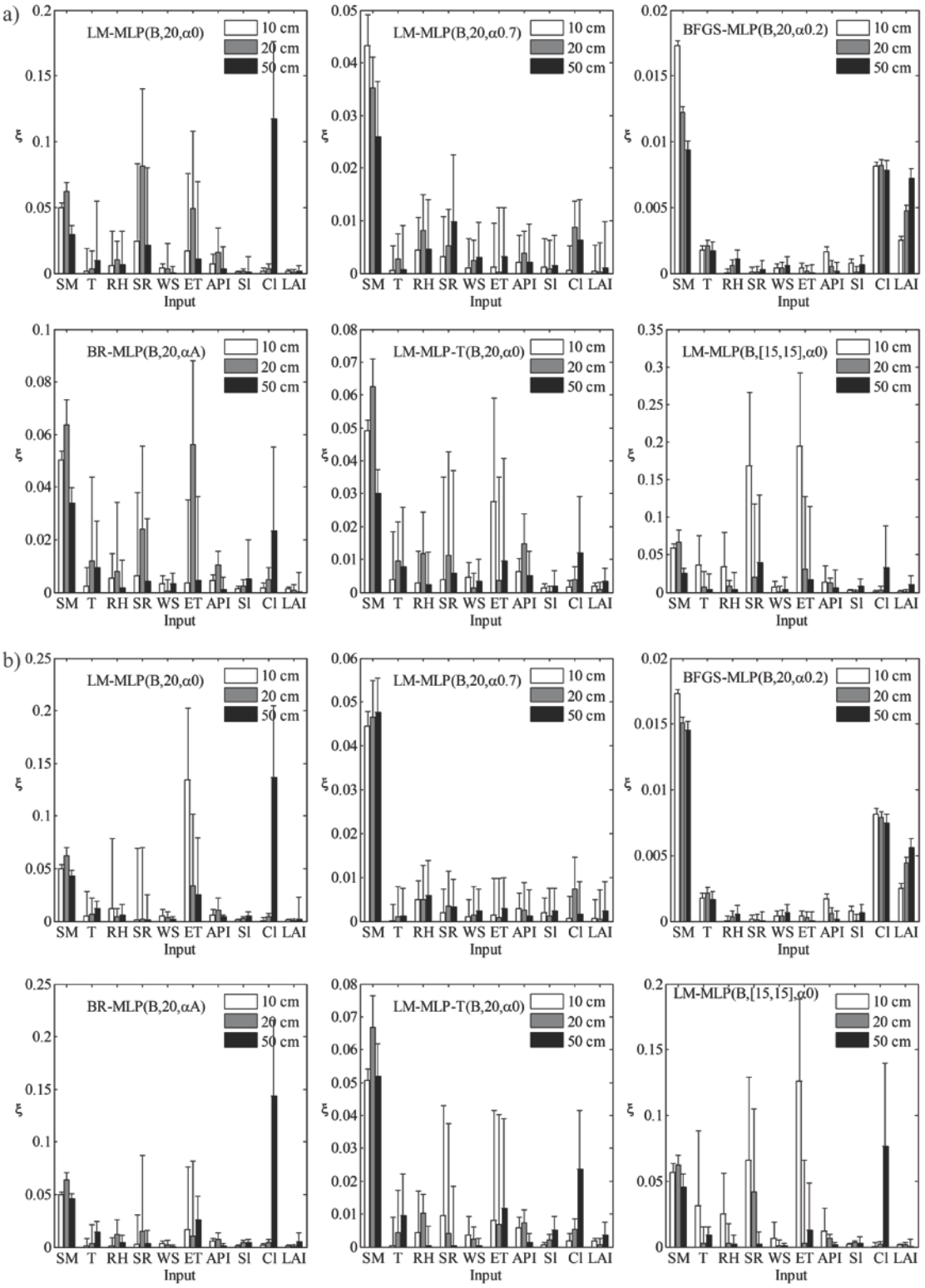

Figure 3-4: Sensitivity of select neural networks using the clamping method for (a) networks trained for direct soil moisture input and (b) networks trained with iterative soil moisture input. The absolute value of the impact ratio is presented, where each bar represents the ensemble mean output and the error bars are one standard deviation of the ensemble mean.

MLP(B,[15,15],α0). The –T in MLP number 8 (Fig 3-4a) indicates the output layer activation function has been changed from the standard linear function to a 'tansig' function. This modification resulted in slightly better accuracy for 20 and 50 cm predictions than the equivalent MLP with the linear output activation function. Neural networks trained with the BFGS algorithm had the most consistent sensitivities amongst the ensemble member networks.

The sensitivity analysis reveals that the most important variables for the prediction of soil moisture in the root-zone are upper layer soil moisture, soil clay content, ET_o and solar radiation. The sensitivity to ET_o likely results from the fact that it is calculated from the other forcing observations and therefore integrates information from the temperature, relative humidity, wind speed and solar radiation.

Since the principle of parsimony indicates preference should be given to simple models, networks were trained using only soil moisture measurements with soil texture or soil moisture measurements with solar radiation and silt content, LM-MLP(M,15,α0) and LM-MLP(S,15,α0) respectively. The skill of these simple networks in estimating subsurface soil moisture was often only slightly worse than networks with the entire baseline input set. This result suggests that the MLPs derive the majority of their estimate from the information provided by the soil moisture measurement and soil texture, with meteorological inputs providing a subtle influence to the soil moisture state. The impact of this was a greater constraint on the neural network prediction, which is evident in the time-series of Fig 3-5, where the use of only soil information tends to result in a network prediction with a slight positive bias and greater stochastic behaviour.

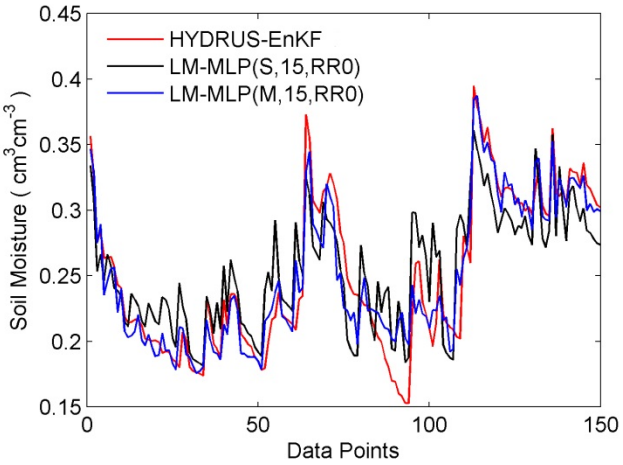

Figure 3-5: Prediction of soil moisture at 20cm from a random training data subset using neural networks with soil and baseline network inputs.

3.5.3 Independent Testing of Neural Network Model

Tables 3-4 and 3-5 summarize the performance metrics of the best performing bagged MLP architectures for the estimation of synthetic subsurface soil moisture from both the McMaster

Table 3-4: Testing Results of the Direct MLP Approach.

MLP	Kelso 2						Governor Rd						Orchard					
#	10		20		50		10		20		50		10		20		50	
	RMSE	R	RMSE	R	RMSE	R	RMSE	R	RMSE	R	RMSE	R	RMSE	R	RMSE	R	RMSE	R
3	0.02	0.95	0.07	0.82	0.05	0.68	0.03	0.95	0.05	0.73	0.06	0.71	0.02	0.96	0.04	0.82	0.05	0.67
4	0.03	0.93	0.07	0.75	0.06	0.52	0.08	0.69	0.05	0.79	0.09	0.35	0.07	0.67	0.04	0.81	0.07	0.37
5	0.02	0.96	0.04	0.86	0.05	0.76	0.02	0.96	0.04	0.86	0.05	0.77	0.02	0.96	0.04	0.85	0.04	0.73
6	0.02	0.96	0.04	0.85	0.06	0.70	0.03	0.96	0.05	0.85	0.05	0.78	0.02	0.96	0.04	0.85	0.04	0.74
7	0.02	0.96	0.04	0.87	0.06	0.75	0.02	0.96	0.04	0.85	0.06	0.73	0.02	0.96	0.04	0.84	0.04	0.73
8	0.02	0.96	0.05	0.85	0.05	0.71	0.02	0.96	0.04	0.79	0.05	0.72	0.02	0.96	0.04	0.83	0.05	0.67
9	0.02	0.95	0.04	0.82	0.06	0.68	0.02	0.96	0.05	0.82	0.07	0.71	0.02	0.96	0.04	0.80	0.05	0.65
10	0.03	0.95	0.05	0.81	0.06	0.65	0.02	0.96	0.05	0.82	0.07	0.68	0.02	0.95	0.04	0.79	0.06	0.60
AVG	0.02	0.96	0.04	0.86	0.05	0.77	0.02	0.96	0.04	0.86	0.05	0.78	0.02	0.96	0.04	0.84	0.04	0.73
	Geneva						Rock Springs						Sunleaf Nursery					
3	0.02	0.98	0.08	0.80	0.05	0.62	0.02	0.88	0.03	0.73	0.03	0.62	0.05	0.95	0.20	0.60	0.33	0.61
4	0.07	0.76	0.08	0.79	0.05	0.40	0.04	0.78	0.04	0.49	0.04	0.45	0.11	0.73	0.14	0.59	0.89	-0.32
5	0.01	0.98	0.03	0.90	0.05	0.67	0.01	0.99	0.02	0.90	0.04	0.54	0.01	0.97	0.04	0.88	0.06	0.67
6	0.02	0.98	0.03	0.91	0.04	0.63	0.01	0.98	0.02	0.87	0.04	0.38	0.02	0.98	0.03	0.90	0.05	0.62
7	0.01	0.98	0.03	0.90	0.05	0.67	0.01	0.99	0.02	0.89	0.04	0.40	0.01	0.98	0.03	0.88	0.07	0.65
8	0.02	0.98	0.06	0.82	0.05	0.61	0.01	0.97	0.03	0.80	0.03	0.58	0.04	0.93	0.10	0.63	0.08	0.48
9	0.01	0.99	0.02	0.93	0.04	0.70	0.01	0.99	0.02	0.92	0.02	0.78	0.01	0.98	0.02	0.92	0.03	0.70
10	0.01	0.98	0.03	0.90	0.04	0.65	0.01	0.99	0.01	0.94	0.02	0.85	0.01	0.98	0.02	0.92	0.10	0.76
AVG	0.01	0.98	0.03	0.91	0.04	0.75	0.01	0.99	0.02	0.92	0.03	0.68	0.01	0.98	0.03	0.90	0.07	0.76

Table 3-5: Testing Results of the Iterative MLP Approach

MLP	Kelso 2						Governor Rd						Orchard					
#	10		20		50		10		20		50		10		20		50	
	RMSE	R	RMSE	R	RMSE	R	RMSE	R	RMSE	R	RMSE	R	RMSE	R	RMSE	R	RMSE	R
3	0.02	0.95	0.05	0.72	0.07	0.28	0.03	0.95	0.06	0.68	0.06	0.61	0.02	0.96	0.04	0.79	0.05	0.64
4	0.03	0.92	0.06	0.64	0.07	0.46	0.08	0.70	0.09	0.45	0.08	0.41	0.06	0.72	0.07	0.49	0.07	0.41
5	0.02	0.96	0.04	0.87	0.06	0.70	0.02	0.96	0.04	0.86	0.06	0.64	0.02	0.96	0.03	0.85	0.04	0.71
6	0.02	0.96	0.04	0.87	0.06	0.66	0.03	0.96	0.05	0.87	0.06	0.51	0.02	0.96	0.04	0.85	0.05	0.70
7	0.02	0.96	0.04	0.87	0.06	0.72	0.02	0.96	0.04	0.86	0.06	0.71	0.02	0.96	0.04	0.85	0.04	0.72
8	0.03	0.96	0.05	0.81	0.07	0.51	0.02	0.95	0.05	0.69	0.06	0.56	0.02	0.95	0.04	0.81	0.05	0.67
9	0.02	0.96	0.05	0.82	0.05	0.71	0.02	0.96	0.05	0.82	0.07	0.70	0.02	0.96	0.04	0.80	0.05	0.64
10	0.03	0.95	0.05	0.81	0.05	0.69	0.02	0.96	0.05	0.81	0.07	0.68	0.02	0.95	0.04	0.78	0.06	0.60
AVG	0.02	0.96	0.04	0.83	0.06	0.73	0.02	0.96	0.04	0.83	0.05	0.75	0.02	0.95	0.04	0.82	0.05	0.70
	Geneva						Rock Springs						Sunleaf Nursery					
3	0.02	0.98	0.06	0.71	0.05	0.46	0.02	0.91	0.04	0.49	0.04	0.61	0.05	0.96	0.08	0.42	1.29	0.37
4	0.06	0.81	0.06	0.66	0.05	0.42	0.06	0.35	0.03	0.81	0.04	0.65	0.12	0.76	0.19	-0.13	0.49	-0.31
5	0.01	0.98	0.03	0.92	0.04	0.69	0.01	0.98	0.02	0.95	0.03	0.67	0.01	0.97	0.04	0.91	0.06	0.69
6	0.02	0.98	0.03	0.92	0.03	0.66	0.01	0.99	0.02	0.94	0.03	0.70	0.02	0.98	0.02	0.92	0.04	0.64
7	0.01	0.98	0.03	0.92	0.04	0.71	0.01	0.98	0.02	0.94	0.03	0.73	0.02	0.96	0.04	0.90	0.06	0.72
8	0.02	0.97	0.05	0.79	0.05	0.54	0.01	0.97	0.02	0.87	0.04	0.66	0.04	0.92	0.07	0.64	0.11	0.14
9	0.01	0.99	0.02	0.92	0.04	0.67	0.01	0.99	0.01	0.95	0.02	0.88	0.01	0.98	0.02	0.93	0.03	0.74
10	0.01	0.98	0.03	0.91	0.04	0.58	0.01	0.99	0.01	0.96	0.02	0.90	0.01	0.98	0.03	0.93	0.07	0.67
AVG	0.01	0.98	0.03	0.90	0.03	0.65	0.02	0.97	0.02	0.93	0.03	0.79	0.01	0.96	0.04	0.82	0.20	0.65

Mesonet and SCAN profiles. Table 3-6 additionally provides the relative bias (RB) of the ensemble mean MLP prediction. Since, the MLP was trained using HYDRUS derived soil profiles, the statistics presented in Tables 3-6 are computed using HYDRUS-EnKF as the "truth value" rather than actual soil moisture observations. This procedure results in the statistical summary representing the MLP's ability to fit soil moisture and ignores potential errors in HYDRUS and the assumptions made in generating the training data. A qualitative comparison to *in situ* observations is made in Fig 3-6.

Table 3-6: Relative Bias of the Ensemble MLP Mean Soil Moisture

	Direct MLP Approach			Iterative MLP Approach		
Depth	10cm	20 cm	50 cm	10 cm	20 cm	50 cm
K2	-0.02	0.02	0.05	-0.02	0.01	0.03
GR	-0.02	-0.02	-0.04	-0.02	-0.004	-0.06
OR	-0.02	-0.02	-0.004	-0.02	-0.03	-0.02
GN	-0.02	0.08	0.05	-0.02	0.08	0.05
RS	-0.02	0.01	0.05	-0.03	0.01	0.04
SL	-0.01	0.15	0.45	0.02	0.23	-19.25

With few exceptions, as the prediction depth increased so did the RMSE and there was an associated decrease in the Pearson correlation coefficient, particularly at the 50 cm depth. Similarly, the relative bias of the ensemble mean in Table 3-6 shows the largest and least consistent bias at the 50 cm depth, whereas the 10 cm depth had a small and relatively consistent bias across all sites. This may be the result of several significant assumptions in the model formulation including the assumption that surface soil moisture observations provide sufficient information for the estimation of soil moisture in deeper soil layers. The latter assumption was lessened by modeling the root-zone soil moisture using an iterative approach. When the iterative approach was applied with the HYDRUS-EnKF soil moisture data and no sub-surface measurement was assumed, the compounding model errors resulted in less predictive skill than was found when only using surface measurements as model input. Therefore, only the results of ANN prediction using surface soil moisture as an input will be presented in further detail.

Figure 3-6 gives a time series representation of the mean neural network output at a) Governor Rd and b) Geneva at each depth considered. These two sites were selected as they have different soils and landscapes (see Table 3-1), therefore highlighting the flexibility of the data-driven approach. At the 10 and 20 cm depths the neural network estimate of soil moisture was

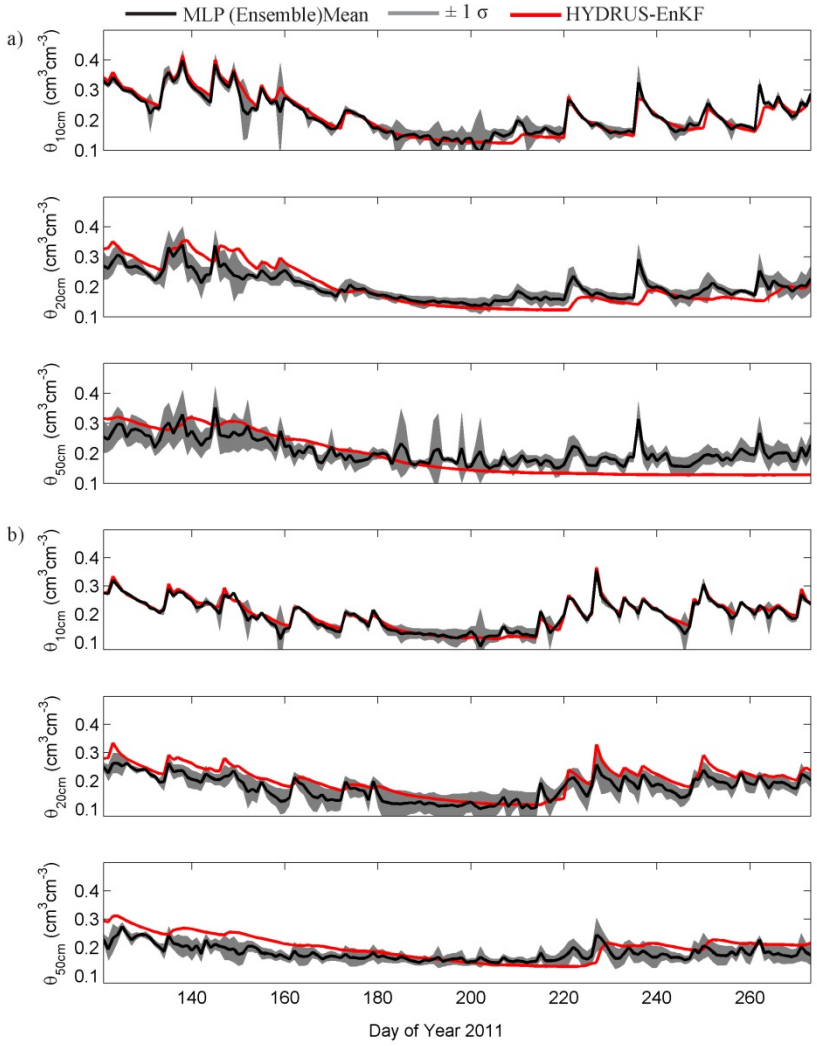

Figure 3-6: Time series of ensemble MLP predictions of soil moisture at (a) GR and (b) GN using MLP with direct surface soil moisture input set.

comparable to that of the HYDRUS-EnKF derived 'truth' value. At a depth of 50 cm, the MLP estimate tended toward values centered about the mean value and did not fully capture the muted seasonality of the soil moisture at this depth. The ensemble representation in Fig 3-6 was derived from an ensemble of the bagged network outputs from Table 3-4. Most MLP configurations capture the soil moisture dynamics in a similar manner, although some configurations were prone to producing erroneous values. Therefore, care must be taken not to rely on a single network configuration as an ensemble of complimentary neural network architectures gives higher skill to the overall model and mitigates the impact of erroneous behaviour in some neural networks. The MLP derived root-zone soil moisture is a good fit to HYDRUS data, but is subject to the error of the HYDRUS training set and has similar errors with respect to observed *in situ* soil moisture. Comparison between the plot of observed soil moisture and the MLP simulation at GR shows the MLP has a dry bias similar to HYDRUS during the early part of the study period, but does capture many of the relative temporal dynamics in terms of wetting and drying periods. A possible explanation for this poor model fit in the early study period, with a better fit at the end

may be the assumption that LAI is constant, where an over-estimated LAI in the early growing season would result in greater evapotranspiration and an erroneously dry simulation. The importance of a variable LAI value should be evaluated in future studies. At GN where the HYDRUS simulation corresponded more closely to observed soil moisture, the MLP also had better performance. Errors in HYDRUS, and therefore the MLP training set, may derive from assumptions of profile homogeneity [Lu et al. 2011], the presumed root-zone distribution, the root-water uptake model, not accounting for air pores preventing saturation in HYDRUS [Ma et al. 2010], the presence of mulch [Calvet and Noilhan 2000] which was ignored in HYDRUS.

3.6 Discussion and Conclusions

Ensemble MLP models were used to estimate root-zone soil moisture at a depth of 10, 20 and 50 cm using observations of soil moisture at a depth of 5 cm. It was shown that the MLP models were able to account for non-linear soil moisture dynamics and interactions with surface meteorological forcing. Several MLP configurations were trained with a bagging approach using data derived from HYDRUS-1D. The ANN models were independently tested against synthetic soil moisture profiles generated by assimilating *in situ* soil moisture observations from three McMaster Mesonet and three USDA SCAN sites in HYDRUS-1D. The model test results were also evaluated against actual soil moisture observations. The proposed ensemble ANN models were effective at modeling soil moisture at the 10 and 20 cm depths at all sites, and the spatial distribution of the validation sites indicates the potential of the ANN based models for soil moisture estimation within a geographic region. These results suggest that ANNs may be a suitable alternative to data assimilation for estimating root-zone soil moisture, with the benefit of requiring significantly less computational resources once trained. The test results also suggest the feasibility of developing a more general ANN based root-zone soil moisture estimation model. However, this does require further research.Elshorbagy and Parasuraman [2008] demonstrated that without the benefit of soil moisture information the ability of a neural network to predict soil moisture in the root-zone is limited. Therefore, it was expected that much of the information used by the MLPs to predict root-zone soil moisture was derived from the surface soil moisture observation and the soil texture as seen in Figs 3-4 and 3-5. The sensitivity to potential evapotranspiration as an input can also be understood due to both its integration of other forcing

variables and the physical importance of evapotranspiration in removing moisture from the soil. This may also indicate why the MLP model performance decreased with depth, where the potential influence of water table interactions, reduced sensitivity to surface forcing and an oversimplified assumption about soil profile characteristics hampered MLP performance. These findings suggest that the ANN methods as presented are most effective for predicting soil moisture in shallow depths (up to 20 cm), whereas methods such as that presented by Wagner et al. [1999] may be more suitable at greater depths.

Previous studies found the use of higher order neural networks produced better results for soil moisture estimation than simple neural networks [Elshorbagy and Parasuraman 2008]. In this application, the LM-MLP(B,[15,15],α0) produced poorer estimates of root-zone soil moisture than simple single layer MLPs. The smaller weight matrices of a single layer network also benefit from less computational cost for training. In terms of network architecture and training, the greatest improvement in performance resulted from regularization of the training cost function either manually or using Bayesian Regularization [MacKay 1992]. Manual selection of the regularization parameter requires a trial and error approach as only certain ratios were found to provide improvements in the network output, while others did not and therefore were not presented. Since individual networks are randomly initialized and are trained independently, the bagged ensemble approach [Brieman 1996], applied herein, provided greater stability in the final model output and minimized erroneous behaviour which was present with some individual networks. To improve the stability and performance of the final model output, the use of an ensemble of MLPs with different architectures and training algorithms is recommended [Abrahart et al. 2012] in addition to the bagging ensemble. The benefits, such as the stability of BFGS training, the optimization of BR and added non-linearity of a tangent sigmoid (‘tansig’) output transfer function, were considered important in this study for achieving the desired level of performance.

Following Das and Mohanty [2006], the soil profile being modeled was considered vertically homogenous and should only be considered valid if the point (grid cell) being modeled is also considered spatially (horizontally) homogenous. Applications of the proposed ensemble ANN method, such as satellite root-zone soil moisture retrieval, additionally require the consideration of the impact of scale and heterogeneity which are not addressed herein.

Since data-driven methods integrate the inter-relationships of training data, they also have the advantage of creating a model which makes few simplifying assumptions about underlying processes. The ANN model generated was based solely on interactions which were extracted from the data set to produce a statistical model of a process. Therefore, the validity of an ANN model is limited to the conditions for which the model is generated. A novel aspect to this research is the demonstration that a simple MLP model can be considered valid over a relatively broad range of physical conditions. In contrast, a physical model with data assimilation is potentially more flexible when conditions change, but subject to the correct mathematical description and selection of processes and parameters involved.

In this research, the limitations of process representation of the physical model, HYDRUS, was also applied to the MLP, since HYDRUS simulations were used as training data. Since many areas of the world do not have extensive soil moisture networks to provide a suitable training set, HYDRUS was evaluated in lieu of observations in an attempt to generate a physically realistic training set. Due to the simplifying assumptions used when generating the generalized HYDRUS training data for ANN training, the soil moisture dynamics at some sites such as K2 and GR were not well captured when compared to the *in situ* observations. Also, the model was tested in a limited geographic region and future analysis will seek to determine the geographic extent of a similar MLP approach.

As was shown by Koster et al. [2009], soil moisture is a model dependent quantity, where the transfer of soil moisture values between models is problematic. This research has demonstrated that deriving ANN training data from HYDRUS is effective for the simulation of soil moisture with characteristics similar to that of HYDRUS, but this data was not always representative of *in situ* observations. It is therefore important, that the network training set be composed of data which is representative of the application for which the networks are proposed to be used, or that a scaling function is used for mapping soil moisture between different applications [Koster et al. 2009]. Further research will be required to determine if this limitation can be overcome in the MLP approach. Potential solutions could involve training the MLP to predict an index such as the soil water index [Wagner et al. 1999] or saturation ratio rather than volumetric soil moisture. To apply the MLP technique to *in situ* soil moisture, training could be carried out using a large set of *in situ* observations such as the SCAN network [Schaefer et al. 2007] or data from the International Soil Moisture Network [Dorigo et al. 2011].

3.7 References

Abrahart, R. J., F. Anctil, P. Coulibaly, C. D. Dawson, N. J. Mount, L. M. See, A. Y. Shamseldin, D. P. Solomatine, E. Toth, and R. L. Wilby, R.L. (2012), Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. Prog. Phys. Geog., 36(4): 480–513, doi: 10.1177/0309133312444943.

Albergel, C., C. Rudiger, T. Pellarin, J.-C. Calvet, N. Fritz, F. Froissard, D. Suquia, A. Petitpa, B. Piguet, and E. Martin (2008), From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations, Hydrol. Earth Syst. Sci. 12(6) 1323-1337, doi:10.5194/hess-12-1323-2008.

Behkhaled, A., B. Remini, and M. Mhaiguene (2004) Influence of antecedent precipitation index on the hydrograph shape, in Hydrology: science and practice for the 21st century, Volume 1. Proceedings of the British Hydrological Society International Conference, 81-87, Imperial College, London.

Bishop, C. M. (1995), Neural Network for Pattern Recognition, Clarendon, Oxford, U. K.

Breiman, L. (1996), Bagging predictors, Mach. Learn., 24(2), 123-140.
Calvet, J.-C., and J. Noilhan (2000), From near-surface to root-zone soil moisture using year-round data, J. Hydrometeorol. 1(5), 393-411, doi: 10.1175/1525-7541(2000)001.

Clark, M. P., D. E. Rupp, R. A. Woods,X. Zheng, R. P. Ibbitta, A. G. Slater, J. Schmidta, M. J. Uddstroma (2008), Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model, Adv. Water Resour. 31: 1309-1324, doi:10.1016/j.adv.watres.2008.06.005.

Das, N. N., and B. P. Mohanty (2006), Root zone soil moisture assessment using remote sensing and vadose zone modeling, Vadose Zone J. 5, 296-309, doi:10.1236.vzj2005.0033.

Dorigo, W. A., W. Wagner, R. Hohensinn, S. Hahn, C. Paulik, A. Xaver, A. Gruber, M. Drusch, S. Mecklenburg, P. van Oevelen, A. Robock, and T. Jackson (2011), The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, Hydrol. Earth Syst. Sci., 15, 1675-1698, doi:10.5194/hess-15-1675-2011.

Dumedah G., and P. Coulibaly (2012), Evolutionary assimilation of streamflow in distributed hydrologic modeling using in-situ soil moisture data. Adv. Water Resour., 53: 231–241, doi:10.1016/j.advwatres.2012.07.012.

Elshorbagy, A., and K. Parasuraman (2008), On the relevance of using artificial neural networks for estimating soil moisture content, J. Hydrol. 362(1), 1-18, doi: 10.1016/j.jhydrol.2008.08.012.

Entekhabi, D., E. G. Njoku, P. E. O'Neill, K. H. Kellogg, W. T. Crow, W. N. Edelstein, J. K. Entin, S. D. Goodman, T. J. Jackson, J. Johnson, J. Kimball, J. R. Piepmeier, R. D. Koster, N. Martin, K. C. McDonald, M. Moghaddam, S. Moran, R. Reichle, J. C. Shi, M. W. Spencer, S. W. Thurman, L. Tsang, and J. Van Zyl (2010), The soil moisture active passive (SMAP) mission, Proc. IEEE, 98(5), 704-716, doi:10.1109/JPROC.2010.2043918.

Feddes, R. A., P. J. Kowalik, and H. Zaradny (1978), Simulation of Field Water Use and Crop Yield, John Wiley & Sons, New York, NY.

Haykin, S. (1999). Neural Networks: A Comprehensive Foundation. Prentice Hall, Upper Saddle River, NJ.

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

Homaee M., R. A. Feddes, and C. Dirksen (2002), Simulation of root water uptake II. Nonuniform transient water stress using different reduction functions. Agric. Water Manage. 57(2), 111–126, doi:10.1016/S0378-3774(02)00071-9.

Jain, A., and S. Srinivasulu (2004), Development of effective and efficient rainfall-runoff models using integration of deterministic, real-coded genetic algorithms and artificial neural network techniques, Water Resour. Res., 40, W04302, doi:10.1029/2003WR002355.

Kerr, Y., P. Waldteufel, J.-P. Wigneron, S. Delwart, F. Cabot, J. Boutin, M.-J. Escorihuela, J. Font, N. Reul, C. Gruhier, S. E. Juglea, M. R. Drinkwater, A. Hahne, M. Martin-Neira, and S. Mecklenburg (2010) The SMOS mission: New tool for monitoring key elements of the global water cycle, Proc. IEEE, 98(5), 666-687, doi:10.1109/JPROC.2010.2043032.

Khan, M. S., and P. Coulibaly (2006), Bayesian neural network for rainfall-runoff modeling, Water Resour. Res., 42, W07409, doi:10.1029/2005WR003971.

Kornelsen, K., and P. Coulibaly (2013a), Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications, J. Hydrol. 476(1), 460-489, doi: 10.1016/j.jhydrol.2012.10.044

Kornelsen, K., and P. Coulibaly (2013b), McMaster Mesonet soil moisture dataset: description and spatio-temporal variability analysis, Hydrol. Earth Syst. Sci., 17(1), 1-18, doi:10.5194/hess-17-1-2013.

Kornelsen, K., and P. Coulibaly (2014), Comparison of interpolation, statistical and data driven methods for imputation of missing values in a distributed soil moisture dataset, J. Hydrol. Eng., 19(1) doi:10.1061/(ASCE)HE.1943-5584.0000767.

Koster, R. D., Z. Guo, R. Yang, P. A. Dirmeyer, K. Mitchell, and M. J. Puma (2009), On the nature of soil moisture in land surface models, J. Climate, 22(16), 4322-4335, doi:10.1175/2009JCLI2832.1.

Kumar, S. V., R. Reichle, R. T. Koster, W. T. Crow, C. D. Peters-Lidard (2010), Role of subsurface physics in the assimilation of surface soil moisture observations, J. Hydrometerol. 10(6), 1534-1547, doi:10.1175/2009JHM1134.1.

Leij, F. J., W. J. Alves, M. Th. van Genuchten, and J. R. Williams (1996), Unsaturated Soil Hydraulic Database, UNSODA 1.0 User's Manual. Report EPA/600/R-96/095, US Environmental Protection Agency, Ada, Oklahoma.

Lu, H., X. Li, Z. Yu, R. Horton, Y. Zhu, Z. Hao, and L. Xiang (2010), Using a H∞ filter assimilation procedure to estimate root zone soil water content, Hydrol. Process. 24(25), 3648-3660, doi:10.1002/hyp.7778.

Lu, H., Z. Yu, Y. Zhu, S. Drake, Z. Hao, E. A. Sudicky (2011), Dual state-parameter estimation of root zone soil moisture by optimal parameter estimation and extended Kalman filter data assimilation, Adv. Water Resour. 34(3) 395-406, doi:10.1016/j.advwatres.2010.12.005.

Ma, Y., S. Feng, D. Su, G. Gao, Z. Huo (2010), Modeling water infiltration in a large layered soil column with a modified Green-Ampt model and HYDRUS-1D, Comput. Electron. Agr. 71S, S40-S47, doi:10.1016/j.compag.2009.07.006.

MacKay, D. J. C. (1992), Bayesian interpolation, Neural Comput., 4(3), 415-557, doi:10.1162/neco.1992.4.3.415.

Moradkhani H., S. Sorooshian, H. Gupta, P. Houser (2005) Dual state–parameter estimation of hydrological models using ensemble Kalman filter. Adv. Water Resour. 28(1), 135-147, doi: doi:10.1016/j.advwatres.2004.09.002.

Nouvellon, Y., A. Begue, M. S. Moran, D. L. Seen, S. Rambal, D. Luquet, G. Chehbouni, and Y. Inoue (2000), PAR extinction in shortgrass ecosystems: effects of clumping, sky conditions and soil albedo, Agr. Forest. Meteorol. 105(1), 21-41, doi:10.1016/S0168-1923(00)00195-4.

Reichle, R., W. T. Crow, R. T. Koster, J. Kimball, and G. De Lannoy (2012), Algorithm Theoretical Basis Document (ATBD) SMAP Level 4 Surface and Root Zone Soil Moisture (L4_SM) Data Product, NASA Goddard Space Flight Center, Greenbelt, MD, USA.

Ritchie, J. T. (1972), Model for predicting evaporation from a row crop with incomplete cover, Water Resour. Res., 8(5), 1204-1213.

Rosenbaum, U., H. R. Bogena, M. Herbst, J. A. Huisman, T. J. Peterson, A. Weuthen, A. W. Western, and H. Vereecken (2012), Seasonal and event dynamics of spatial soil moisture patterns at the small catchment scale, Water Resour. Res., 48, W10544, doi:10.1029/2011WR011518.

Schaap, M. G., F. J. Leij, M. Th. Van Genuchten (2001), ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions, J. Hydrol. 251(1) 163-176.

Schaefer, G. L., M. H. Cosh, and T. J. Jackson (2007), The USDA natural resources conservation service soil climate analysis network (SCAN), J. Atmos. Ocean. Tech. 24(12), 2073-2077, doi:10.1175/2007JTECHA930.1.

Scurlock, J. M. O., G. P. Asner, and S. T. Gower (2001), Worldwide Historical Estimates and Bibliography of Leaf Area Index, 1932-2000. ORNL Technical Memorandum TM-2001/268, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Mirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010), Investigating soil moisture-climate interactions in a changing climate: A review, Earth-Sci. Rev., 99(1) 125-161, doi:10.1016/j.earscirev.2010.02.004.

Singh, V. P. (2010), Entropy theory for movement of moisture in soils, Water Resour. Res., 46, W03516, doi:10.1029/2009WR008288.

Srivastava, S. K., N. Yograjan, V. Jayaraman, P. P. Nageswara Rao, and M. G. Chandrasekhar (1997), On the relationship between ERS-1 SAR/Backscatter and surface/sub-surface soil moisture variations in vertisols, Acta Astronaut., 40(10), 693-699.

Van Genuchten, M. Th. (1980), A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Am. J. 44(5), 892-898, doi: 10.2136/sssaj1980.03615995004400050002x.

van Genuchten, M. Th. (1987), A numerical model for water and solute movement in and below the root zone. Research Report No 121, U.S. Salinity laboratory, USDA, ARS, Riverside, California.

Walker, J. P., G. R. Willgoose, and J. D. Kalma (2001a), One-dimensional soil moisture retrieval by assimilation of near-surface observations: a comparison of retrieval algorithms. Adv. Water Resour. 24(6), 631-650, doi: 10.1016/S0309-1708(00)00043-9.

Walker, J. P., G. R. Willgoose, and J. D. Kalma (2001b), One-dimensional soil moisture profile retrieval by assimilation of near-surface measurements: A simplified soil moisture model and field application. J. Hydrometeor. 2, 356-373, doi: 10.1175/1525-7541(2001)002.

Wagner, W., G. Lemoine, and H. Rott (1999), A method for estimating soil moisture from ERS scatterometer and soil data. Remote Sens. Environ. 70(2), 191-207, doi:10.1016/S0034-4257(99)00036-X.

Wang, W., P. Jones, and D. Partridge (2000), Assessing the impact of input features in a feedforward neural network, Neural Comput. & Applic. 9(2), 101-112, doi: 10.1007/PL00009895.

Chapter 4: Potential of Bias Correction for Downscaling Satellite Observations

Summary of Paper 3: Kornelsen, K.C., Cosh, M.H. and Coulibaly, P. (submitted) Potential of Bias Correction for Downscaling Satellite Observations, Journal of Geophysical Research: Atmosphere, Manuscript Number: 2014JD022683

Summary:

This research evaluated a novel application of existing bias correction techniques for downscaling watershed average soil moisture and brightness temperature (TB). Using data from two well instrumented USDA watersheds and the Community Microwave Emission Model (CMEM), the watershed was studied for temporal persistence of soil moisture and TB. The mean squared difference (MSD) between the watershed average soil moisture or TB and that of individual sites was decomposed to partition the role of bias, variance and conditional bias on the MSD. Using a linear rescaling bias correction technique, the soil moisture and TB at both polarizations and multiple incidence angles were downscaled from the watershed scale to the point scale.

The key findings of this research were:

- Both soil moisture and TB are temporally stable with increasing (decreasing) variability with incidence angle for H (V) polarization.
- Bias was the primary contributor to the MSD between the watershed average values and those at the local scale. Differences that resulted from timing, variance and conditional bias were relatively consistent at all incidence angles for TB, whereas bias was incidence angle dependent.
- The correction of bias had little impact on correlation, but resulted in significant improvement in RMSD between the watershed scale soil moisture or TB and that of the local sites.
- The correction of biases can be considered a reliable method for downscaling watershed scale soil moisture and TB for hydrological applications.

4.1 Abstract

Passive microwave satellites such as SMOS (Soil Moisture and Ocean Salinity) or SMAP (Soil Moisture Active Passive) observe brightness temperature (TB) and retrieve soil moisture at a spatial resolution greater than most hydrological processes. Bias correction is proposed as a simple method to disaggregate soil moisture to a scale more appropriate for hydrological applications. Temporal stability of soil moisture and TB was demonstrated at the Little Washita and Little River Experimental Watersheds using *in situ* observations and the Community Microwave Emissions Model. Decomposition of the mean squared difference (MSD) between the watershed average soil moisture and TB showed bias was a major contributor to differences between watershed average and local scale soil moisture and TB, particularly at sites with high MSD. The mean RMSD between watershed average and local soil moisture was 0.04 m^3m^{-3} and 0.06 m^3m^{-3} at Little River and Little Washita respectively. Following a simple bias correction the RMSD was reduced to 0.03 m^3m^{-3} at both sites. Considering multiple incidence angles at both horizontal and vertical polarization, bias correction of watershed average TB_V reduced the RMSD by approximately 75% and 45% and TB_H RMSD by 68% and 36% for Little River and Little Washita respectively at all incidence angles. Therefore, at sub-satellite grid scale, bias correction can be considered a viable technique for downscaling passive microwave observations and soil moisture retrievals.

4.2 Introduction

Soil moisture is an important geophysical state variable that influences hydro-climatic processes at the Earth's surface [Seneviratne et al., 2010]. The assimilation of soil moisture information has been demonstrated to improve numerical weather prediction [Drusch, 2007], hydrological forecasting [Bronstert et al., 2012] and climate modelling [Seneviratne et al., 2013] amongst other applications. Monitoring of soil moisture for geophysical applications is best conducted with remote sensing observations due to the high spatial heterogeneity of soil moisture; typically at microwave wavelengths using either active microwave remote sensing [Baghdadi et al., 2007; Wagner et al., 1999] or passive radiometers [Bolton et al., 2010; Kerr et al., 2010; Entekhabi et al., 2010]. Observations from passive microwave systems such as the Soil Moisture and Ocean

Salinity (SMOS) [Kerr et al., 2010] or the Soil Moisture Active Passive (SMAP) [Entekhabi et al., 2010] missions have the advantage of shorter revisit times and higher radiometric accuracy, whereas higher resolution soil moisture retrievals from synthetic aperture radar (SAR) typically have long revisit periods making their use for operational applications impractical [Kornelsen and Coulibaly, 2013a]. A further consideration for operational data assimilation systems is the type of information to be assimilated. The relatively high latency in delivering a soil moisture retrieval product to end-users necessitates the assimilation of observations, such as brightness temperature (TB), and the use of observation operators such as the Community Microwave Emission Model (CMEM) [Drusch et al., 2009] for operational applications. For global climate modeling and numerical weather prediction, computational limits result in the spatial scale of passive satellite observations being commensurate with model grid resolution. However, applications such as flood forecasting and numerical weather prediction, particularly for convective storms, are expected to benefit when a model is run at a spatial scale that is higher than the soil moisture information that can be produced in short revisit times by SMOS, the SMAP radiometer or the MetOp Active Scatterometer (ASCAT) [Wagner et al., 1999; Bartalis et al., 2008]. Dumedah et al. [2013] demonstrated that the uncertainty involved in oversampling SMOS TB from its native 43 km resolution to the 15 km discrete global grid the data are distributed on, was comparable to the 2 K uncertainty of the instrument itself. However, even this spatial resolution is coarser than many hydrological processes which occur at scales of less than 1 km.

The discrepancy in scale between some applications and satellite observations requires that the satellite product be scaled down to a resolution appropriate for hydrological applications. Several different approaches to downscale soil moisture information have been presented. These methods can be categorized as those based on the relationship between soil moisture and land surface temperature/evaporation using optical and thermal remote sensing; those based on microwave remote sensing; and those based on statistical relationships which we consider to include data assimilation.

The method of Merlin et al. [2010; 2012] is based on the principal that there is a physical relationship between soil temperature and soil moisture which is described by a semi-empirical soil evaporative efficiency model [Merlin et al., 2012]. A similar approach is the triangle-based methods first developed by Chauhan et al. [2003], which are based on the triangular or

trapezoidal shaped relationship between land surface temperature, vegetation content and soil moisture [Chauhan et al., 2003; Kim and Hogue, 2012]. These methods rely on concurrent optical and thermal observations to downscale soil moisture and suffer from limitations of optical remote sensing, such as the impact of cloud cover. They have an advantage in that they do not require any ground or model based information.

The downscaling approach of the SMAP mission is to couple active and passive microwave observations [Entekhabi et al., 2010]. SMAP is unique in that the satellite platform carries both active and passive microwave instruments which share the same antenna and make concurrent observations. This allows higher resolution SAR backscatter to be used to downscale radiometer TB [Das et al., 2014]. Complimenting passive observations with other currently available SAR observations is impractical since the long revisit period results in limited sampling [Kornelsen and Coulibaly, 2013a].

Statistical downscaling methods represent a broad spectrum of approaches, which typically involve the calibration of a statistical relationship with high resolution proxies or soil moisture information. Mascaro et al. [2011] and Parada and Liang [2003] downscaled soil moisture by describing the multi-fractal relationship between soil moisture at different resolutions. Parada and Liang [2003] used ARMA models and wavelet techniques to downscale soil moisture to progressively finer scales. Mascaro et al. [2011] calibrate a downscaling model that uses predictors such as topography and soil texture to create a multi-fractal downscaling cascade to predict soil moisture. Wilson et al. [2005] applied a simpler approach where topographic indices and topographic information are used to spatially distribute soil moisture. It was found that the use of surface information provided some skill in distributing soil moisture, whereas further skill was achieved by including the mean residual pattern [Wilson et al., 2005]. This result is indicative of the soil moisture property of temporal persistence [Vachaud et al., 1985], which states that the relative difference of soil moisture at a particular location in a watershed is stable with respect to the mean watershed state. Temporal persistence has been demonstrated at the field, watershed and regional scale in different environments all over the globe for soil moisture [i.e. Cosh et al., 2006; Kornelsen and Coulibaly, 2013b; Brocca et al., 2010; Cosh et al., 2004; Vanderlinden et al., 2012]. Parada and Liang [2008] compared soil moisture assimilation using the multi-scale Kalman filter and the assimilation of coarse resolution soil moisture with temporal ratios and determined that the addition of temporal

persistence recovered much of the information lost by the coarse resolution soil moisture. Loew and Mauser [2008] demonstrated that when *a priori* information was available, the property of temporal persistence allowed soil moisture to be effectively downscaled using a simple linear relationship.

The primary purpose of many of the previous studies was to provide state updates to a data assimilation system to improve hydro-climatic forecasts. Most variants of land surface data assimilation schemes use derivations of the Kalman Filter for sequential data assimilation, which requires that both the model and satellite observations share the same state space. Therefore, it is common practice that a bias correction be included as part of the observation operator for data assimilation [Drusch et al., 2009; Reichle and Koster, 2004], and is necessary for both direct assimilation of soil moisture [Reichle and Koster, 2004] and indirect assimilation of TB [Drusch et al., 2009]. This paper aims to demonstrate that, for the purposes of data assimilation; bias correction techniques can effectively 'disaggregate' both microwave TB and soil moisture. The methodology will be demonstrated using two well instrumented watersheds operated by the US Department of Agriculture (USDA) [Cosh et al., 2006; Bosch et al., 2007]. Temporal persistence of soil moisture has already been demonstrated at these watersheds [Cosh et al., 2008; Cosh et al., 2006; Bosch et al., 2006; Heathman et al., 2009] and will be extended in this research by determining if temporal stability is also demonstrated by TB. Brightness temperature can also be expected to display temporal persistence since it is strongly influenced by soil moisture particularly at L-Band [Drusch et al., 2009] and variations due to changes in vegetation density occur slowly compared to soil moisture. To the authors knowledge this is the first attempt to analyze TB for temporal persistence.

Since high resolution TB measurements from field campaigns have finite temporal sampling, CMEM is used to simulate TB rather than using temporally limited observations. The efficacy of the proposed approach will also be demonstrated through the decomposition of the mean squared difference (MSD) into its principal components [Murphy, 1988] and downscaling watershed average soil moisture and TB using a simple bias correction method.

4.3 Study Areas and Data

This study will examine two dense soil moisture networks from experimental watersheds operated by the USDA Agricultural Research Service; Little Washita (LW), OK and Little River (LR), GA shown in Fig. 4-1. Both watersheds are of similar order of magnitude to a SMOS DGG or ASCAT grid cell. These networks are located in agricultural areas and have previously been used for the validation of AMSR-E soil moisture [Jackson et al., 2010] and SMOS soil moisture [Jackson et al., 2012]. Previous studies have demonstrated the representativeness and accuracy of these sites and analyzed the soil moisture scaling properties [Cosh et al., 2008]. Following Jackson et al. [2010; 2012] the watershed average was calculated based on a weighted mean of reliable stations using Thiessen polygons. It was assumed that the soil moisture measured at each station was representative of the entire Thiessen polygon, and was considered adequate for the purposes of demonstrating bias correction as a downscaling technique. All data were analyzed for the period of 1 April 2010 – 31 September 2010. This time period approximately coincides with the spring/summer seasons in the study areas and decreases the likelihood of errors induced by frost conditions.

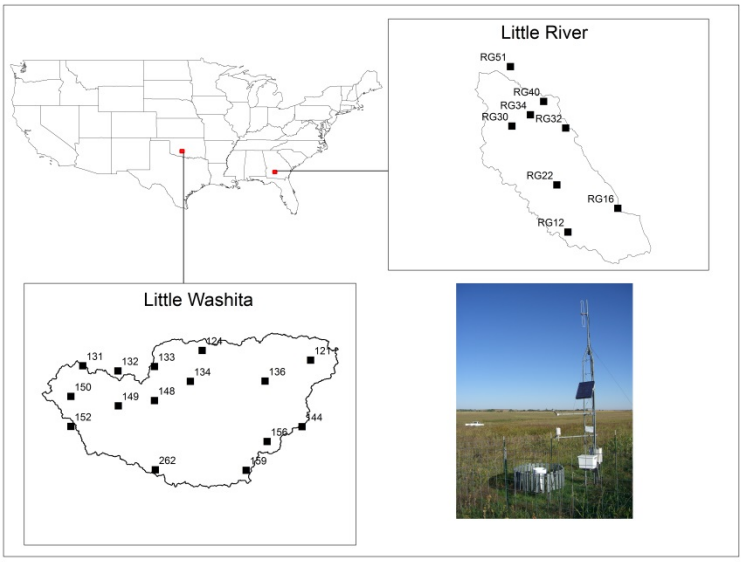

Figure 4-1: The Little Washita and Little River soil moisture networks. The filled boxes show only the stations considered in this study. The lower right image shows a typical station located at Little Washita.

Little Washita is located in southwestern Oklahoma, where the monitored study area covers 610 km^2 [Cosh et al., 2006]. The topography is gently rolling and its dominant land uses are range, pasture and cropland. High vegetation is sparse and spread across a few woody savannas. The soil textures range from sands to clays, although at most study sites the soils were sandy loam to silt loam. Soil moisture observations are available at 20 sites in LW of which 15

are used in this study, and 5 sites are considered to be less reliable [Jackson et al., 2010]. More information on the LW watershed can be found in Cosh et al. [2006], Jackson et al. [2010], Heathman et al. [2009] and references therein.

The Little River Experimental Watershed is a 334 km^2 agricultural watershed in southern Georgia [Bosch et al., 2007]. Land cover in the watershed is dominated by agricultural land uses and mixed agriculture with natural vegetation [Broxton et al., 2014]. Agriculture consists primarily of row crops and natural vegetation throughout the watershed is mainly upland forests [Sullivan et al., 2007]. The soil texture was sampled during the SMEX03 field campaign and was found to be loamy sand at each study site [Cosh and Birch, 2008]. Of the 19 soil moisture profiles available at LR during the study period, only 8 were considered to be reliable enough for inclusion in this analysis [Jackson et al., 2010].

Besides soil moisture, ancillary data at the USDA experimental watersheds was required as input to CMEM. Leaf area index (LAI) values were retrieved from the MODIS MOD15A3 1 km product which was extracted at each sampling location and smoothed following Fang et al. [2008]. Land cover was derived from a MODIS land cover climatology (2001-2010) produced by Broxton et al. [2014]. The fraction of each land cover type was extracted for every Thiessen polygon and converted to the ECOCLIMAP land categories required by CMEM. Soil texture for the LW and LR watersheds were collected during the SMEX03 field experiments [Cosh and Birch, 2008]. A single 2 m air temperature observation from a Soil Climate Analysis Network (SCAN) [Schaefer et al., 2007] station in each watershed was considered to be representative of the entire watershed. Given the limited size of each study area this assumption was not considered limiting. Using these data as input to CMEM, TB was simulated for each individual station. Watershed average TB was calculated as a weighted mean using the Thiessen polygon method.

4.4 Methods

4.4.1 Community Microwave Emissions Modeling Platform

The Community Microwave Emissions Modeling platform (CMEM) was introduced by Drusch et al. [2009] and de Rosnay et al. [2009] as the European Centre for Medium-Range Weather Forecasting (ECMWF) contribution to the SMOS mission. In this study, CMEM was used to

model microwave emissions in L-Band (1.4 GHz) using the Microwave Emission of the Biosphere (LMEB) model [Wigneron et al., 2007]. Smooth surface emissivity was determined using the Fresnel law and rough surface emissivity was calculated following Wigneron et al. [2007]. The atmospheric opacity model of Pellarin et al. [2003] was chosen and effective temperature was determined with the parameterization of Holmes et al. [2006] following the recommendation of de Rosnay and Wigneron [2006]. To convert the dielectric constant to soil moisture, the model of Mironov et al. [2004] was chosen. CMEM was used to simulated brightness temperature at incidence angles of 20°, 30°, 40°, 50° and 60° for both polarizations to cover the range of incidence angles available from SMOS [Kerr et al., 2010] and the SMAP nominal 40° incidence angle [Entekhabi et al., 2010].

4.4.2 Analysis Methods

The concept of temporal persistence [Vachaud et al., 1985] is closely related to that of bias, suggesting there is a time invariant relationship, or bias, between the soil moisture at a particular location and that of the watershed mean value. Most applications of temporal persistence [i.e. Jacobs et al., 2004; Cosh et al., 2008; Brocca et al., 2010] propose finding a site or small set of sites that are representative of mean watershed conditions by identifying sites whose mean relative difference $\left(\overline{\delta}_{i,j}\right)$ is close to zero and for which there is small variance in the relative difference $\left(\sigma(\delta)^2_{i,j}\right)$. The relative difference is defined as the difference between the soil moisture at point *i* in watershed *j* at time *t* and that of the mean value of all the points in *j* at *t*

$$\overline{\delta}_{i,j} = \frac{1}{n}\sum_{i=1}^{n}\frac{\theta_{i,j,t} - \overline{\theta}_{j,t}}{\overline{\theta}_{j,t}} \tag{4-1}$$

$$\sigma(\delta)^2_{i,j} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{\theta_{i,j,t} - \overline{\theta}_{j,t}}{\overline{\theta}_{j,t}} - \overline{\delta}_{i,j}\right)^2 \tag{4-2}$$

where $\overline{\theta}_{j,t}$ is the watershed mean soil moisture/TB at *t*, and *n* is the number of temporal observations. Jacobs et al., [2004] propose that sites best representing the mean are those with the lowest root mean squared error of mean relative difference, which is a joint metric that incorporates the mean relative difference as an indication of bias and the variance of the relative difference as an indication of stability. It is defined as:

$$RMSE_{i,j} = \left(\overline{\delta}_{i,j}^2 + \sigma(\delta)_{i,j}^2\right)^{1/2} \tag{4-3}$$

The impact of bias when comparing a point location to the watershed average can also be determined by decomposing the mean squared difference into its constituent components [Murphy, 1988; Weglarczyk, 1998] as:

$$MSD = \sigma_{i,j}^2 \left(\left(1 - R^2\right) + \left(\frac{\sigma_j}{\sigma_{i,j}} - R \right)^2 + \frac{B^2}{\sigma_{i,j}^2} \right) \tag{4-4}$$

where σ is the standard deviation, the sole subscript j indicates the watershed average timeseries, R is the linear correlation coefficient and B is the bias between the observed values at i and the watershed mean values. This decomposition of MSD contains three additive non-dimensional terms. $IA = 1 - R^2$ is a measure of differences in the covariability between the observations at location i and the watershed mean value. The second term,

$$IB = \left(\frac{\sigma_j}{\sigma_{i,j}} - R \right)^2$$

is a nondimensional measure of the conditional bias, which is the covariability between the differences between the two datasets and the average watershed observations [Weglarczyk, 1998]. In model calibration problems, ideal conditions would result in the elimination of this term as model error would should not covary with model output. Applied to MSD, the covariance of the watershed average and observed values are expected and cannot be eliminated. The final term,

$$IC = \frac{B^2}{\sigma_{i,j}^2}$$

provides a measure of the impact of bias on the mean squared difference.

A simple bias correction method was chosen to demonstrate the efficacy of the proposed approach. Following Draper et al. [2009], the watershed mean values will be matched to the local grid/site climatology using:

$$\tilde{\theta}_{j(i)} = \mu_{i,j} + \frac{\sigma_{i,j}}{\sigma_j}\left(\theta_{i,j} - \mu_j\right) \tag{4-5}$$

where the notation $\tilde{\theta}_{j(i)}$ is the estimate created by bias correcting the averaged watershed values to match those of location *i*, and μ is the temporal mean. This bias correction method rescales the mean and standard deviation of the average watershed soil moisture/TB to that of the individual sites/grid cells. The performance of this method for bias correction was found to be comparable to more advanced methods, but is subject to appropriate estimates of the parameters (μ, σ) to be effective [Kornelsen and Coulibaly, manuscript in preparation, 2014]. Under conditions where a bias correction is appropriately applied, that is the bias is eliminated and standard deviations are matched, term *IB* of Eq. (4-4) is simplified to $(1-R)^2$ and term *IC* is eliminated. Therefore, simplifying and rearranging Eq. (4-4) produces an *a priori* estimate of the MSD following bias correction:

$$MSD \approx \sigma_{i,j}^2 (2-2R) \tag{4-6}$$

Under conditions where the variability of local observations and watershed mean values are similar, Eq. 4-6 provides an accurate estimate of MSD as the magnitude of values from their mean remains unaltered. If the variances of the two datasets are highly dissimilar, the process of bias correction will alter the correlation between the two datasets and Eq. 4-6 can be considered a naïve *a priori* estimate of the MSD.

4.5 Results

4.5.1 Temporal Stability

Temporal stability of soil moisture at LR has previously been analyzed for the study period of June 2003 [Bosch et al., 2006] during the SMEX03 campaign. The temporal stability of soil moisture presented here covers a longer time period (Apr-Sep 2010) and was restricted to those sites which are considered more reliable in terms of the accuracy of soil moisture measurements [Jackson et al., 2010]. Sites that are wetter than the watershed mean value were found to have greater variability of relative difference than dry sites and were located in areas of natural vegetation [Bosch et al., 2006]. The anomalously wet soil moisture at RG16 is believed to result from micro-topography storing water near the surface resulting low correlation with watershed average soil moisture (R = 0.68). At depths of 20 cm and 50 cm the correlation and RMSD of soil moisture at RG16 to the watershed average was comparable to that of other the sites,

supporting the conclusion that micro-topography was the cause of the wetness. Temporal persistence of soil moisture at LW was consistent with previous findings that the watershed was temporally stable and the correlation between individual sites and the watershed mean typically greater than 0.75 [Cosh et al., 2006].

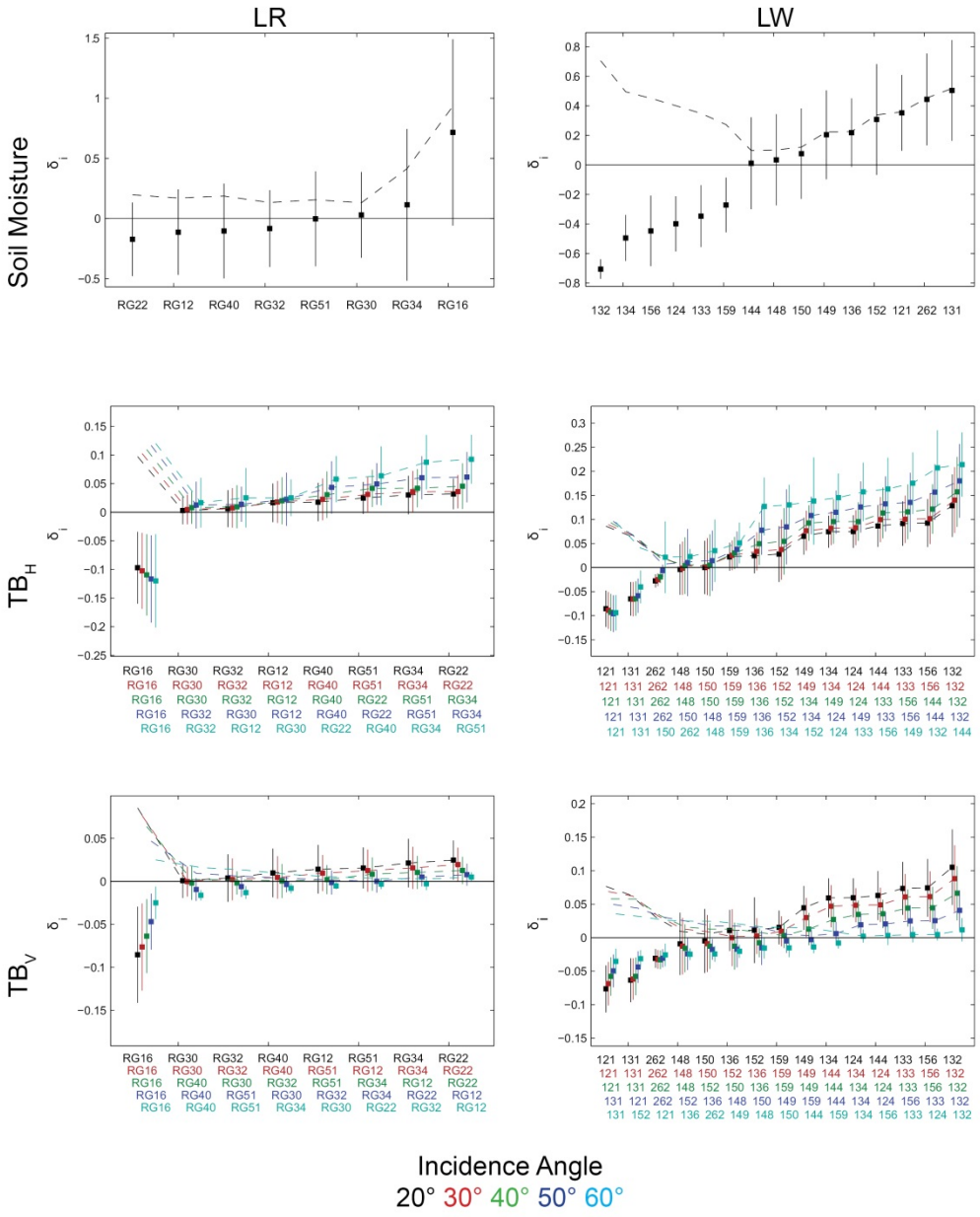

Figure 4-2: Temporal stability of soil moisture and TB. The box is the mean relative difference and the error bars are one standard deviation of the relative difference. The dashed lines are RMSE (Eq. 5-3) where low RMSE indicates a local site is representative of the watershed average.

Brightness temperature was also found to display properties of temporal stability, where the spatial pattern was opposite that of soil moisture. As expected wet sites had lower than average TB and dry sites had higher than average TB; since wet soils have lower L-band emissions than dry soils [Kerr et al., 2010]. While the mean relative difference of all three sites was influenced by soil moisture, L band emissions are also influenced by vegetation, dielectric roughness and effective temperature [Wigneron et al., 2007] all of which contribute to the order of the sites in Fig 4-2. These influences resulted in the ranks of the stations being different for TB than they were for soil moisture, particularly amongst sites with similar soil moisture conditions but different surface characteristics. The impact on site rank was more noted at LR than LW since there was greater variability in land cover at these sites.

The angle of incidence also had a strong impact on the strength of the temporal persistence and only a minor impact on the ranks of the stations. For horizontally polarized emissions, increased incidence angles resulted in greater relative difference of each site from the mean value. Also, higher incidence angles resulted in greater variability in the relative difference at each site. The opposite was found for vertically polarized TB, where high incidence angles resulted in relatively small differences from the basin mean TB_V and low variability. A sensitivity analysis to determine the relative impact of influential variables/parameters on the relative difference of TB with differing incidence angles is beyond the scope of this study; however, theoretical considerations from the literature suggest two factors which likely resulted in the observed patterns. The effect of dielectric roughness has been found to increase(decrease) H(V) polarized emissions with increasing incidence angle and is partially related to roughness induced by soil moisture [Wigneron et al., 2001; Shi et al., 2002; Wigneron et al., 2007]. The second factor; optical depth (τ_p), is modeled in L-MEB as constant for H polarization, but increases with incidence angle for V polarization [Wigneron et al., 2007]. The combined decrease in observed soil emissions due to dielectric roughness and attenuation, combined with the increased vegetation emission contribution at V polarization with increasing incidence angle could explain the lower relative difference and standard deviation found in Fig 4-2.

4.5.2 MSD Decomposition

Figure 4-3 shows the decomposition of the mean squared difference at LR and LW. For both TB and soil moisture the contribution of errors in variance/timing (*IA*) to the total MSD was

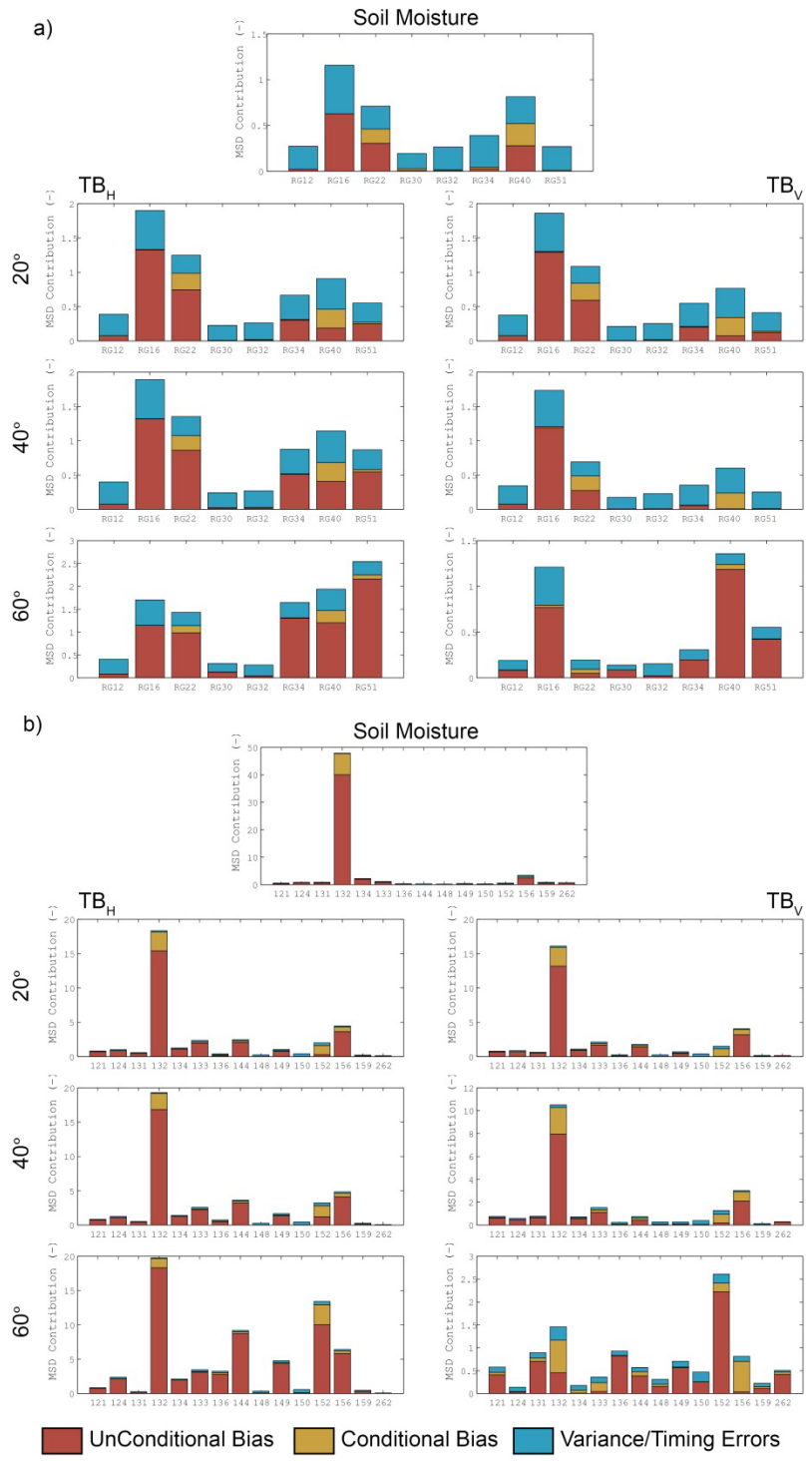

Figure 4-3: Decomposition of the mean squared difference between watershed average TB and soil moisture and that of the local sites at LR (top) and LW (bottom).

consistent across all sites in each study area. The contribution of the conditional bias (*IB*) was found to vary between sites, but was consistent between TB_H and TB_V and multiple incidence angles within each site. The greatest variation in the relative contribution to MSD was the contribution from the unconditional bias (*IC*). Figure 4-3 demonstrates that high MSD at sites which were markedly different from the mean watershed conditions were dominated by the impact of bias. There was a slight positive relationship between the bias contribution to MSD, LAI and the fraction of high vegetation, which resulted in high unconditional bias at RG16 and RG40 at LR.

As incidence angle increased, sites with relatively high LAI and

greater fractions of high vegetation also had small increases in MSD associated to increased bias. This resulted in a large increase in bias at LW 152 with increasing incidence angle. As expected based on Fig. 4-2, at V polarization the MSD decreased with incidence angle, however, the relative decomposition of MSD between sites was similar at each incidence angle. This is best exemplified looking at the magnitude of the scale of the TB_V MSD in Fig. 4-3.

4.5.3 Bias Correction

For each site the watershed average soil moisture/TB was compared against the local site observations and the watershed average bias was corrected using Eq. 4-5 to match that of the local measurements at each site. The performance in terms of RMSD and correlation for each site prior to and following bias correction can be seen in Fig. 4-4. At LW and LR, the variance of the watershed averaged values and the local measurements were similar and resulted in no significant change in correlation at any site following bias correction. At LR prior to bias correction, the RMSD between the watershed average and the local site soil moisture was below the 0.04 m^3m^{-3} target at all sites except RG16 and RG34. These sites had the lowest correlation with the watershed average soil moisture and were wetter and more variable than the other sites. Bias correction resulted in lower RMSE at all sites and only RG16 was above the target with an RMSD of 0.05 m^3m^{-3}. At LW the RMSD at 13 of 15 sites was above the 0.04 m^3m^{-3} target prior to bias correction, with a mean RMSD of 0.055 m^3m^{-3}. After the bias of the watershed mean soil moisture was corrected to match each site, the resulting RMSD at all sites was below the target, with a mean RMSD of 0.028 m^3m^{-3}.

In terms of TB the correction of bias between the watershed mean TB and that of the local site resulted in improved site representativeness as was expected from Fig. 4-3. Figure 4-3 also demonstrates that prior to bias correction the watershed average $TB_{H(V)}$ was less(more) representative of local site conditions as incidence angle increased. Correction of the bias in watershed average TB resulted in negligible (<1 K) dependence between TB and incidence angle at H polarization, while some incidence angle dependence remained at V polarization. The RMSD of TB_H was approximately 7 K and 9 K for LW and LR respectively. This was considerably above the radiometric accuracy of both SMAP (1.5 K) [Das et al., 2014] and SMOS (2 K) [Kerr et al., 2010]. The discrepancy was attributed to differences in the temporal characteristics between the watershed average TB and that of the local sites. The vegetation

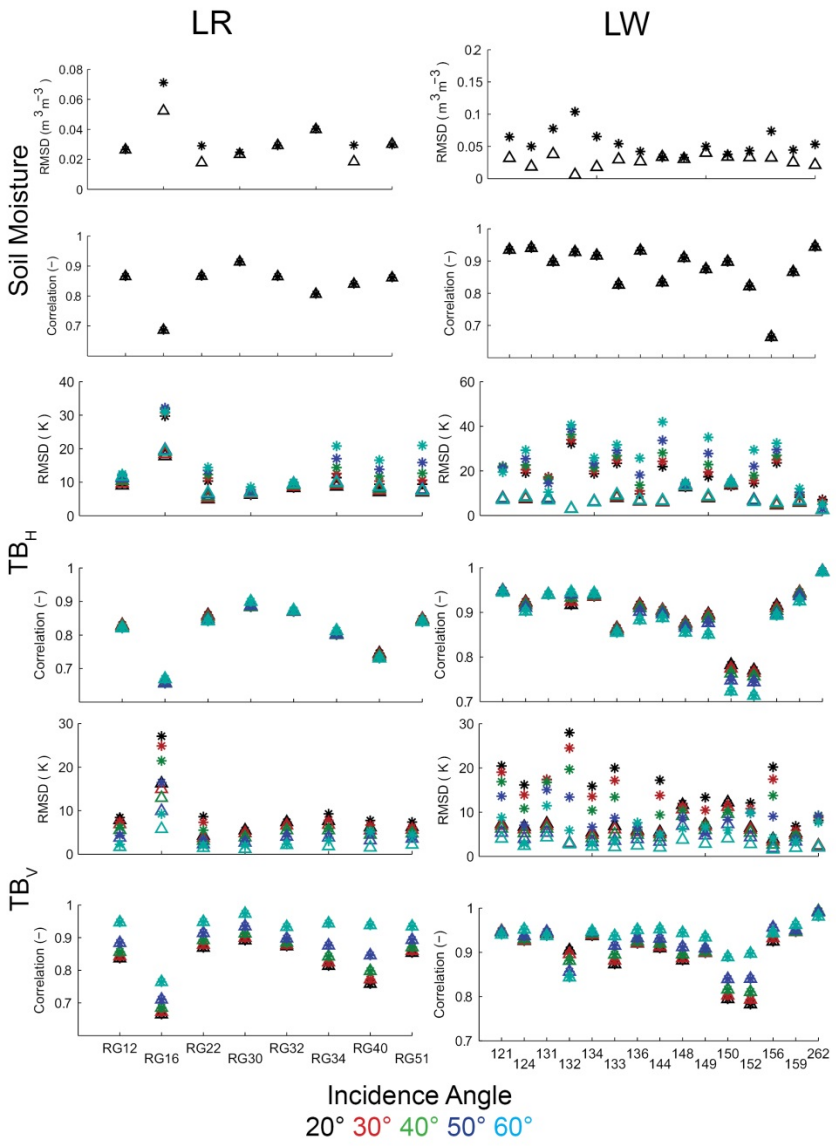

Figure 4-4: Performance of watershed average soil moisture/TB for predicting local site conditions prior to (asterisk) and after (triangle) bias correction.

contribution to and attenuation of TB could explain the higher RMSD at LR which had a greater vegetation content than LW. Prior to bias correction the standard deviation of the RMSD was 6 K at 20° incidence angle for H polarization and increased to 11 K at 60°. Following bias correction the standard deviation of RMSD between the sites was ~3 K. For V polarization the standard deviation of RMSD ranged from 6 K at low incidence angles to 3 K at high angles. Following bias correction the standard deviation was approximately 3 K and 1 K for low and high incidence angles respectively. These results demonstrate that bias correcting the watershed average TB at each incidence angle not only increased the representativeness of the average watershed TB for each site, but also resulted in consistent performance between sites at both H and V polarizations.

4.6 Discussion and Conclusions

Soil moisture and brightness temperature at two USDA ARS watersheds were shown to display properties of temporal stability. Through the decomposition of the mean squared difference

between the watershed average TB/soil moisture and that at local sites, it was determined that when the MSD was large; a major contribution to MSD was the unconditional bias.

A limitation of this study was the assumption that a single air temperature measurement was representative of the small watersheds and ground temperature could be used as a proxy for vegetation canopy temperature. This assumption may lead to a TB error of approximately 3 K [Wigneron et al., 2007], but is expected to be consistent across the study domain. In this study, both watersheds are considered as nominal targets for soil moisture retrievals. Future research should also determine if TB is temporally stable under non-nominal conditions such as higher fractional vegetation cover, mixed soil types and mixed topography. The bias correction approach to downscaling assumes that precipitation and evapotranspiration are major drivers of the spatial distribution of soil moisture. The influence of other drivers, such as micro-topography at RG16, can be expected to deteriorate the effectiveness of the technique by decreasing the correlation between a site and the watershed average soil moisture. The assumption was also made that precipitation was uniform at the scale of the watershed and did not homogenize the soil moisture pattern. The limitations of this assumption were not explicitly assessed herein, although it is less likely to be limiting at smaller scales. It has also been demonstrated that following large rainfalls temporally persistent spatial patterns return quickly [Kornelsen and Coulibaly, 2013b].

The mean relative difference of TB at each site from the average TB was impacted by the incidence angle of the observation. At H(V) polarization the surface was more homogenous at low(high) incidence angles. The associated MSD between the mean state and that of local sites increased with incidence angle at H polarization and decreased with incidence angle at V polarization. The incidence angle dependence of MSD was mainly the result of unconditional bias which was easily corrected.

The results demonstrate that when *a priori* information is available about the surface state, such as the model state in a land data assimilation approach, standard bias correction methods can be effectively used to apply satellite observations at sub-grid scales. This ensues because much of the spatial difference between sites can be attributed to temporally persistent biases in both soil moisture and brightness temperature. The bias correction applied herein was a simple mean-variance matching approach, although better results could be expected with advanced bias correction techniques [i.e. Drusch et al., 2005; Kornelsen and Coulibaly,

manuscript in preparation, 2014]. Similar to other downscaling methods, the proposed approach cannot be assumed to identify the 'true' soil moisture state at high resolution, but rather a higher resolution state that contains the biases of the reference dataset, in this case the point scale measurements. This contrasts with downscaling methods based on satellite observations which maintain the biases of the remotely sensed measurements [Merlin et al., 2012].

4.7 References

Baghdadi, N., M. Aubert, O. Cerdan, L. Franchisteguy, C. Viel, E. Martin, M. Zribi, and J. F. Desprats (2007), Operational mapping of soil moisture using synthetic aperture radar data: Application to the Touch Basin (France). Sensors 7(10), 2458-2483, doi:10.3390/s7102458.

Bartalis Z., W. Wagner, V. Naeimi, S. Hasenauer, K. Scipal, H. Bonekamp, J. Figa, and C. Anderson (2008), Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT), Geophys. Res. Lett. 34, L20401, doi:10.1029/2007GL031088.1.

Bolton, J. D., W.T. Crow, X. Zhan, T.J. Jackson, C.A. and Reynolds (2010), Evaluating the utility of remotely sensed soil moisture retrievals for operational agricultural drought monitoring. IEEE J. Sel. Topics Appl. Earth. Observ. 3(1) 57-66, doi: 10.1109/JSTARS.2009.2037163

Bosch, D.D., V. Lakshmi, T.J. Jackson, M. Choi and J.M. Jacobs (2006), Large scale measurements of soil moisture for validation of remotely sensed data: Georgia soil moisture experiment of 2003, J. Hydrol. 323(1), 120-137, doi: 10.1016/j.jhydrol.2005.08.024

Bosch, D.D., J.M. Sheridan, R.R. Lowrance, R.K. Hubbard, T.C. Strickland, G.W. Feyereisen and D.G. Sullivan (2007), Little River Experimental Watershed database, Water Resour. Res.43, W09470, doi:10.1029/2006WR005844.

Broxton, P.D., X. Zeng, D. Sulla-Menashe and P.A. Troch (2014), A global land cover climatology using MODIS data, J. Appl. Meteor. Climatol. 53(6), 1593-1605, doi:10.1175/JAMC-D-13-0270.1.

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

Bronstert A., B. Creutzfeldt, T. Graeff, I. Hajnsek, M. Heistermann, S. Itzerott, T. Jagdhuber, D. Kneis, E. Lück, D. Reusser and E. Zehe (2012), Potentials and constraints of different types of soil moisture observations for flood simulations in headwater catchments, Nat. Hazards, 60(3), 879-914, doi:10.1007/s11069-011-9874-9.

Chauhan, N.S., S. Miller, and P. Ardanuy (2003), Spaceborne soil moisture estimation at high resolution: A microwave-optical/IR synergistic approach, Int. J. Remote Sens., 24(22), 4599–4622.

Cosh, M.H., T.J. Jackson, R. Bindlish, and J.H. Prueger (2004), Watershed scale temporal and spatial stability of soil moisture and its role in validating satellite estimates, Remote Sens. Environ, 92(4), 427-435, doi:10.1016/j.rse.2004.02.016.

Cosh, M.H., T.J. Jackson, P. Starks, and G. Heathman (2006), Temporal stability of surface soil moisture in the Little Washita River watershed and its applications in satellite soil moisture product validation, J. Hydrol., 323(1), 168-177, doi:10.1016/j.hydrol.2005.08.020.

Cosh, M.H., T.J. Jackson, S. Moran and R. Bindlish (2008), Temporal persistence and stability of surface soil moisture in a semi-arid watershed, Remote Sens. Environ., 112(2), 304-313, doi:10.1016/j.rse.2007.07.001.

Cosh, M. and J. Birch. 2008. SMEX03 Ancillary Soil Characteristics Data. [Soil Texture]. Boulder, Colorado USA: NASA DAAC at the National Snow and Ice Data Center.

Das, N.N., D. Entekhabi, E.G. Njoku, J.J.C. Shi, J.T. Johnson and A. Colliander (2014), Tests of the SMAP combined radar and radiometer algorithm using airborne field campaign observations and simulated data, IEEE Trans. Geosci. Remote Sens. 52(4), 2018-2028, doi:10.1109/TGRS.2013.2257605.

de Rosnay, P., J-P. Wigneron, T. Holmes and J.-C. Calvet (2006), Parameterizations of the effective temperature for L-band radiometry. Inter-comparison and long term validation with SMOSREX field experiment, in Thermal Microwave Radiation- Applications for Remote Sensing, edited by Mätzler, C., P.W. Rosenkranz, A. Battaglia, and J.P. Wigneron, IET Electromagnetic Waves Series 52, London, UK.

de Rosnay P., M. Drusch, A. Boone, G. Balsamo, B. Decharme, P. Harris, Y. Kerr, T. Pellarin, J. Polcher and J.-P. Wigneron (2009), The AMMA Land Surface Model Intercomparison Experiment coupled to the Community Microwave Emission Model: ALMIP-CMEM", J. Geophys. Res., 114, D05108, doi:10.1029/2008JD010724.

Draper, C.S., J.P. Walker, P.J. Steinle, R.A.M. de Jeu and T.R.H. Holmes (2009), An evaluation of AMSR-E derived soil moisture over Australia, Remote Sens. Environ., 113(4), 703-710, doi:10.1016/j.rse.2008.11.011.

Drusch M., E.F. Wood, and H. Gao (2005), Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture, Geophys. Res. Lett. 32, L15403, doi:10.1029/2005GRL023623.

Drusch, M. (2007). Initializing numerical weather prediction models with satellite-derived soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and the TMI soil moisture data set. J. Geophys. Res., 112, D03102, doi:10.1029/2006JD007478.

Drusch, M., T. Holmes, P. de Rosnay and G. Balsamo (2009) Comparing ERA-40 based L-band brightness temperatures with Skylab observations: A calibration / validation study using the Community Microwave Emission Model, J. Hydromet. 10(1), 213-226, doi: 10.1175/2008JHM964.1.

Dumedah, G., J.P. Walker and C. Rüdiger (2013), Can SMOS data be used directly on the 15-km Discrete Global Grid?, IEEE Trans. Geosci. Remote Sens. 52(5), 2538-2544, doi:10.1109/TGRS.2013.2262501.

Entekhabi, D., E.G. Njoku, P.E. O'Neill, K.H. Kellogg, W.T. Crow, W.N. Edelstein, J.K. Entin, S.D. Goodman, T.J. Jackson, J. Johnson, J. Kimball, J.R. Piepmeier, R.D. Koster, N. Martin, J.K McDonald, M. Moghaddam, S. Moran, R. Reichle, J.C. Shi, M.W. Spencer, S.W. Thurman, L. Tsang, and J. van Zyl (2010). The Soil Moisture Active Passive (SMAP) Mission. Proceedings of the IEEE, 98(5), 704–716, doi:10.1109/JPROC.2010.2043918.

Fang, H. L., S. L. Liang, J. R. Townshend, and R. E. Dickinson (2008), Spatially and temporally continuous LAI data sets based on an integrated filtering method: Examples from North America, Remote Sens. Environ., 112(1), 75– 93, doi:10.1016/j.rse.2006.07.026.

Heathman G.C., M. Larose, M.H. Cosh and R. Bindlish (2009), Surface and profile soil moisture spatio-temporal analysis during an excessive rainfall period in the Southern Great Plains, USA, Catena 78(1), 159-169, doi: 10.1016/j.catena.2009.04.002.

Holmes T.R.H., P. de Rosnay, R.A.M. de Jeu, J.-P. Wigneron, Y. Kerr, J.-C. Calvet, M.J. Escorihuela, K. Saleh and F. Lemaître (2006), A new parameterization of the effective temperature for L band radiometry, Geophys. Res. Lett. 33, L07405, doi:10.1029/2006GL025724.

Jackson, T.J., M. Cosh, R. Bindlish, P. Starks, D. Bosch, M. Seyfried, D. Goodrich, S. Moran and J. Du, (2010) Validation of Advanced Microwave Scanning Radiometer soil moisture products, IEEE Trans. Geosci. Remote Sens. 48(12), 4256–4272, doi:10.1109/TGRS.2010.2051035.

Jackson, T. J., R. Bindlish, M.H. Cosh, Z. Tianjie, P.J. Starks, D.D. Bosch, M. Seyfried, M.S. Moran, D.C. Goodrich, Y.H. Kerr and D. Leroux (2012), Validation of Soil Moisture and Ocean Salinity (SMOS) soil moisture over watershed networks in the U.S. IEEE Trans. Geosci. Remote Sens. 50(5), 1530–1543, doi:10.1109/TGRS.2011.2168533.

Jacobs, J.M., B.P. Mohanty, E.-C. Hsu and D. Miller (2004), SMEX02: Field scale variability, time stability and similarity of soil moisture, Remote Sens. Environ. 92(4), 436-446, doi:10.1016/j.rse.2004.02.017.

Kerr, Y. H., P. Waldteufel, J.-P. Wigneron, S. Delwart, F. Cabot, J. Boutin, M.-J. Escorihuela, J. Font, N. Reul, C. Gruhier, S.E. Juglea, M.R. Drinkwater, A. Hahne, M. Martin-Neira and S.

Mecklenburg (2010), The SMOS mission: New tool for monitoring key elements ofthe global water cycle. Proceedings of the IEEE, 98(5), 666–687, doi: 10.1109/JPROC.2010.2043032.
Kim J. and T.S. Hogue (2012), Improving spatial soil moisture representation through integration of AMSR-E and MODIS products, IEEE Trans. Geosci. Remote Sens. 50(2), 446-460, doi:10.1109/TGRS.2011.2161318.

Kornelsen, K. C., and P. Coulibaly (2013a) Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications. J. Hydrol. 476(1), 460-489, doi:10.1016/j.hydrol.2012.10.044.

Kornelsen, K.C. and P. Coulibaly (2013b), The McMaster Mesonet soil moisture dataset: description and spatio-temporal variability analysis, Hydrol. Earth Syst. Sci. 17, 1589-1606, doi:10.5194/hess-17-1589-2013.

Loew A. and W. Mauser (2008), On the disaggregation of passive microwave soil moisture data using a priori knowledge of temporally persistent soil moisture fields, IEEE Trans. Geosci. Remote Sens. 46(3), 819-834, doi:10.1109/TGRS.2007.914800.

Mascaro, G., E.R. Vivoni and R. Deidda (2011), Soil moisture downscaling across climate regions and its emergent properties, J. Geophys. Res. 116, D22114, doi:10.1029/2011JD016231.

Merlin O., C. Rüdiger, A. Al Bitar, P. Richaume, J.P. Walker and Y. Kerr (2012), Disaggregation of SMOS soil moisture in southeastern Australia, IEEE Trans. Geosci. Remote Sens. 50(5), 1556-1571, doi:10.1109/TGRS.2011.2175000.

Merlin, O., A. Al Bitar, J.P. Walker and Y. Kerr (2010), An improved algorithm for disaggregating microwave-derived soil moisture based on red, near-infrared and thermal-infrared data, Remote Sens. Environ. 114(10), 2305-2316, doi:10.1016/j.rse.2010.05.007.

Mironov, V.L., M.C. Dobson, V.H. Kaupp, S.A. Komarov and V.N. Kleschenko (2004), Generalized refractive mixing dielectric model for moist soils, IEEE Trans. Geosci. Remote Sens. 42(4), 773-758, doi:10.1109/TGRS.2003.823288.

Murphy, A.H. (1988), Skill scores based on the mean square error and their relationships to the correlation coefficient, Monthly Weather Rev. 116, 2417-2424.

Parada, L.M. and X. Liang (2003), A downscaling framework for L band radiobrightness temperature imagery, J. Geophys. Res. 108, D228863, doi:10.1029/2003JD003637.

Parada L.M., and X. Liang (2008), Impacts of spatial resolutions and data quality on soil moisture data assimilation, J. Geophys. Res. 113, D10101, doi:10.1029/2007JD009037.

Pellarin, T., J.-P. Wigneron, J.-C. Calvet, and P. Waldteufel (2003) Global soil moisture retrieval from a synthetic L-band brightness temperature data set, J. Geophys. Res. 108(D12), 4364, doi:10.1029/2002JD003086.

Reichle, R.H. and R.D. Koster (2004), Bias reduction in short records of satellite soil moisture, Geophys. Res. Lett. 31, L19501, doi:10.1029/2004GL020938.

Schaefer, G.L., M.H. Cosh and T.J. Jackson (2007), The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). J. Atmos. Oceanic Tech. 24(12), 2073-2077, doi:10.1175/2007JTECHA930.1.

Seneviratne, S.I., T. Corti, E.L. Davin, M. Hirschi, E.B. Jaeger, I. Lehner, B. Orlowski, and A. J. Teuling (2010), Investigating soil moisture-climate interactions in a changing climate: A review, Earth Sci. Rev. 99(1), 125-161, doi:10.1016/j.earscirev.2010.02.004.

Seneviratne, S.I., M. Wilhelm, T. Stanelle, B. van den Hurk, S. Hagemann, A. Berg, F. Cheruy, M.E. Higgins, A. Meier, V. Brovkin, M. Claussen, A. Ducharne, J.-L. Dufrense, K. L. Findell, J. Ghattas, D.M. Lawrence, S. Malyshev, M. Rummukainen and B. Smith (2013), Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment. Geophys. Res. Lett. 40(19) 5212-5217, doi: 10.1002/grl.50956.

Shi, J., K.S. Chen, Q. Li, T.J. Jackson, P.E. O'Neill, and L. Tsang (2002), A parameterized surface reflectivity model and estimation of bare-surface soil moisture with L-band radiometer. IEEE Trans. Geosci. Remote Sens. 40(12), 2674−2686.

Sullivan, D.G., H.L. Batten, D. Bosch, J. Sheridan and T. Strickland (2007), Little River Experimental Watershed, Tifton, Georgia, United States: A geographic database, Water Resour. Res. 43, W09471, doi:10.1029/2006WR005836.

Vachaud, G., A. Passerat De Silans, P. Balabanis and M. Vauclin (1985), Temporal stability of spatially measured soil water probability density function, Soil Sci. Soc. Am. J. 49, 822-828.

Vanderlinden, K., H. Vereecken, H. Hardelauf, M. Herbst, G. Martinez, M.H. Cosh and Y.A. Pachepsky (2012), Temporal stability of soil water contents: A review of data and analyses, Vadose Zone J. 11(4) doi:10.2136/vzj2011.0178.

Wagner, W., G. Lemoine and H. Rott (1999), A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data, Remote Sens. Environ. 70(1), 191–207.

Weglarczyk, S. (1998), The interdependence and applicability of some statistical quality measures for hydrological applications, J. Hydrol. 206(1), 98-103.

Wigneron, J. -P., L. Laguerre, and Y. Kerr (2001), Simple modeling of the L band microwave emission from rough agricultural soils. IEEE Trans. Geosci. Remote Sens. 39(8), 1697−1707.

Wigneron, J.-P., Y. Kerr, P. Waldteufel, K. Saleh, M.J. Escorihuela, P. Richaume, P. Ferrazzoli, P. de Rosnay, R. Gurney, J.-C. Calvet, J.P. Grant, M. Guglielmetti, B. Hornbuckle, C. Mätzler, T. Pellarin and M. Schwank (2007), L-band microwave emission of the biosphere (L-MEB) model: Description and calibration against experimental data sets over crop fields. Remote Sens. Environ. 107(4), 639–655, doi:10.1016/j.rse.2006.10.014.

Wilson, D.J., A.W. Western and R.B. Grayson (2005), A terrain and data-based method for generating the spatial distribution of soil moisture, Adv. Water Resour. 28(1), 43-54, doi:10.1016/j.advwatres.2004.09.007.

Chapter 5: Downscaling Satellite Soil Moisture and Brightness Temperature for a Land-Surface Model using Bias Correction

Summary of Paper 4: Kornelsen, K.C., Davison, B., and Coulibaly, P. (submitted) Downscaling Satellite Soil Moisture and Brightness Temperature for a Land-Surface Model using Bias Correction, IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing.

Summary:

This research expands on that presented in Chapter 5 by relaxing assumptions about spatial homogeneity of LAI and evaluating the technique on a fully distributed dataset derived from the MESH hydrological land surface scheme. MESH was set up and calibrated on the Hamilton-Halton Watershed domain and represents both the highest resolution and most detailed application of MESH to date in terms of accounting for sub-grid heterogeneity. MESH was coupled to CMEM to produce online distributed TB simulations at both polarizations and multiple incidence angles. MESH was validated using data from the McMaster Mesonet and MESH-CMEM was validated using multi-angular TB observations and soil moisture retrievals from SMOS. Both simulated and observed soil moisture and TB at the SMOS grid cell scale was downscaled to the scale of individual grid cells (~800 m) using CDF matching.

The primary results of this research were:

- In comparison to SMOS observations, MESH-CMEM simulated TB at all incidence angles that were consistent with theoretical considerations and skill comparable to other CMEM applications.
- MESH-CMEM TB was more consistent with SMOS observations than MESH soil moisture was with SMOS soil moisture retrievals. This results from the use of unrelated ancillary data in the SMOS retrieval algorithm and supports other considerations that data assimilation should use observed TB and not retrieved soil moisture.
- Downscaling SMOS TB and soil moisture using CDF matching to MESH-CMEM grid cell scale resulted in improved performance at nearly all grid cells. The performance of the downscaled variables at individual grid cells was comparable to the global (SMOS scale) performance, demonstrating information recovery through downscaling.

5.1 Abstract

The assimilation of soil moisture and brightness temperature (TB) are expected to improve the modelling of land surface processes, but are only available at a resolution that is far coarser than the scale of many hydrological processes. Due to systematic differences between model states and satellite observations, bias correction is a necessary step in land data assimilation schemes and was evaluated as a method to downscale coarse scale satellite observations to fine scale model grid cells (~800 m). This was done by coupling the Modélisation Environmentale Communautaire–Sufrace Hydrolgy (MESH) Hydrological Land-Surface Scheme to the Community Microwave Emissions Model (CMEM) to simulate soil moisture and TB. MESH-CMEM was validated and found to be in good agreement with observations from the Soil Moisture and Ocean Salinity (SMOS) satellite at the scale of SMOS data products ($R \approx 0.55$), with simulated TB being better correlated than soil moisture retrievals. Following downscaling, TB and soil moisture retrievals had comparable performance to coarse resolution SMOS data. Bias correction of TB was more reliable than soil moisture. These findings indicate that both TB and soil moisture retrievals can be assimilated in a land surface model at moderate to high resolution with a simple observation operator.

5.2 Introduction

Soil moisture is an important geophysical state variable which influences runoff, evapotranspiration and energy fluxes. The assimilation of soil moisture information can benefit hydrological predications of floods and drought [1]-[3], agricultural applications [4], [5] and numerical weather prediction [6], [7]. An important source of soil moisture information for watershed to continental scale applications is remote sensing from microwave satellites such as the Advanced Scatterometer (ASCAT) [8], the Soil Moisture and Ocean Salinity (SMOS) satellite [9] or the Soil Moisture Active Passive (SMAP) mission [10]. Validation studies have demonstrated that observations from these satellites can capture the temporal variability of soil moisture [11]-[14]; however, their use in hydrological and agricultural predictions is limited because of the discrepancy in scale between the satellite products (> 25 km) and that of hydrological processes (< 1 km) [15].

The discrepancy in scale between applications and soil moisture products has resulted in several methods to disaggregate soil moisture retrievals to a higher spatial resolution. In the context of the SMOS mission, [15], [16], [17] have developed a method to disaggregate SMOS soil moisture in semi-arid environments taking advantage of the relationship between surface soil moisture and evaporative efficiency using optical remote sensing data. There have also been many empirical approaches to downscale soil moisture using model information, land surface databases and multi-scale statistical relationships and data-driven approaches [18]- [20]. In many cases, despite the disaggregation of soil moisture, biases of the sensor(s) used to create the products remain and must be corrected prior to data assimilation [15]. Similar to soil moisture retrievals, brightness temperature (TB) observations from remote sensing sources are biased compared to modelled TB and require correction prior to assimilation [21], [22].

Using data from two U.S. Department of Agriculture (USDA) experimental watersheds [23] and the Community Microwave Emission Model (CMEM) [21], [24] demonstrated that the temporal variability of both soil moisture and TB at the watershed scale was representative of the temporal variability at local sites. It was also shown that much of the spatial differences at these sites could be attributed to a temporally persistent bias, which could be easily removed with a simple bias correction method [24].

The prior analysis of [24] made several assumptions about the spatial representativeness of in situ observations and some ancillary data required by CMEM. This study will expand upon the work of Kornelsen et al. [24] by coupling CMEM with the Modélisation Environmentale Communautaire – Surface Hydrology (MESH) Hydrological Land-Surface Scheme (HLSS) [25], [26]. The lateral overland and subsurface flow that distinguishes MESH from the CLASS LSS has been found to be effective for predicting streamflow in many parts in Canada [25]-[32]. The use of a distributed HLSS removes assumptions made about the spatial representativeness of point scale soil moisture and air temperature and allows for verification of the results of [24] using a dataset with greater spatial diversity than is available with in situ measurements. Further, TB observations and soil moisture retrievals from SMOS will be downscaled to model grid cell resolution (~800 m) for comparison. The study domain is mid-latitude temperate watershed in Southern Ontario, Canada with primarily agricultural land use mixed with forests and urban areas and uneven terrain. This landscape allows for the evaluation of bias correction for downscaling satellite observations in differing types of environments.

5.3 Study Area and Data

5.3.1 Study Area

The selected study area was in the Hamilton-Halton (HH) watersheds on the north-western shore of Lake Ontario. The model domain was located in a temperate region of the mid-latitudes and has imperfectly drained loam soils. The land cover is dominated by agricultural land uses, where wheat, corn and soy are major crops. The study domain also contains many areas of mixed forest cover and some urban areas as seen in Fig. 6-1. Regions with these characteristics are rarely studied for soil moisture research and calibration/validation activities, largely due to the lack of dense networks. The MESH model domain extends over several large watersheds; however, for consistency a subset of the model domain indicated in Fig. 6-1 was used for this study. The study domain covers an area approximately 20 km × 20 km and was chosen to cover mainly agricultural areas with some forest stands, while minimizing the impact of urban areas. The Niagara Escarpment provides a topographic barrier between the west and east of the study domain. Wetlands and small bodies of open water were most prevalent in the western part of the study area atop the escarpment. The escarpment also divides soil types, where loam soils dominate to the west and clay loam soils are found in the lower plains to the east. Silt-loam soils are found in between and are associated with forested areas along the escarpment and in river flood plains.

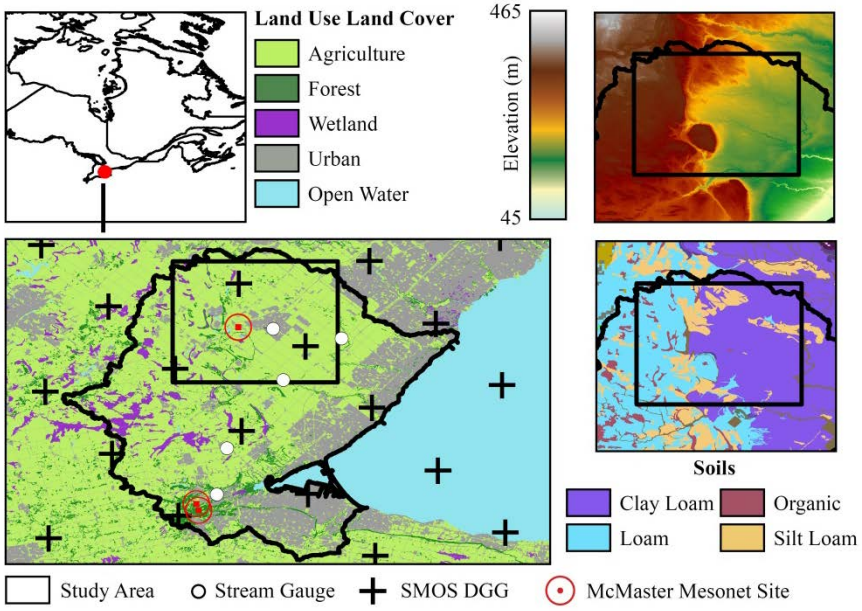

Figure 5-1: The Hamilton-Halton Watershed MESH-CMEM model domain, with the subset-study area indicated.

The HH study domain was larger than the discrete global grid (DGG) that SMOS products are distributed on but smaller than the 43 km SMOS ground resolution. The discrepancy in scale between the HH study domain and the SMOS ground resolution is not considered a hindrance for two reasons. First, [33] assessed the uncertainty of airborne radiometer measurements aggregated to SMOS scale and concluded that SMOS data could be used directly on the DGG scale because the uncertainty of oversampling was lower than the instrument radiometric uncertainty. Second, [11] explained that 80% of the signal at the centre of a grid node is associated to a circular area of 20 km around the node centre.

A long continuous record of hourly soil moisture observations was available for these watersheds [34], which were used to calibrate and validate the MESH model. The McMaster Mesonet consists of three sites located in different land-cover/soil conditions, each with Campbell Scientific CS616 time domain reflectometry (TDR) probes arranged in a square array with nine soil profile stations at six depths from 10 cm to 100 cm. Soil moisture has been recorded hourly with minor gaps since 2007 [34]. Stream flow observations for model calibration were retrieved from the Environment Canada HYDAT database (https://www.ec.gc.ca/rhc-wsc/).

5.3.2 SMOS Data

SMOS was launched by the European Space Agency (ESA) in 2009 and began normal operations in 2010. The sole instrument on SMOS is a 1.4 GHz (L-band) 2-D interferometric radiometer that has a constant Earth fixed forward tilt angle of 32.5° providing an angular coverage of -10° to +60° [9]. By acquiring multiple images of the same area on the Earth's surface during each overpass, SMOS provides multi-angular data for each DGG at both horizontal and vertical polarizations. The SMOS orbit crosses the equator each day at 6 A.M. (ascending) and 6 P.M. (descending) and images the globe once every three days [9]. In this study, reprocessed (Version 5.05) multi-angular TB observations from the Level-1C (L1C) products are used for analysis with MESH and CMEM simulations. All SMOS L1C data are filtered using the SMOS flags for RFI, border field of view, and sun points.

The multi-angular TB information from SMOS were used to retrieve soil moisture in the SMOS level 2 (L2) data products by iterative minimization of a cost function between observed and modelled TB which accounts for soil and vegetation characteristics [35]. The primary forward model used for soil moisture retrieval is the L-Band Microwave Emission of the

Biosphere model (L-MEB) [36]. Data in this study were from a reprocessing campaign using version 5.51 of the operational soil moisture processor. A major update of the ver. 5.51 processor over previous versions was the switch to the dielectric model of [37] which is expected to provide better results particularly under dry conditions. All SMOS L2 data products were filtered to remove retrievals with a probability of RFI greater than 0.1, and data quality index (DQX) greater than 0.1.

Both L1C and L2 SMOS data products are oversampled from the ~43 km SMOS sampling resolution to the 15 km DGG. To mitigate the potential impacts of this oversampling, the mean TB and soil moisture of the three DGG nodes located within the study area were considered to be representative of the 'true' observations of SMOS. For SMOS TB observations, all data were binned to the nearest integer incidence angle using a bin size of ±0.5°.

5.4 Models and Methods

5.4.1 CMEM

The Community Microwave Emissions Modelling platform (CMEM) [21], [22] was developed at the European Centre for Medium-Range Weather Forecasting (ECMWF) as a forward operator for passive microwave TB. CMEM is modular software and integrates many of the models/procedures used in the SMOS L2 soil moisture processor. CMEM was used to model microwave emissions in L-Band using L-MEB [36] to model the soil and vegetation contribution to top of the atmosphere TB. Smooth surface emissivity was modelled using the Fresnel law and rough surface emissivity was modelled following [36]. The atmospheric opacity was determined as in [38]. The parameterization of [39] was used to calculate effective temperature following the recommendation of [40]. Similar to ver. 5.51of the SMOS L2 processor the model of [37] was selected to convert the dielectric constant to soil moisture. CMEM was used to simulated brightness temperature at incidence angles of 20°, 30°, 40°, 50° and 60° for both polarizations to cover the range of incidence angles available from SMOS [9] and the SMAP nominal 40° incidence angle [10].

5.4.2 MESH

The Modélisation Environmentale Communautaire (MEC) Surface and Hydrology (MESH) is a semi-distributed Hydrologic Land-Surface Scheme first introduced by [26] with continued development at Environment Canada [25]. The current version of MESH was developed by implementing lateral surface and sub-surface hydrological processes into version 3.6 of the Canadian Land Surface Scheme (CLASS) [41] and applying the hydrological routing schemes used by WATFLOOD [39]. In comparison to 1D soil-vegetation atmosphere transfer schemes such as CLASS, each grid cell in MESH is sloped to better simulate surface runoff using Manning's equation and interflow in the shallow surface layer is generated by a parameterization of Richard's equation [26]. A tile approach is used to account for heterogeneity within each grid cell, where the spatial coverage of each hydrologically important landscape type is tracked within each grid. MESH was calibrated using the grouped-response unit (GRU) approach [42], where the landscape was categorized into representative groups (.i.e. by soil, land use, etc.) and the group model parameters calibrated simultaneously across the entire domain. This study used seven GRUs: agriculture on clay loam soils, agriculture on loam soils, agriculture on silt loam soils, mixed forest, wetlands, open water and impervious surfaces. The sub-grid movement of water and energy are calculated for each tile/GRU by CLASS routines and are aggregated by spatial weighting for each grid cell. Water is then routed between grid cells by the WATFLOOD routing scheme using Manning's equation combined with a simplified continuity equation [27].
MESH was run as an offline model and was forced with output from the Global Environmental Multiscale model (GEM) [43], [44] and the Canadian Precipitation Analysis (CaPA) [45]. The representativeness of GEM and CaPA data for the HH watershed were verified using weather stations of the McMaster Mesonet. The model was spun-up for six months prior to calibration or the model run from which data were analysed.

To facilitate the modelling of TB, MESH was coupled to CMEM to produce TB for both H and V polarization at multiple incidence angles at each time-step of MESH. All input variables required for CMEM, except ECOCLIMAP land cover types, are extracted from MESH at run-time and TB was simulated for each tile within a MESH grid cell. For analysis the tiles were aggregated using a spatially weighted mean based on tile coverage to grid cell resolution. Urban areas in MESH were treated as bare surfaces in CMEM and represent a potential source of error when TB values are aggregated to a large area.

5.4.3 MESH Calibration

In this study MESH was setup to operate with a grid size of 0.00826° (~800 m) over the study domain. The model was calibrated using a two-step approach following [30], where first vegetation parameters for 1D CLASS were calibrated and these parameters were then applied in MESH for the calibration of parameters that impacted the 2D hydrology.

Table 5-1: CLASS vegetation parameters for each land-cover type from the first calibration step.

Parameter	Description	Calibrated value			
		Needle-leaf	Broad-leaf	Crops	Grass
PAMX	Annual max plant area index [ ]	2.0	2.9	4.0	5.0
PAMN	Annual min plant area index [ ]	1.9	0.2	0.0	1.3
LNZ0	Natural Log of max roughness length [ ]	-5.5	0.1	-4.6	-2.0
ALVC	Average visible albedo fully-leafed [ ]	0.05	0.06	0.06	0.03
ALI	Average near-IR albedo fully-leafed [ ]	0.2	0.2	0.3	0.3
ROOT	Annual max rooting depth [m]	1.1	1.5	1.2	0.5
CMAS	Annual max canopy mass [kg m^{-2}]	28.5	30.0	3.8	1.9
RSMN	Min stomatal resistance [s m^{-1}]	193	116	85	76
QA50	Reference shortwave radiation [Wm^{-2}]	30	25	30	21
VPDA	Vapour pressure deficit coefficient [ ]	0.94	0.66	0.50	0.98
VPDB	Vapour pressure deficit coefficient [ ]	1.27	0.46	1.00	0.62
PSGA	Soil moisture suction coefficient [ ]	122	121	100	76
PSGB	Soil moisture suction coefficient [ ]	5	5	5	3

First, representative stations of the McMaster Mesonet [34] were identified for each vegetation type; needle leaf, broad leaf, crops and grass. The parameters which impact vegetation and evapotranspiration in CLASS were calibrated using a semi-supervised approach using the Dynamically Dimensioned Search algorithm (DDS) [46] in the OSTRICH software package [47] to find a set of parameters that minimized the mean squared error of root-zone soil moisture for all sites in a vegetation category. The CLASS parameters are given in Table 5-1. Parameters were manually checked for consistency with published CLASS parameters [41] and were rejected and automatic calibration conducted again if the parameters were beyond the tolerance specified in Table 5-1. If multiple sites for a land cover type were available, this process continued until the parameter set resulted in good performance at all sites. This allowed for the

identification of parameters representing the needle leaf, broad leaf and grass vegetation categories to be applied to each GRU. For crops and urban land covers, where no information was available, literature CLASS parameters were used [41]. In the second calibration step, DDS was used to calibrate the parameters that govern overland flow, interflow, baseflow and routing using DDS to maximize the mean Nash-Sutcliffe efficiency from five stream gauges in the watershed seen in Fig. 5-1. These parameters are identified in Table 5-2. The model was calibrated on a 3.07 GHz Intel Xeon workstation running Windows 7 with 24 GB of memory. Each model run took approximately 2.1 hours and the calibration was conducted using 1650 model runs which required approximately 3 500 hours of computer time. This process was repeated four times using random initial conditions and the best parameter set selected herein.

Table 5-2: Calibrated MESH hydrology parameters for each GRU from the second calibration step.

Parameter	Description	GRU 1	GRU 2	GRU 3	GRU 4	GRU 5	GRU 6	GRU 7
DRN	Drainage Index [ ]	0.29	0.005	0.05	0.95	$6.0E^{-4}$	0.5	0.07
DD	Drainage density [km km^{-2}]	47.4	60.4	8.5	21.7	68.8	40.3	22.2
XLSP	Average overland slope [ ]	0.06	0.004	0.03	0.5	$6.0E^{-4}$	0.7	0.6
GRKF	K^*/Ks [ ]	0.2	0.3	0.01	0.01	0.8	0.2	0.02
MANN	Manning's n	1.4	1.3	0.6	1.2	0.07	1.9	0.53
WFCI	Ks^* [ms^{-1}]	0.67	0.33	0.39	0.42	$1.0E^{-5}$	0.03	0.01

***K is the horizontal hydraulic conductivity at depth and Ks is the horizontal saturated hydraulic conductivity at the surface**

5.4.4 Analysis Methods

A decomposition of the mean squared difference was used to determine the structure of the differences between SMOS and fine scale soil moisture or TB [24]. The mean squared difference:

$$\mathrm{MSD} = \frac{1}{n}\sum_{i=1}^{n}\left(x_j - y\right)^2 \tag{5-1}$$

was decomposed following [48] and [49] as:

$$\mathrm{MSD} = \sigma_x^2\left(IA + IB + IC\right) \tag{5-2}$$

where σ_x^2 is the variance of the grid cell values and *IA*, *IB* and *IC* are three additive non-dimensional components which describe the MSD. The variable *y* denotes the satellite scale value of the variable *x* at the j^{th} fine scale grid cell. *IA* is expressed as:

$$IA = 1 - R^2 \tag{5-3}$$

where *R* is the correlation coefficient. *IA* is a measure of the differences in covariability or timing and variance differences between the fine and satellite resolution values. The second term,

$$IB = \left(\frac{\sigma_y}{\sigma_x} - R \right)^2 \tag{6-4}$$

is a non-dimensional measure of the covariability between the two data sets termed the conditional bias [49]. It is an indication of the multiplicative bias between the satellite and fine scale values. The final term,

$$IC = \frac{B^2}{\sigma_x^2} \tag{5-5}$$

quantifies the impact of the unconditional bias, commonly referred to simply as bias, where *B* is the bias.

The correction of biases between the SMOS scale and MESH-CMEM grid scale was carried out using the cumulative distribution function (CDF) matching technique of [50]. The data were sorted and the differences fit with a third order polynomial model which is applied to the satellite resolution data. The advantage of this method is that it matches high order statistical moments as opposed to only the first two moments which are matched with simpler techniques [51].

Besides the MSD, the following comparative metrics were used to quantify the global differences between SMOS and MESH-CMEM at both scales. The root mean squared difference (RMSD) is the square root of the MSD and is a global measure of the misfit between two sets of data scaled to data units. As demonstrated by Eq. (5-2), the MSD and its root are influenced by (unconditional) bias, the impact of which can be mitigated using the unbiased RMSD.

$$\text{ubRMSD} = \left(\text{RMSD}^2 - B^2 \right)^{1/2} \tag{5-6}$$

The final metric, is the well-known correlation coefficient:

$$R = \frac{\text{E}\left(x_j - \overline{x} \right)\left(y - \overline{y} \right)}{\sigma_x \sigma_y} \tag{5-7}$$

where E is the expectation operator and the over bar denotes the temporal mean value.

5.5 Results

5.5.1 MESH Calibration and Validation

To limit the number of free parameters being calibrated in MESH at any given time, a two-step approach was adopted. Individual sites of the McMaster Mesonet were selected based on their representativeness of the land cover types represented by CLASS; which is the core of the MESH model. Following the GRU concept, model parameters were calibrated to find a single set of parameters for each generic land cover type which produced the lowest RMSE of rootzone soil moisture at each site, using forcing data from 2010. Figure 5-2 shows the point scale validation of rootzone soil moisture during the year 2011 for each CLASS/MESH land cover type. The RMSE of rootzone soil moisture were approximately 0.02 m^3m^{-3} and 0.04 m^3m^{-3} during the calibration and validation periods respectively using the CLASS model.

In step-two, the parameters from step-one were held constant and MESH was calibrated by maximizing the Nash-Sutcliffe efficiency (NSE) at five stream gauges. The resulting NSE values were 0.62, 0.12, 0.40, 0.20 and 0.43 for the calibration period. This was lower than previous studies [27]; however, the sub-basins of the stream gauges are also considerably smaller and have less flow than prior MESH applications and the Niagara Escarpment prominant in the model domain. This makes direct comparison difficult.

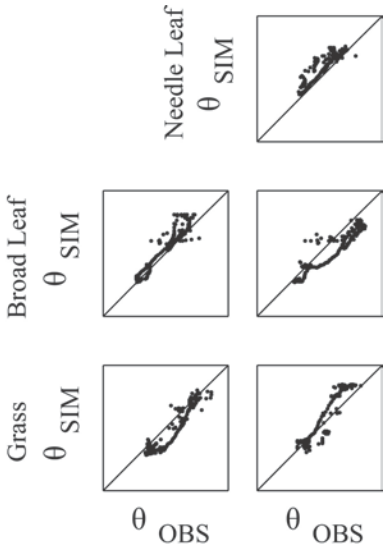

Figure 5-2: Scatter plots of point scale rootzone soil moisture for CLASS calibration of the needle leaf, broad leaf and grass GRUs. When soil moisture observations were available for multiple sites within a GRU the same parameter set was applied at both sites.

Since the soil moisture calibration in step-one was a point calibration procedure, MESH can also be spatially validated by comparing the grid cell soil moisture with the mean soil moisture from all McMaster Mesonet sites found within a particular grid cell as shown in Fig. 5-3. There was a slight depth discrepancy where model output from the first layer representing 0-10 cm whereas the McMaster Mesonet observations are at a depth of 10 cm. The expected result from this discrepancy is a slight dry bias in the model output, which was found with the CLASS

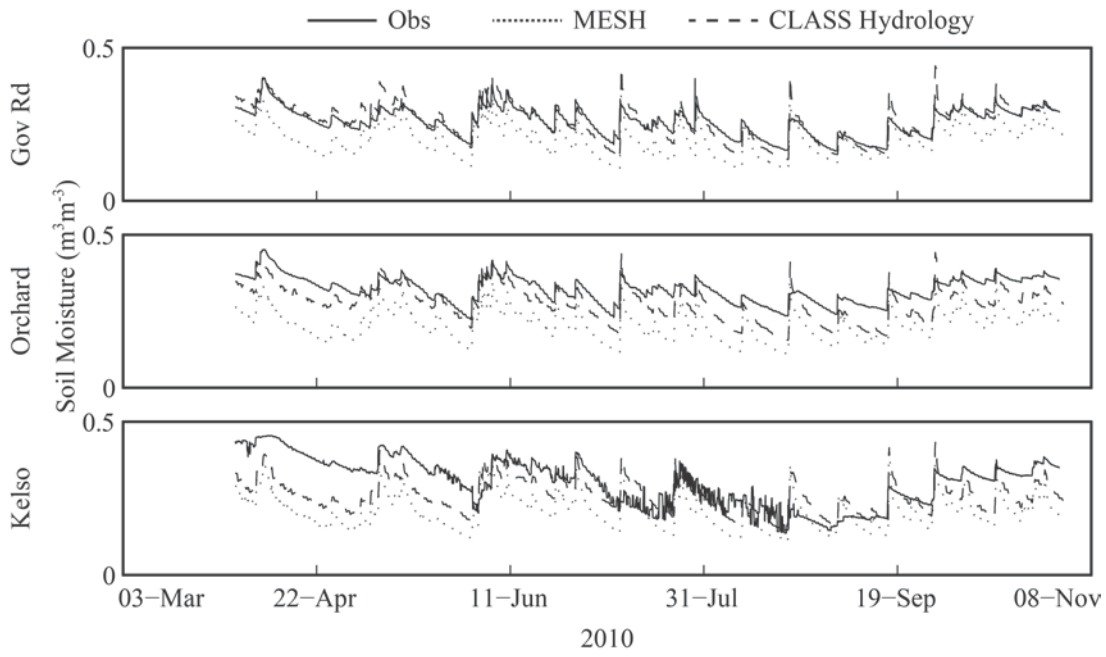

Figure 5-3: Grid cell soil moisture validation of the 0-10cm MESH layer and the 10cm observations at the three McMaster Mesonet sites. The McMaster Mesonet observations consist of the mean of the nine individual stations within each grid cell. MESH was validated using sloped grid cells and 2D interflow in the surface layer (MESH) and flat grid cells without interflow (CLASS Hydrology).

output. It was noted that MESH soil moisture in the upper soil layer had a large dry bias at each of the sites. In order to diagnose the source of this bias, the lateral flow of soil moisture in MESH was switched off; resulting in the model being representative of CLASS with hydrologic routing [25]. The result was a significant reduction in the bias and decrease in RMSE when comparing grid cell soil moisture to the average soil moisture of all McMaster Mesonet sites within a particular grid cell. As a consequence, the loss of lateral soil water movement decreased the quality of streamflow simulation, similar to [25]. These findings were believed to result from the two step calibration procedure and demonstrate the importance of multi-objective optimization for calibration of HLSS's for proper water cycle partitioning. The cost/benefits of such an approach are reserved for future research and are beyond the scope of this paper.

Table 5-3: Comparison of performance of MESH and CLASS-like operation of MESH, using model grid cell 0-10cm soil moisture and spatial mean 10cm observations form the McMaster Mesonet.

Site	RMSE	R	Bias
$GovRd_{MESH}$	0.07	0.80	-0.06
$GovRd_{CLASS}$	0.03	0.86	0.01
$Orchard_{MESH}$	0.12	0.80	-0.11
$Orchard_{CLASS}$	0.05	0.88	-0.05
$Kelso_{MESH}$	0.10	0.60	-0.07
$Kelso_{CLASS}$	0.06	0.67	-0.01

Since the primary purpose of this research was the simulation of soil moisture and TB, results were used from the CLASS-like operation of MESH, which will hereafter be referred to as MESH, since this was found to produce better soil moisture simulations for the study period as seen in Table 5-3.

5.5.2 MESH-CMEM Validation

To validate the MESH-CMEM coupling, simulations from MESH-CMEM in the study area were spatially aggregated using an arithmetic mean for each time step for comparison with SMOS TB observations. The time-series representation of TB in Fig. 5-4, demonstrates that there is good agreement between the temporal pattern and relative magnitude of TB between MESH-CMEM and

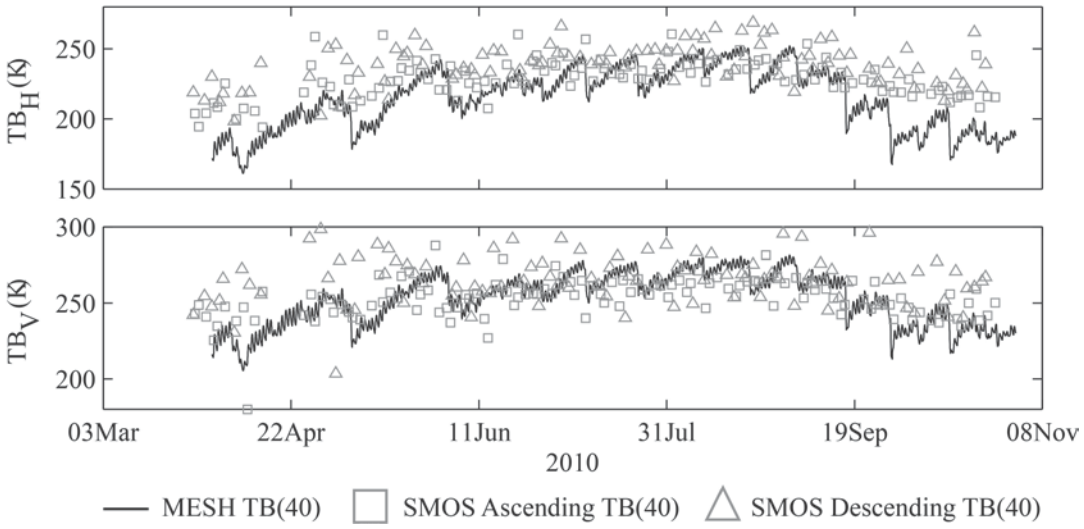

Figure 5-4: Temporal validation of MESH-CMEM TB at 40° incidence angle averaged over the entire study area using SMOS L1C observations.

SMOS. Observations from SMOS ascending orbits had greater consistency with MESH-CMEM simulations for both polarizations and were less biased. There was a discrepancy between modelled TB and SMOS observations during April and October of the study period, where modelled TB was consistently lower than SMOS observations. The difference may result from a spring/autumn discrepancy in soil moisture simulated by MESH on clay-loam soils and observed soil moisture, represented by the Kelso site in Fig. 5-3. Similarly, MESH determines vegetation growth using periods with temperatures above a vegetation growth threshold [41], resulting in the late leaf budding when compared to leaf area-index observations from MODIS Fig. 5-5.

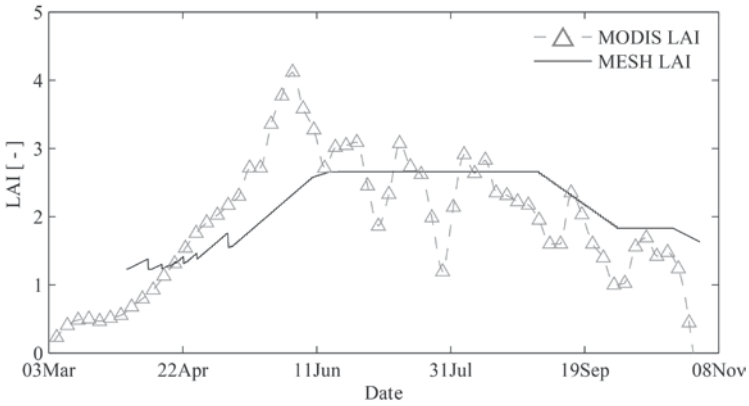

Figure 5-5: MESH LAI, which increases based on growing degree days, compared to MODIS LAI retrieved from Aqua and Terra (MOD15A3).

These factors combined would result in the under-estimation of TB by MESH-CMEM by under-estimating the soil moisture contribution to TB and the direct vegetation contribution to TB.

The scatterplots in Fig. 5-6 reveal that there was good agreement between MESH-CMEM and SMOS observations, particularly for the ascending overpasses (R≈0.54) which is consistent with the findings of [21] and [52] prior to calibration of radiative transfer model parameters. The regression lines in Fig. 6-6 show that both overpasses are relatively consistent for TB_H, whereas TB_V observations from the ascending overpass were more consistent with modelled TB_V (R≈0.5) than TB_V observations from the descending overpass (R≈0.3). Rowlandson et al. [53] compared SMOS morning and evening overpasses and suggest a number of reasons for differences seen, including polarization dependent directional RFI, the presence of dew on the canopy in the morning, differences in soil moisture and diurnal changes in vegetation optical depth based on

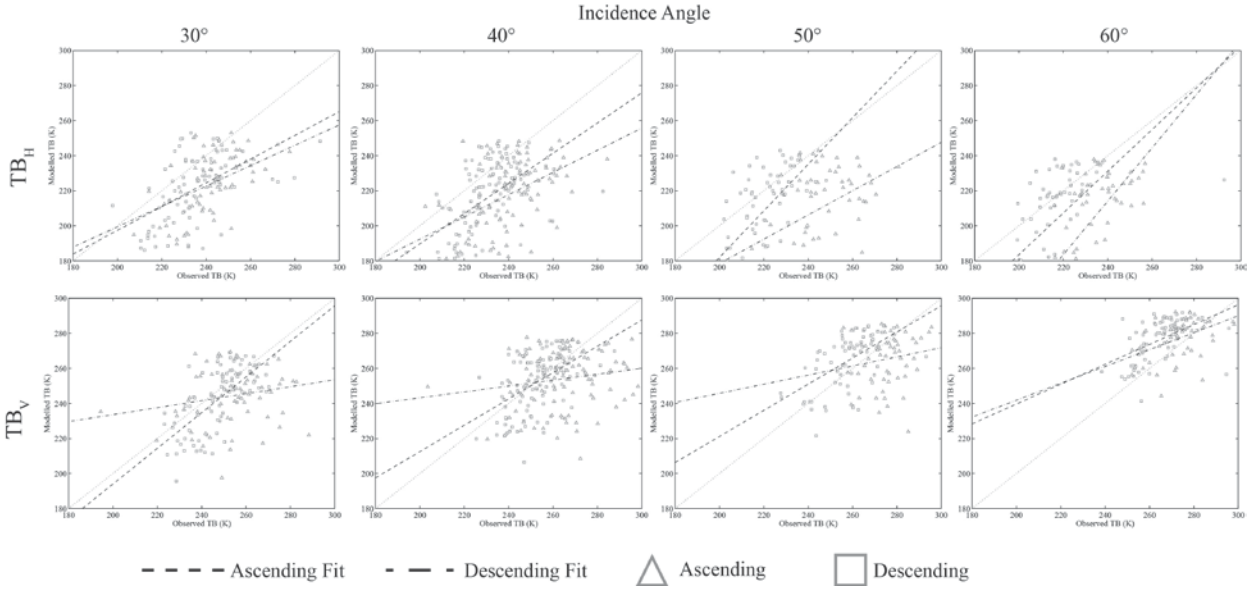

Figure 5-6: Scatter plots with best fit lines of concurrent MESH-CMEM TB simulations and SMOS L1C observations. For comparison SMOS data were divided between ascending and descending overpasses.

plant water loss during the daytime. It was noted that the relation between plant tissue water content and TB emission is understood theoretically, but has not been well quantified [53]. Given that the discrepancy between ascending and descending overpasses is greater at V polarization than H polarization and previous findings that TB_V is more influenced by vegetation than TB_H [36]; the impact of diurnal vegetation water content may be a likely candidate for the discrepancy. While canopy water content is modelled by MESH, it is not represented explicitly in CMEM but would be observed by SMOS. Regardless of the cause of this discrepancy, which cannot be confidently determined with available data, it is clear that for the purposes of bias correction; ascending and descending overpasses must be separately considered as each have different biases [12], [13], [53].

MESH-CMEM was used to simulate TB at incidence angles from 20° to 60° in increments of 10° for comparison with angular signatures observed by SMOS. TB from both SMOS and MESH-CMEM were fit by quadratic function and a comparison made anytime greater than 15 incidence angles were available from the SMOS observations. A subset of these comparisons can be seen in Fig. 5-7 for ascending and descending overpasses that were close in time. For all SMOS angular signatures compared to MESH-CMEM see the supplementary material. MESH-CMEM produced multi-angular TB simulations which were consistent with

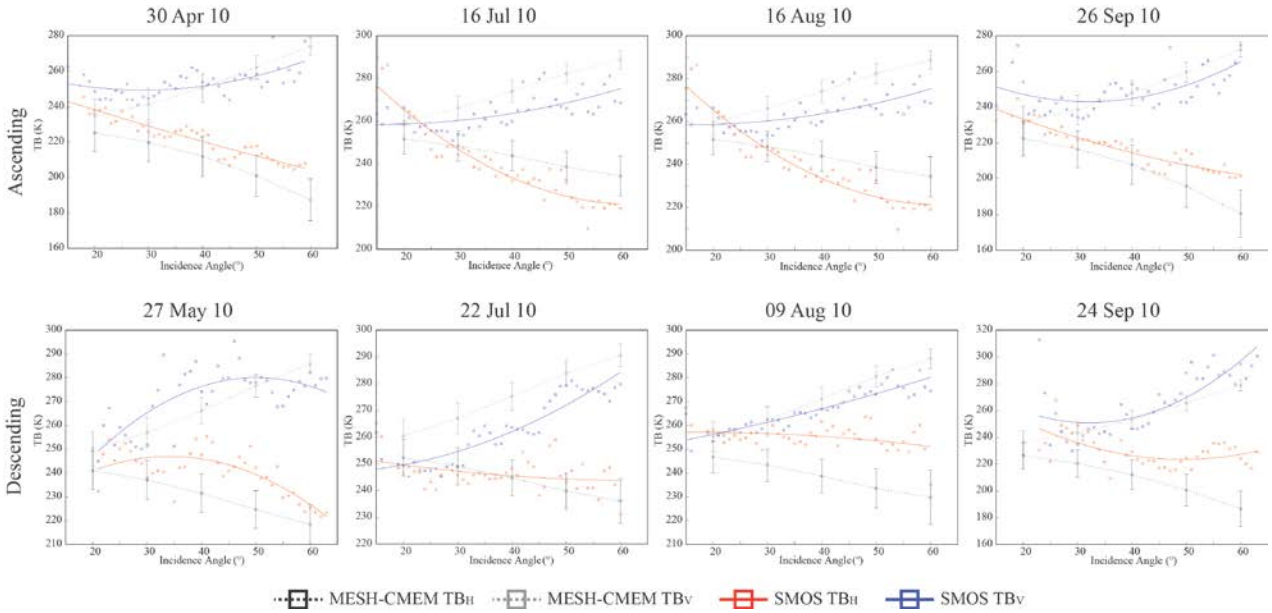

Figure 5-7: Angular TB profiles on selected dates showing mean MESH-CMEM TB with 1σ error bars and SMOS observations. For a comparison of all concurrent MESH-CMEM and SMOS angular profiles see the supplementary material.

theoretical considerations. TB_H and TB_V shifted vertically with soil moisture and the angle between TB_H and TB_V increased with vegetation biomass [36]. The multi-angular TB profile was also found to be in good agreement with SMOS observations, particularly for ascending overpasses. For many descending overpasses, the angle between TB_H and TB_V was smaller for SMOS observations than MESH-CMEM simulations, and in some instances SMOS $TB_{H(V)}$ had little angular dependence for descending overpasses (See supplementary material). This was presumed to result from shortcomings in SMOS observations from influences such as RFI.

5.5.3 Mean Squared Difference Decomposition

The mean squared difference between the grid cell mean soil moisture / TB and that of individual grid cells is decomposed using Eq. (5-2) in Fig. 5-8. For both soil moisture and $TB_{H(V)}$, those grid cells with the largest RMSD had high sub-grid impervious surfaces, open water or wetlands fractions. The impervious surface fraction particularly contributed to the presence of timing and random errors components of the MSD for TB. CMEM does not consider urban areas in the model, which resulted in the MESH-CMEM coupling treating urban areas as bare soils for which evaporation occurs quickly and which hold little water. The result was that soil moisture in these regions during periods without rain was continually near the wilting point. This approach was slightly different than the SMOS L2 retrieval algorithm approach to dealing with urban areas, which assigns a default contribution of TB for urban areas [35]. With either approach, the result can be expected to be

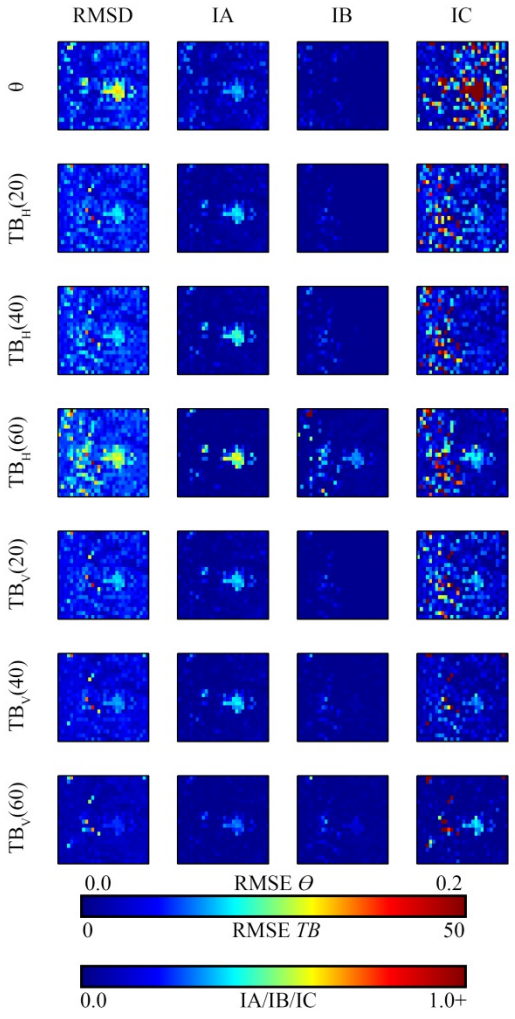

Figure 5-8: RMSD and decomposition of the MSD of MESH-CMEM study area mean and grid cell soil moisture and TB.

similar, where urban areas have little temporal correlation with DGG-scale values and are strongly biased. RMSD and *IC* were also found to be high in forested areas. Since the vegetation optical depth at nadir for H polarization for low vegetation and both polarizations for forests are modelled as constant [36] differences in *IC* likely results from the impact of greater vegetation biomass removing water from the soil. The result was a positive relationship between unconditional bias and sub-grid forest cover fractions for cells that had greater than 25% forest coverage at TB_H at all angles and TB_V at 20°-40° ($R \approx 0.6$, $p < 0.01$).

The decomposed MSD demonstrates that a large proportion of the error was attributed to unconditional bias at all incidence angles [24]. An underlying assumption of using bias correction for disaggregating soil moisture and TB, is that the satellite-scale temporal observations are representative of biased grid-scale dynamics. Kornelsen et al. [24] demonstrated this assumption was valid for relatively homogenous low vegetation semi-arid regions. In this study, with the exception of urban and wetland areas, the random/timing errors and conditional biases are small and were distributed randomly across the study domain. Therefore, even in a heterogeneous environment, the assumption that spatial differences are driven largely by biases was valid.

5.5.4 Bias Correction

Given that observations from SMOS ascending overpasses were found to be more consistent with modelled TB (Fig. 5-6, 5-7) only the ascending overpasses were considered for further analysis. Also, to avoid unfairly skewing statistical analysis, grid cells that had a fraction of urban areas or open water that was greater than 0.5 were removed from the analysis. The filtering removed 45 grid cells out of a potential 660 grid cells. The density-scatter plots in Fig. 5-9 show the relationship between individual MESH grid cells and the mean modelled state, raw SMOS observations and the mean MESH state and SMOS observations whose CDF's were matched to the climatology of individual grid cells. Since all of the grid cells in the model domain were considered in Fig. 5-9, this and the resulting performance statistics in Table 5-4 provide an indication of the global performance of the bias correction technique.

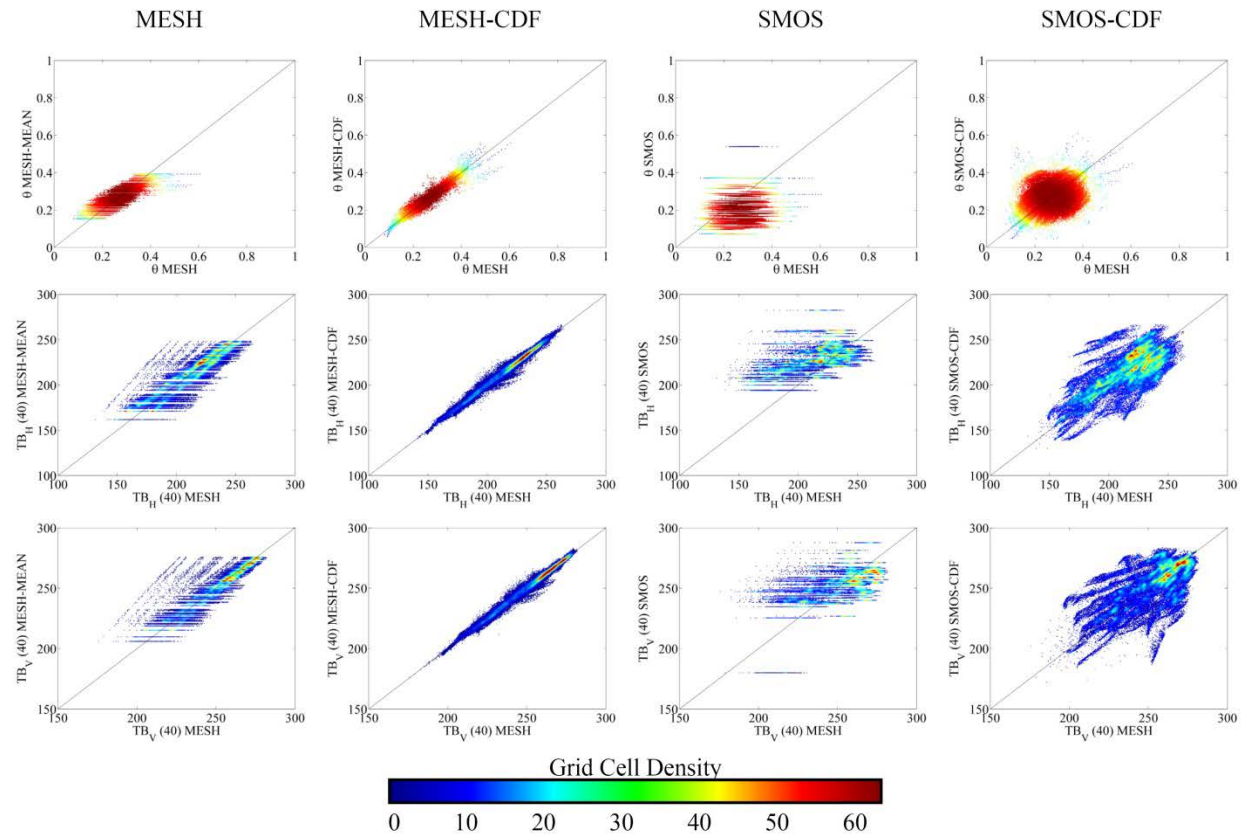

Figure 5-9: Density-scatter plots of satellite scale and MESH-CMEM grid scale data for all 615 grid cells (urban areas and open water removed) for the study period of Apr 1 – Oct 31 2010. High density indicates more individual points in the scatter plot at a particular XY co-ordinate.

Table 5-4: Validation of MESH-CMEM using SMOS observations.

Variable	R		RMSD (m^3m^{-3}; K)		N		ubRMSD (m^3m^{-3}; K)	
	ASC	DSC	ASC	DSC	ASC	DSC	ASC	DSC
Soil Moisture	0.19	0.12	0.11	0.11	109	82	0.08	0.1
$TB_H(20)$	0.42	0.69	28.7	19.3	50	42	18.6	13.2
$TB_H(30)$	0.56	0.57	22.2	24.1	80	69	17.3	16.8
$TB_H(40)$	0.57	0.54	22.5	28.5	114	101	17.3	15.8
$TB_H(50)$	0.64	0.45	22.9	41.7	68	58	19.6	21.9
$TB_H(60)$	0.59	0.67	28.2	37.6	69	58	24.6	22.7
$TB_V(20)$	0.31	0.45	23.3	21.3	50	43	17.0	16.6
$TB_V(30)$	0.69	0.20	13.9	22.3	81	67	12.9	20.2
$TB_V(40)$	0.58	0.17	14.2	23.3	115	98	14.1	20.5
$TB_V(50)$	0.55	0.16	14.8	19.1	68	56	13.3	15.9
$TB_V(60)$	0.51	0.44	16.2	12.2	68	54	11.2	11.5

Table 5-5: Comparison of MESH-CMEM average and SMOS ascending overpass TB/soil moisture with MESH-CMEM grid scale simulations prior to and following CDF matching bias correction.

Variable	R				ubRMSD (m^3m^{-3}; K)			
	MESH	MESH CDF	SMOS	SMOS CDF	MESH	MESH CDF	SMOS	SMOS CDF
Soil Moisture	0.90	0.98	0.14	0.31	0.03	0.01	0.09	0.07
$TB_H(20)$	0.91	0.99	0.42	0.58	7.7	2.7	19.9	17.5
$TB_H(30)$	0.93	0.99	0.50	0.62	8.3	2.7	20.2	19.3
$TB_H(40)$	0.93	0.99	0.57	0.72	8.7	2.8	19.4	17.7
$TB_H(50)$	0.94	0.99	0.62	0.71	9.6	2.8	22.2	21.0
$TB_H(60)$	0.94	0.99	0.56	0.72	11.0	2.7	27.8	24.6
$TB_V(20)$	0.92	0.99	0.34	0.45	6.9	2.5	18.2	18.9
$TB_V(30)$	0.35	0.99	0.65	0.73	6.5	2.3	15.1	14.7
$TB_V(40)$	0.95	0.99	0.55	0.63	5.7	2.1	15.7	16.0
$TB_V(50)$	0.69	0.99	0.56	0.63	4.8	1.7	14.2	14.4
$TB_V(60)$	0.95	0.99	0.50	0.66	3.9	1.1	12.0	10.7

The horizontal banding in Fig. 5-9 subplots that compared MESH grid cells to the mean model state or raw SMOS observations was the result of multiple grid cells being represented by a single coarse resolution value. If a particular grid cell was well correlated with the coarse resolution value, but biased the result would be a series of points which were parallel the ideal 1:1 line and horizontally displaced. Matching the CDF of the mean MESH model state and SMOS observations to that of individual grid cells resulted in a global correlation greater than 0.99 for MESH and an improvement in SMOS correlation by approximately 0.17 for soil moisture and 0.12 for TB as seen in Table 5-5. The improvement in correlation with bias correction was a result of the two datasets sharing the same range, thereby matching the variance in the denominator of Eq. (5-7) and reducing the differences in the numerator. This was particularly true of the noisy SMOS observations. Comparing the global performance measures in Table 5-5 to that of the SMOS product-scale performance in Table 5-4, shows that the high resolution bias corrected SMOS observations globally had similar performance to comparisons at coarse resolution. The variability of performance between SMOS observations after bias correction and individual model grid cells is shown in Fig. 5-10. For intermediate incidence angles, bias correction resulted in a significant improvement in both correlation and RMSE. The

large decrease in RMSE between the SMOS observations before and after bias correction was largely attributed to the reduction in bias, although there was a slight reduction in unbiased RMSE for most incidence angles.

5.6 Discussion and Conclusions

This study presented the coupling and validation of a land-surface hydrological model (MESH) [25], [26] with the Community Microwave Emission Model [21], [22]. Comparison with SMOS TB observations demonstrated the MESH-CMEM was able to well represent the temporal pattern and angular profile of top of the atmosphere TB for SMOS ascending overpasses, although differences were noted for descending overpasses, particularly for vertically polarized TB (Fig. 5-6).

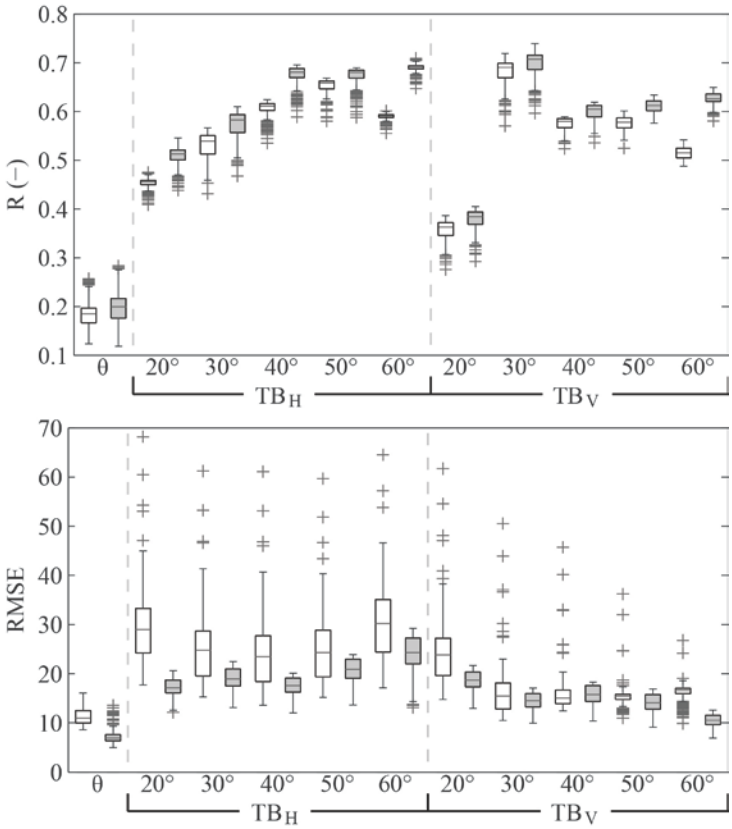

Figure 5-10: Performance of SMOS soil moisture retrievals and TB observations compared to individual MESH-CMEM grid cells prior to (white) and following (grey) bias correction. To fit the range of the figure, the soil moisture values were multiplied by 100. Therefore, soil moisture RMSE of 10m^{3}m^{-3} in the figure is actually 0.1m^{3}m^{-3}. The centre of the box is the median and the edges are the 25th and 75th percentiles. The whiskers extend 2.7σ and the crosses are considered outliers.

A surprising result was the difference in performance, particularly in terms of correlation, between SMOS L2 soil moisture and SMOS L1C TB. It was initially expected that soil moisture retrievals would have similar performance to TB observations, particularly after bias correction, since the CMEM model is similar in formulation to the SMOS retrieval algorithm [21], [35], [36] and MESH soil moisture was similar to in situ observations (Fig. 5-2 & 5-3). It is possible that the lower performance of the SMOS retrieval algorithm compared to MESH was the result of errors in the MESH model itself. Modelled soil moisture may have been representative at the individual sites of Fig. 5-3, but due to calibration or structural model errors was considerably different at other grid cells. Since soil moisture away from these sites is unknown, it is possible

that SMOS is more representative of the true soil moisture state than the model. However, this cause is unlikely given the similarity in MESH-CMEM TB to SMOS observations.

A more plausible explanation is that the difference was induced by the method of soil moisture retrieval. The SMOS L2 processor searches for a set of parameters that minimizes the difference between the observed TB, via a fitted curve, and the simulated TB, similar to Fig. 5-7. During many of the overpasses in this study, SMOS TB_V was less sensitive to incidence angle (flat fitted curve) than MESH-CMEM and TB at both polarizations was often biased. Since the retrieved soil moisture is influenced by TB at all incidence angles and both polarizations [35] systematic differences in TB during a particular overpass could result in less consistency in the soil moisture retrievals. In contrast, TB at a given incidence angle and polarization are not impacted by inter-dependencies and differences could be more reliably corrected. Since the differences in angular profiles of TB between MESH-CMEM and SMOS were not consistent, that is biases and slope changed frequently, it is difficult to specify an underlying cause. These differences may be the result of external noise, such as RFI, structural deficits in L-MEB which are yet to be described [53], lower soil moisture signal due to dense vegetation during the peak growing season [14], [54], [55] or scaling issues [56]. Underlying causes will be more easily identified with data that have higher radiometric accuracy [57] or greater agreement with theoretical expectations [58].

Since the correction of TB has been demonstrated herein to be more reliable than soil moisture, this study supports other evidence and theoretical/practical considerations that suggest modelling approaches that integrate information from satellites such as SMOS are recommended to use TB observations rather than soil moisture retrievals (i.e. [7], [59]-[61]). The other considerations include latency of soil moisture retrievals and the inconsistency between auxilary data sources in retrieval products and the current model application [59]-[61]. The latter reason is the likely cause of higher correlation and lower relative error of TB compared to retrieved soil moisture in this study.

This study also expanded on the previous research of [24] by demonstrating the validity of the bias correction approach for downscaling satellite observations of TB and retrieved soil moisture. Particularly, [24] demonstrated the concept in relatively homogenous and semi-arid watersheds with relatively few point scale measurements. In contrast, the method was examined herein, using a distributed model, in a mid-latitude watershed with imperfectly drained soil and

mixed land-use, soil types and topography as the sub-SMOS grid scale. Using the theoretical basis of the temporal stability concept [62], results further demonstrate that spatial soil moisture variability can be generally explained by differences in temporal mean and variance, between watershed (SMOS) scale and fine scale soil moisture. Correcting the bias between SMOS L1C TB observations and individual grid cells with a resolution of approximately 800 m resulted in comparable performance to correcting SMOS TB to modelled TB at scales similar to SMOS data products.

The primary motivation for the use of bias correction to disaggregate passive microwave observations and soil moisture retrievals is the relative simplicity of the method and the existing requirement for bias correction in data assimilation applications [24]. It should be noted that while the disaggregation method is easy to implement, in contrast to other disaggregation methods, the requirement of pre-determined land surface conditions (i.e. from a land surface model) is complex and potentially time consuming. This requirement also provides the largest theoretical obstacle to the implementation of the method, in that it must be assumed that the HLSS in use is actually representative of the spatial distribution of soil moisture and errors in modelled soil moisture primarily result from shortcomings in precipitation, temperature, etc. that are used to drive the model. Given a well calibrated model, the assumption that spatial differences can be accounted for by bias is well justified by the abundance of literature supporting the concept of temporal stability at multiple scales and with different data types [34], [62]-[64]. This assumption was also well supported by the implementation of the distributed HLSS herein where the correlation between soil moisture and TB at all grid cells were biased by highly correlated (Fig. 5-9).

Nevertheless, some studies have demonstrated that certain basins do not demonstrate temporal stability [65] and the use of remote sensing only techniques [15], [19], [66], [67] may uncover structural deficiencies in the model which cannot be determined with bias correction. It is expected that the approach demonstrated herein, will be complimentary to remote sensing based disaggregation approaches, since remote sensing approaches are independent of model errors, but are biased by satellite states and errors [15] and the bias correction technique is dependent on model state space but is subject to less noise (Fig. 5-9).

Holding the assumptions of temporal persistence to be true, the results of this study indicate that scale poses less of a barrier to soil moisture or TB assimilation than the uncertainty inherent in satellite observations or our current operational modelling of microwave emissions.

5.7 References

[1] H. Moradkhani, "Hydrologic remote sensing and land surface data assimilation," Sensors, vol. 8, pp. 2986-3004, May 2008.

[2] J. Komma, G. Blöschl, and C. Reszler, "Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting," J. Hydrol. vol. 357, pp. 228-242, May 2008.

[3] O. Merlin, A. Chehbouni, G. Boulet, and Y. Kerr, "Assimilation of disaggregated microwave soil moisture into a hydrological model using coarse-scale meteorological data," J. Hydromet. vol. 7, pp. 1308-1322, Jan. 2006.

[4] A.L. Barbu, J.-C. Calvet, J.-F. Mahfouf, and S. Lafont, "Integrating ASCAT surface soil moisture and GEOV1 leaf area index into the SURFEX modelling platform: a land data assimilation application over France," Hydrol. Earth Syst. Sci., vol. 18, pp. 173-192, Jan. 2014.

[5] G.C. Heathman, P.J. Starks, L.R. Ahuja, and T.J. Jackson, "Assimilation of surface soil moisture to estimate profile soil water content," J. Hydrol., vol. 279, pp. 1-17, Feb. 2003.

[6] P. de Rosnay, M. Drusch, D. Vasiljevic, G. Balsamo, C. Albergel, and L. Isaksen, "A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF," Q. J. R. Meteorol. Soc., vol. 139, pp. 119-1213, Jul. 2013.

[7] G. Balsamo, J.-F. Mahfouf, S. Bélair, and G. Deblonde, "A land data assimilation system for soil moisture and temperature: An information content study," J. Hydromet., vol. 8, pp. 1225-1242, Dec. 2007.

[8] Z. Bartalis, W. Wagner, V. Naeimi, S. Hasenauer, K. Scipal, H. Bonekamp, J. Figa, and C. Anderson, "Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT)," Geophys. Res. Lett., vol. 34, L20401, Oct. 2007.

[9] Y.H. Kerr, P. Waldteufel, J.-P. Wigneron, S. Delwart, F. Cabot, J. Boutin, M.J. Escorihuela, J. Font, N. Reul, C. Gruhier, S. E. Juglea, M.R. Drinkwater, A. Hahne, M. Martin-Neira, and S. Mecklenburg, "The SMOS mission: New tool for monitoring key elements of the global water cycle," Proc. IEEE, vol. 98, no. 5, pp. 666–687, May 2010.

[10] D. Entekhabi, E.G. Njoku, P.E. O'Neill, K.H. Kellogg, W.T. Crow, W.N. Edelstein, J.K. Entin, S.D. Goodman, T.J. Jackson, J. Johnson, J. Kimball, J.R. Piepmeier, R.D. Koster, N. Martin, K.C. McDonald, M. Moghaddam, S. Moran, R. Reichle, J.C. Shi, M.W. Spencer, S.W.

Thurman, L. Tsang, and J. van Zyl, "The Soil Moisture Active Passive (SMAP) mission," Proc. IEEE, vol. 98, no. 5, pp. 704–716, May 2010.

[11] A. Al Bitar, D. Leroux, Y. H. Kerr, O. Merlin, P. Richaume, A. Sahoo, and E. F. Wood, "Evaluation of SMOS soil moisture
products over continental U.S. using the SCAN/SNOTEL network," IEEE Trans. Geosci. Remote Sens., vol. 50, no. 5, pp. 1572-1586, May 2012.

[12] T. W. Collow, A. Robock, J. B. Basara, and B. G. Illston, "Evaluation of SMOS retrievals of soil moisture over the central United States with currently available in situ observations," J. Geophys. Res., vol. 117, no. D9, pp. D09113, May 2012.

[13] T. J. Jackson, R. Bindlish, M. H. Cosh, T. Zhao, P. J. Starks, D. D. Bosch, M. Seyfried, M. S. Moran, D. C. Goodrich, Y. H. Kerr, and D. Leroux, "Validation of Soil Moisture and Ocean Salinity (SMOS) soil moisture over watershed networks in the U.S.," IEEE Trans. Geosci. Remote Sens., vol. 50, no. 5, pp. 1530-1543, May 2012.

[14] M. Pan, A.K. Sahoo, E.F. Wood, A. Al Bitar, D. Leroux, and Y.H. Kerr, "An initial assessment of SMOS derived soil moisture over the Continental United States," IEEE J. Sel. Topics Appl. Earth Observ. vol. 5, no. 5, pp. 1448-1457, Oct. 2012.

[15] O. Merlin, C. Rüdiger, A. Al Bitar, P. Richaume, J.P. Walker, and Y.H. Kerr, "Disaggregation of SMOS soil moisture in Southeastern Australia," IEEE Trans. Geosci. Remote Sens., vol. 50, no. 5, pp. 1556-1571, May 2012.

[16] O. Merlin, J.P. Walker, A. Chehbouni, and Y. Kerr, "Towards deterministic downscaling of SMOS soil moisture using MODIS derived soil evaporative efficiency," Remote Sens. Environ., vol. 112, no. 10, pp. 3935-3946, Oct. 2008.

[17] M. Piles, N. Sánchez, M. Vall-Ilossera, A. Camps, J. Martínez-Fernández, J. Martínez, and V. González-Gambau, "A downscaling approach for SMOS land observations: Evaluation of high-resolution soil moisture maps over the Iberian Peninsula," IEEE J. Sel. Topics Appl. Earth Observ. vol. 7, no. 9, pp. 3845-3857, Sep. 2014.

[18] G. Mascaro, E.R. Vivoni, and R. Deidda "Soil moisture downscaling across climate regions and its emergent properties," J. Geophys. Res. 116, D22114, Nov. 2011.

[19] D.J. Wilson, A.W. Western, and R.B. Grayson, "A terrain and data-based method for generating the spatial distribution of soil moisture," Adv. Water Resour., vol. 28, no. 1, pp.43-54, Jan. 2005.

[20] S. Chakrabarti, T. Bongiovanni, J. Judge, L. Zotarelli, and C. Bayer, "Assimilation of SMOS soil moisture for quantifying drought impacts on crop yield in agricultural regions," IEEE J. Sel. Topics Appl. Earth Observ. vol. 7, no. 9, pp. 3867-3879, Sep. 2014.

[21] M. Drusch, T. Holmes, P. de Rosnay, and G. Balsamo, "Comparing ERA-40 based L-Band brightness temperatures with Skylab observations: A calibration/validation study using the Community Microwave Emission Model," J. Hydromet.,vol. 10, pp. 213-226, Jun. 2009.

[22] P. de Rosnay, M. Drusch, A. Boone, G. Balsamo, B. Decharme, P. Harris, Y. Kerr, T. Pellarin, J. Polcher, and J.-P. Wigneron, "AMMA land surface model intercomparison experiment coupled to the Community Microwave Emission Model: ALMIP-MEM," J. Geophys. Res., vol. 114, pp. D05108, Mar. 2009.

[23] T. J. Jackson, M. Cosh, R. Bindlish, P. Starks, D. Bosch, M. Seyfried, D. Goodrich, S. Moran, and J. Du, "Validation of advanced Microwave Scanning Radiometer soil moisture products," IEEE Trans. Geosci. Remote Sens., vol. 48, no. 12, pp. 4256–4272, Dec. 2010.

[24] K.C. Kornelsen, M. Cosh, and P. Coulibaly, "Potential of bias correction for downscaling satellite observations," submitted to J. Geophys. Res., 2014.

[25] A. Pietroniro, V. Fortin, N. Kouwen, C. Neal, R. Turcotte, B. Davison, D. Verseghy, E.D. Soulis, R. Caldwell, N. Evora, and P. Pellerin, "Development of the MESH modelling system for hydrological ensemble forecasting of the Laurentian Great Lakes at the regional scale," Hydrol. Earth Syst. Sci., vol. 11, pp. 1279-1294, May 2007.

[26] E.D. Soulis, K.R. Snelgrove, N. Kouwen, F. Seglenieks, and D.L. Verseghy, "Towards closing the vertical water balance in Canadian Atmospheric Models: Coupling of the land surface scheme CLASS with the distributed hydrological model WATFLOOD," Atmos. Ocean, vol. 38, no. 1, pp. 251-269, Jan. 2000.

[27] A. Haghnegahdar, B.A. Tolson, B. Davison, F.R. Seglenieks, E. Klyszejko, E.D. Soulis, V. Fotrin, and L.S. Matott, " Calibrating Environment Canada's MESH Modelling System over the Great Lakes Basin," Atmos. Ocean, vol. 52, no. 4, pp. 281-293, Aug. 2014.

[28] P. Pellerin, A. Pietroniro, S. Bélair, V. Fortin, D. Charpentier, B. Bilodeau, I. Doré, J. Töyrä, M. Carrera, B. Davison, B. Toth, and S. Marin, "The development of water-balance indicators in an agriculture-dominated watershed using coupled hydrometeorological modeling and data assimilation," Gatineau, Canada, National Agri- Environmental Standards Initiative Synthesis Report 16, Environment Canada.

[29] E.D. Soulis, and F.R. Seglenieks, "The MAGS integrated modeling system, in: Woo, M.K. (Ed.), Cold Regions Atmospheric and Hydrologic Studies: The Mackenzie GEWEX Experience, vol. II. London, United Kingdom: Springer-Verlag, 2007, pp. 445–474.

[30] P. Dornes, B.A. Tolson, B. Davison, A. Pietroniro, J.W. Pomeroy, and P. Marsh, "Regionalisation of land surface hydrological model parameters in subarctic and arctic environments," Phys. Chem. Earth, vol. 33, no. 17-18, pp. 1081-1089, Dec. 2008.

[31] B. Davison, S. Pohl, P. Domes, P. Marsh, A. Pietroniro, and M. MacKay, "Characterizing snowmelt variability in a land-surface-hydrologic model," Atmos. Ocean, vol. 44, no. 3, pp. 271–287, Nov. 2006.

[32] S. Pohl, B. Davison, P. Marsh, and A. Pietroniro, "Modelling spatially distributed snowmelt and meltwater runoff in a small Arctic catchment with a hydrology –land surface scheme (WATCLASS)," Atmos. Ocean, vol. 43, no. 3, pp. 193–211, Nov. 2005.

[33] G. Dumedah, J. P. Walker, and C. Rüdiger, "Can SMOS data be used directly on the 15-km discrete global grid?" IEEE Trans. Geosci. Remote Sens., vol. 52, no. 5, pp. 2538-2544, May 2014.

[34] K.C. Kornelsen, and P. Coulibaly, "McMaster Mesonet soil moisture dataset: description and spatio-temporal variability analysis," Hydrol. Earth Syst. Sci., vol. 17, no. 1, pp. 1-18, Jan. 2013.

[35] Y.H. Kerr, P. Waldteufel, P. Richaume, J.-P. Wigneron, P. Ferrazzoli, A. Mahmoodi, A. Al Bitar, F. Cabot, C. Gruhier, S.E. Juglea, D. Leroux, A. Mialon, and S. Delwart, "The SMOS soil moisture retrieval algorithm," IEEE Trans. Geosci. Remote Sens., vol. 50, pp. 1384–1403, May 2012.

[36] J.-P. Wigneron, Y. Kerr, P. Waldteufel, K. Saleh, M.-J. Escorihuela, P. Richaume, P. Ferrazzoli, J. P. Grant, B. Hornbuckle, P. de Rosnay, J.-C. Calvet, T. Pellarin, R. Gurney, and C. Mätzler, "L-band microwave emission of the biosphere (L-MEB) model: Results from calibration against experimental data sets over crop fields," Remote Sens. Environ., vol. 107, no. 4, pp. 639–655, Apr. 2007.

[37] V.L. Mironov, M.C. Dobson, V.H. Kaupp, S.A. Komarov, and V.N. Kleschenko, "Generalized refractive mixing dielectric model for moist soils," IEEE Trans. Geosci. Remote Sens. vol. 42, no. 42, pp. 773-758, Apr. 2004.

[38] T. Pellarin, J. P. Wigneron, J. C. Calvet, M. Berger, H. Douville, P. Ferrazzoli, Y. H. Kerr, E. Lopez-Baeza, J. Pulliainen, L. P. Simmonds, and P. Waldteufel, "Two-year global simulation of L-band brightness temperatures over land," IEEE Trans. Geosci. Remote Sens., vol. 41, no. 9, pp. 2135–2139, Sep. 2003.

[39] T. R. H. Holmes, P. de Rosnay, R. de Jeu, J.-P. Wigneron, Y. H. Kerr, J.-C. Calvet, M. J. Escorihuela, K. Saleh, and F. Lemaitre, "A new parameterization of the effective temperature for L-band radiometry," Geophys. Res. Lett., vol. 33, p. L07 405, Apr. 2006.

[40] P. de Rosnay and J.-P. Wigneron, "Parameterizations of the effective temperature for L-band radiometry. Inter-comparison and long term validation with SMOSREX field experiment," in Radiative Transfer Models for Microwave Radiometry, C.Mätzler, Ed. Stevenage, U.K.: Inst. Elect. Eng., 2006.

[41] D.L. Verseghy, "CLASS- The Canadian Land Surface Scheme, Version 3.5, Technical Documentation," Climate Research Division, Science and Technology Branch, Environment Canada, 183 pp., 2011.

[42] N. Kouwen, E.D. Soulis, A. Pietroniro, J. Donald, and R.A. Harrington, "Grouped Response Units for Distributed Hydrologic Modeling," J. Water Resour. Planning Manage., vol. 119, no. 3, pp. 289–305, Jun. 1993.

[43] J. Côté, S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, "The operational CMC-MRB Global Environmental Multiscale (GEM) Model. Part I: Design considerations and formulation," Mon. Weather Rev., vol. 126, no. 6, pp. 1373-1395, Jun. 1998.

[44] J. Côté, J.-G. Desmarais, S. Gravel, A. Méthot, and A. Patoine, "The operational CMC-MRB Global Environmental Multiscale (GEM) Model. Part II: Results," Mon. Weather Rev., vol. 126, no. 6, pp. 1397-1418, Jun. 1998.

[45] J.-F. Mahfouf, B. Brasnett, and S. Gagnon. "A Canadian Precipitation Analysis (CaPA) Project: Description and Preliminary Results," Atmos. Ocean, vol. 45, no. 1, pp. 1 – 17, Jan. 2007.

[46] B.A. Tolson, and C.A. Shoemaker, "Dynamically dimensioned search algorithm for computationally efficient watershed model calibration," Water Resour. Res., vol. 43, pp. W01413, Jan. 2007.

[47] L.S. Mattot, "An optimization software tool: Documentation and user's guide," Buffalo, NY, USA, State University of New York at Buffalo, 2007.

[48] A.H. Murphy, "Skill scores based on the mean square error and their relationships to the correlation coefficient," Mon. Weather Rev. vol. 116, no. 12, pp. 2417-2424, Dec. 1988.

[49] S. Weglarczyk, "The interdependence and applicability of some statistical quality measures for hydrological applications," J. Hydrol. vol. 206, no. 1, pp. 98-103, Jan. 1998.

[50] M. Drusch, E.F. Wood, and H. Gao, "Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture," Geophys. Res. Lett., vol. 32, pp. L15403, Aug. 2005.

[51] C.S. Draper, J.P. Walker, P.J. Steinle, R.A.M. de Jeu, R.A.M., and T.R.H. Holmes, "An evaluation of AMSR-E derived soil moisture over Australia," Remote Sens. Environ., vol. 113, no. 4, pp. 703-710, Apr. 2009.

[52] G.J.M. De Lannoy, R.H. Reichle, and V.R.N. Pauwels, "Global calibration of the GEOS-5 L-Band microwave radiative transfer model over nonfrozen land using SMOS observations," J. Hydromet. vol. 14, no. 3, pp. 765-785, Jun. 2013.

[53] T.L. Rowlandson, B.K. Hornbuckle, L.M. Bramer, J.C. Patton, and S.D. Logsdon, "Comparisons of evening and morning SMOS passes over the Midwest United States," IEEE Trans. Geosci. Remote Sens., vol. 50, no. 5, pp. 1544-1555, May 2012.

[54] R. Rahmoune, P. Ferrazzoli, Y.H. Kerr, and P. Richaume, "SMOS level 2 retrieval algorithm over forests: Description and generation of global maps," IEEE J. Sel. Topics Appl. Earth Observ. vol. 6, no. 3, pp. 1430-1439, Jun. 2013.

[55] R. Rahmoune, P. Ferrazzoli, Y.K. Singh, Y.H. Kerr, P. Richaume, and A. Al Bitar, "SMOS retrieval results over forests: Comparisons with independent measurements," IEEE J. Sel. Topics Appl. Earth Observ. vol. 7, no. 9, pp. 3858-3866, Sep. 2014.

[56] W.L. Crosson, A.S. Limaye, and C.A. Laymon, "Impacts of spatial scaling errors on soil moisture retrieval accuracy at L-band," IEEE J. Sel. Topics Appl. Earth Observ. vol. 3, no. 1, pp. 67-80, Mar. 2010.

[57] A.M. Zurita, I. Corbella, M. Martín-Neira, M.A. Plaza, F. Torres, and F.J. Benito, "Towards a SMOS operational mission: SMOSOps-Hexagonal," vol. 6, no. 3, pp. 1769-1780, Jun. 2013.

[58] T. Zhao, J. Shi, R. Bindlish, T.J. Jackson, Y.H. Kerr, M.H. Cosh, Q. Cui, Y. Li, C. Xiong, and T. Che, "Refinement of SMOS multiangular brightness temperature toward soil moisture retrieval and its analysis over reference targets," IEEE J. Sel. Topics Appl. Earth Observ. accepted.

[59] W.A. Lahoz, and G.J.M. De Lannoy, "Closing the gaps in our knowledge of the hydrological cycle over land: Conceptual problems," Surv. Geophys., vol. 35, no. 3, pp. 623-660, May 2014.

[60] M. Durand, E. Kim, and S.A. Margulis, "Radiance assimilation shows promise for snowpack characterization," Geophys. Res. Lett., vol. 367, no. 2, pp. L02503, Feb. 2009.

[61] W.T. Crow, and E.F. Wood, "The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using Ensemble Kalman filtering: as case study based on ESTAR measurements during SGP97," Adv. Water Resour., vol. 26, no. 1, pp. 137-149, Jan. 2003.

[62] G. Vachaud, A. Passerat De Silans, P. Balabanis, and M. Vauclin, "Temporal stability of spatially measured soil water probability density function," Soil Sci. Soc. Am. J. vol. 49, no.4 , pp. 822-828, Jul. 1985.

[63] L. Brocca, F. Melone, T. Moramarco, and R. Morbidelli, "Spatial-temporal variability of soil moisture and its estimation across scales," Water Resour. Res., vol. 46, no. 2, W02516, Feb. 2010.

[64] M.H. Cosh, T.J. Jackson, P. Starks, and G. Heathman, "Temporal stability of surface soil moisture in the Little Washita River watershed and its applications in satellite soil moisture product validation," J. Hydrol., vol. 323, no. 1, pp. 168-177, Jan. 2006.

[65] B. Mohanty and T.H. Skaggs, "Spatio-temporal evolution and time-stable characteristics of soil moisture within remote sensing footprints with varying soil, slope and vegetation," Adv. Water Resour., vol. 24, no. 9-10, pp. 1051-1067, Nov. 2001.

[66] J. Kim and T.S. Hogue, "Improving spatial soil moisture representation through integration of AMSR-E and MODIS products," IEEE Trans. Geosci. Remote Sens. vol. 50, no. 2, pp. 446-460, Feb. 2012.

[67] N.S. Chauhan, S. Miller, and P. Ardanuy, "Spaceborne soil moisture estimation at high resolution: A microwave-optical/IR synergistic approach," Int. J. Remote Sens., vol. 24, no. 22, pp. 4599–4622, Jul. 2003.

Chapter 6: Improved Bias Correction Methods for Satellite Soil Moisture Retrievals

Summary of Paper 5: Kornelsen, K.C. and Coulibaly, P. (Submitted) Improved Bias Correction Methods for Satllite Soil Moisture Retrieval, Remote Sensing of Environment, Manuscript Number: RSE-D-14-01088.

Summary:

The work of this final research paper focuses on the method of bias correction itself. SMOS soil moisture was compared to previously examined SCAN sites that are distributed throughout the United States for the years 2010 and 2011. A new approach was proposed for enhancing currently existing bias correction methods by bootstrap resampling the soil moisture data. The resampled data displays similar statistical properties (i.e. mean and variance) to the original datasets, but the temporal dependency between the SMOS and SCAN data is removed. The resampled data sets were used to find the parameters for the bias correction and the robustness of the bias correction was evaluated. It is expected that the results from this study will improve the error characterization in future data assimilation applications.

The key finds of this research were:

- Bias correction techniques, particularly CDF matching, could induce multiplicative bias in the residuals of the bias corrected data. The result was that bias correction induced systematic underestimation of values above the mean and systematic overestimation of values below the mean. This cause was the sorting and ranking used in CDF matching.
- Resampling the data sets removed the temporal dependency but had similar mean and variance. The resulting bias correction parameters could eliminate the bias in the mean and match the variance of the two datasets while significantly reducing the multiplicative bias. Data that were bias corrected with the new method had higher correlation at the cost of slightly higher RMSE.
- Differences in SCAN observations due, to differing hydrological conditions, or differences in SMOS retrievals between years resulted in bias correction parameters that were not applicable during different time periods for the same site.
- The results show that care must be taken in the selection of bias correction parameters to avoid erroneous data errors. The Kolmogorov-Smirnov test is recommended to ensure appropriate parameter selection and successful correction of biases.

6.1 Abstract

Soil moisture is a principle component of the Earth's climate and hydrological systems that is difficult to monitor and model due to high variability, uncertainty in land surface characterization and uncertainty in soil moisture forcing. Satellite soil moisture retrievals and brightness temperature observations, such those available from the Soil Moisture and Ocean Salinity (SMOS) mission, can be a valuable source of information for data assimilation and merging with other satellite retrieval datasets. To correct for biases in these data sets, bias correction methods such as cumulative distribution function (CDF) matching and linear rescaling are used to map satellite soil moisture climatology to that of *in situ* or model values. This study compared SMOS retrievals to soil moisture observations from the SCAN network for a calibration year 2010 and validation year 2011 before and after bias correction. Additive bias between SMOS retrievals and SCAN observations was removed by standard bias correction techniques and a new resampling approach was found to reduce multiplicative biases. The improved bias correction technique is promising for applications in data assimilation.

6.2 Introduction

Soil moisture is a principal component of the Earth's climate and hydrological systems. The state of moisture in the soil controls the hydrological and energy interactions between the atmosphere, vegetation and soil at the Earth's surface, which drives the balances of water and energy. For this reason the state of soil moisture is important for both scientific and operational applications such as numerical weather prediction (Drusch, 2007), flood forecasting (Bronstert et al. 2012; Berthet et al. 2009), and climate modelling (Seneviratne et al. 2013).

Monitoring and predicting soil moisture for scientific and operational purposes is a difficult task since *in situ* networks have poor spatial coverage and models suffer from errors in meteorological forcing, land surface characterization and simplifications of process descriptions. Recently, indirect measurements of soil moisture have become available through the use of active and passive microwave remote sensing on various platforms. Wagner et al. (1999) presented a change detection algorithm for the active microwave Advanced SCATterometer (ASCAT) on-board the MetOp (Meteorological Operational) satellite and several soil moisture

retrieval algorithms have been implemented for the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) on-board the National Aeronautics and Space Administration's (NASA) Aqua satellite (Owe et al. 2001; Owe et al. 2008; Njoku et al. 2003). On the 5th of November 2009, the European Space Agency (ESA) launched the Soil Moisture and Ocean Salinity (SMOS) mission (Kerr et al. 2010), the first dedicated soil moisture satellite, and in November 2014, NASA has scheduled the launch of the Soil Moisture Active Passive (SMAP) satellite (Entekhabi et al. 2010). Both SMOS and SMAP have L-band radiometers which are considered ideal for the monitoring of soil moisture as they penetrate both the atmosphere and thin vegetation and operate in a protected wavelength which should have minimal radio frequency interference (RFI) (Kerr et al. 2010; Entekhabi et al. 2010).

Validation of satellite soil moisture retrievals is an important step in understanding the quality of retrieval results and characterizing errors that may be present. With respect to SMOS validation, the quality of SMOS soil moisture retrievals have been evaluated using *in situ* networks in Spain (Sánchez et al. 2012), Germany (Dall' Amico et al. 2012), Denmark (Bircher et al. 2012), Italy and Luxembourg (Lacava et al. 2012), Canada (Gherboudj et al. 2012), the United States (Al Bitar et al. 2012; Collow et al. 2012; Jackson et al. 2012), and Australia (Su et al. 2013). Large scale evaluation of SMOS data has also been conducted by comparing SMOS soil moisture retrievals to soil moisture products from ASCAT and AMSR-E as well as land surface data assimilation system (LDAS) outputs (Leroux et al. 2013; Al-Yaari et al. 2014) and by determining the impact of SMOS retrieved soil moisture in a simple data assimilation system (Pan et al. 2012). Similar efforts have been made for soil moisture products from ASCAT (i.e. Brocca et al. 2011; Bartalis et al. 2007) and AMSR-E (i.e. Jackson et al. 2012; Brocca et al. 2011; Pan et al. 2012). Synthesis of these various results reveals that SMOS soil moisture products have good temporal correlation to observed or modelled soil moisture (Al Bitar et al. 2012; Jackson et al. 2012; Su et al. 2013; Lacava et al. 2012) and particularly outperform other soil moisture retrievals as vegetation density increases (Pan et al. 2012; Al-Yaari et al. 2014). Under nominal conditions, SMOS retrievals are close to meeting the target root mean squared error (RMSE) of $0.04 m^3 m^{-3}$ (Jackson et al. 2012; Al Bitar et al. 2012; Sánchez et al. 2012) although a persistent bias in SMOS, and other, soil moisture retrievals is a consistent issue that remains to be addressed (Sánchez et al. 2012; Jackson et al. 2012; Al Bitar et al. 2012; Su et al. 2013) and in some cases may be informative of algorithm performance (Jackson et al. 2012).

Biases and systematic differences between satellite retrieved soil moisture and the reference soil moisture are problematic for many applications such as LDAS and the blending of soil moisture products. For LDAS applications, several studies have shown that the assimilation of satellite data can improve the characterization of the surface states (Das et al. 2008; Draper et al. 2012; de Rosnay et al. 2013; Reichle and Koster, 2005; Albergel et al. 2010; Reichle et al. 2008), however, a fundamental assumption of the Kalman filter, and many of its derivatives, is that observation noise is mean zero Gaussian with a given covariance Rk. There is a similar requirement when blending several satellite soil moisture products to generate products representative of soil moisture climatology (Liu et al. 2011; Liu et al. 2012; Yilmaz et al. 2012). To correct for the presence of bias and variance errors Reichle and Koster (2004) and Drusch et al. (2005) proposed matching the cumulative distribution function (CDF) of observed satellite data to the model climatology as observation operators for the direct assimilation of satellite soil moisture. This technique has been adopted in many studies for comparison of soil moisture retrieval performances (Lacava et al. 2012; Su et al. 2013; Brocca et al. 2010), data assimilation (Crow and van den Berg, 2010; Draper et al. 2011) and for blending AMSR-E, ASCAT and other soil moisture products (Liu et al. 2011; Liu et al. 2012). Linear rescaling has also been used to correct the climatology of satellite soil moisture data where the mean and standard deviation of the satellite soil moisture are rescaled to match that of the *in situ* reference dataset (Draper et al. 2009; Brocca et al. 2010; Su et al. 2013). Both bias correction techniques assume that the reference dataset does not contain noise and errors in the retrieved soil moisture can be ignored in the bias correction technique. Not accounting for the possibility of errors in the data will be shown to result in undesirable conditions such as multiplicative bias in the corrected dataset. A property often neglected in cal/val campaigns.

The purpose of this study is to evaluate the presence of, and propose a correction for multiplicative bias in satellite retrieved and renormalized soil moisture. While the analysis will be presented using SMOS retrieved soil moisture over Soil Climate Analysis Network (SCAN) (Schaefer et al. 2007) sites in the continental U.S., it is expected the results are indicative of errors which may be present in other applications and with other sensors (Su et al. 2013). The SCAN sites chosen have been previously selected and validated in a node-site comparison by Al Bitar et al. (2012). To expand upon the previous validation and identify the types of error present, this study will renormalize the SMOS soil moisture to that of the concurrent SCAN

observations. An analysis of the residuals will be used to demonstrate the presence of multiplicative bias before and after the climatology of the SMOS data have been matched to that of *in situ* observations. An assessment of the robustness of bias correction methods will also be made by temporal cross-validation, where the correction parameters will be calibrated for retrievals made during calendar year 2010 and validated during 2011.

6.3 Study Areas and Soil Moisture Measurements

The SCAN network was designed by the Natural Resources Conservation Service (NRCS) to support natural resources assessments, conservation and water resources management within the U.S. In contrast to many networks which cover a limited spatial extent, SCAN sites have been placed at selected sites distributed around the United States to collect hourly atmospheric, soil moisture and soil temperature data in different climate, physiographic and soil regions (Schaefer et al. 2007). Soil moisture at each SCAN site is collected by Stevens Hydra Probes at depths of approximately 5, 10, 20, 50 and 100 cm (Schaefer et al. 2007), of which the ~5 cm depth is analyzed herein. After filtering for SCAN sites that had sensors at the 5 cm depth and were not located in areas of strong topography or dense vegetation Al Bitar et al. (2012) focused their attention on 13 SCAN sites and 4 SNOwpack TELemetry (SNOTEL) sites for SMOS validation and analysis, of which data from the SCAN sites are re-visited herein. The use of the same study sites allows

Table 6-1: Site Information

DGG Number	SCAN ID	State	Land Cover
172276	2160	UT	Grassland
186675	2018	WY	Grass
187758	2002	MN	Crop
203609	2093	KS	Crop
203626	2001	NE	Crop
218480	2092	KS	Crop
219434	2168	NM	Grassland
231355	2079	KY	Grassland
235420	2030	AR	Crop
237475	2084	MS	Crop
240043	2024	MS	Grassland
241088	2053	AR	Crop
5023541	2051	FL	Wetland

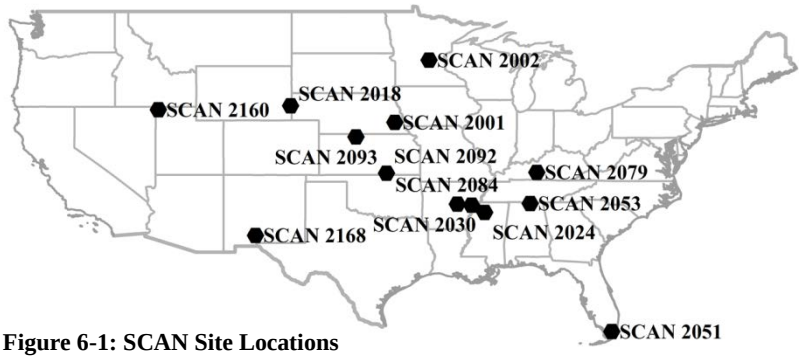

Figure 6-1: SCAN Site Locations

for insights gained from the previous work to be applied in this study. The distribution of the selected sites and some basic information about the sites can be found in Fig. 6-1, Table 6-1 and in Al Bitar et al. (2012).

All SCAN data were downloaded from the International Soil Moisture Network (ISMN; Dorigo et al. 2010) which is a data hosting facility for soil moisture data from various networks around the globe. In addition to quality control by the contributing network the ISMN performs basic quality control and homogenization of all its soil moisture datasets (Dorigo et al. 2010).

6.3.1 SMOS Soil Moisture Products

The SMOS Level 2 (L2) User Data Product (SML2UDP) provided by ESA from the second data reprocessing campaign was used in this study (ver. 5.51). Soil moisture, optical thickness and other geophysical variables are retrieved based on the inversion of the LMEB radiative transfer model (Kerr et al. 2012; Wigneron et al. 2007) by minimizing a cost function from multi-angular brightness temperature (TB) observations from the microwave imaging radiometer with aperture synthesis (MIRAS) (Kerr et al. 2012). SMOS soil moisture are provided on the (Icosahedral Snyder Equal Area Earth) ISEA 4H9 fixed grid with nodes equally spaced at 14.989 km (Kerr et al. 2010), although each SMOS grid is actually the weighted mean of an area with a radius of approximately 42 km around the node center (Al Bitar et al. 2012). The SMOS ascending and descending half-orbits coincide with approximately 0600 h and 1800 h local solar time (Kerr et al. 2010) where differences in the redistribution of water, heterogeneity of surface temperature caused by daytime heating and sensor orientation during the overpass all contribute to differences in accuracy of the retrieval between ascending and descending overpasses (Rowlandson et al. 2012; Collow et al. 2012; Jackson et al. 2012). Since the goal of this study is to assess the errors associated with bias correction and many researchers have reported a greater accuracy and stability of SMOS ascending overpasses (Jackson et al. 2012; Rowlandson et al. 2012), the SMOS data are separated and only ascending half-orbits are considered. Several indicators of quality are provided with SML2UDP including the data quality index, (DQX) which is a measure of the retrieved standard deviation reflecting radiometric uncertainty and information on the presence of radio frequency interference (RFI). Individual soil moisture retrievals were only considered for analysis if the soil moisture was successfully retrieved had a

DQX less than 0.1 m^3m^{-3} and the probability of radio frequency interference (RFI) in an individual retrieval was below 0.075.

A major difficulty with the validation of satellite soil moisture products is the significant mismatch between point scale *in situ* observations and areal averaged soil moisture retrievals. In previous studies this issue has been addressed using multiple *in situ* sensors within a grid footprint (Jackson et al. 2012; Collow et al. 2012; Su et al. 2013; Bircher et al. 2013), by renormalization of the climatology (Parrens et al. 2014; Brocca et al. 2011; Lacava et al. 2012; Su et al. 2012; Draper et al. 2009) or by selecting representative nodes for a point-node comparison (Al Bitar et al. 2012). This study will use the representative nodes selected by Al Bitar et al. (2012) as an initial baseline and will match the climatology of the SMOS footprint soil moisture to that of the *in situ* measurement by CDF matching (Drusch et al. 2005) and Linear Rescaling (Draper et al. 2009).

6.4 Analysis and Methods

6.4.1 Bias Correction and Error Characterization

Consider two independent datasets of the same geophysical variable sampled at concurrent instances in time. The reference dataset R is assumed to be related to the observed or bias corrected dataset θ by the linear function:

$$R(t) = c + b\theta(t) + \varepsilon \tag{6-1}$$

where c is a constant that results in additive bias to θ with respect to R, the parameter b is the function slope resulting in multiplicative bias and ε is random error $N(0,\sigma)$. Renormalization of the dataset θ' is accomplished with a function f and parameters Θ

$$\theta' = f(\theta, R, \Theta) \tag{6-2}$$

designed to map the values of θ to those of R. Under ideal circumstances, that is without the random error component, the renormalized dataset would exactly match the reference dataset. However, under the constraints of imperfect conditions bias correction methods seek to transfer the statistical properties of the reference set to θ', and to reduce the terms c and b as near as possible to zero. The tolerance for remaining additive or multiplicative bias following bias correction depends on the application.

The two presented bias correction methods, CDF matching and linear rescaling, both reduce the additive bias to near zero and constrain the variance of the observations to that of the reference dataset. The fundamental difference between both methods is that CDF matching, under ideal circumstances, will exactly reproduce the reference CDF thereby matching all statistical moments. Linear rescaling will only match the first two statistical moments to the reference dataset, but with a shape that is reminiscent of the observed data, particularly in the tails of the distribution. The specification of terms is therefore important, where both methods are referred to as bias correction, and CDF matching is reserved for techniques that preserve higher order moments as well. This has not always been the case in practice.

There are various performance evaluation statistics available for the validation/comparison of retrieved soil moisture. The three most commonly used measures are the RMSE (6-3), the correlation coefficient (6-4) and the bias (6-5):

$$RMSE = \left(\frac{1}{N} \sum_{i=1}^{N} \left(R - \theta \right)^2 \right)^{\frac{1}{2}} \tag{6-3}$$

$$r = \frac{\mathrm{cov}(R, \theta)}{\sigma_R \sigma_\theta} \tag{6-4}$$

$$B = \mu_R - \mu_\theta \tag{6-5}$$

where N is the sample size and *cov* is the covariance. The RMSE is a measure of global fit between two datasets; the correlation coefficient r is a measure of the linear dependence between two variables and is not affected by changes in bias or the slope of the linear fit; and bias measures the consistent difference between two datasets at the mean values μ. In addition to the use of statistical measures, an analysis of residuals will be presented to demonstrate the presence of systematic errors with respect to the SCAN observations. Since the focus of this research is the correction of biases, the spatial mismatch between SMOS and SCAN observations will be neglected.

6.4.2 CDF Matching

Cumulative Distribution Function matching (CDFM) techniques have been used to reduce the bias between satellite observed soil moisture and that of reference values from land surface models or *in situ* measurements (Reichle and Koster, 2004; Draper et al. 2011; Crow and van den

Berg, 2010). Various implementations are available to match the CDFs to two datasets such as piecewise linear interpolation (Liu et al. 2011) and the use of fitted polynomials (Drusch et al. 2005). CDF matching in this study was carried out by fitting a third order polynomial to the ranked satellite soil moisture and the differences between ranked observed and ranked satellite soil moisture (Drusch et al. 2005).

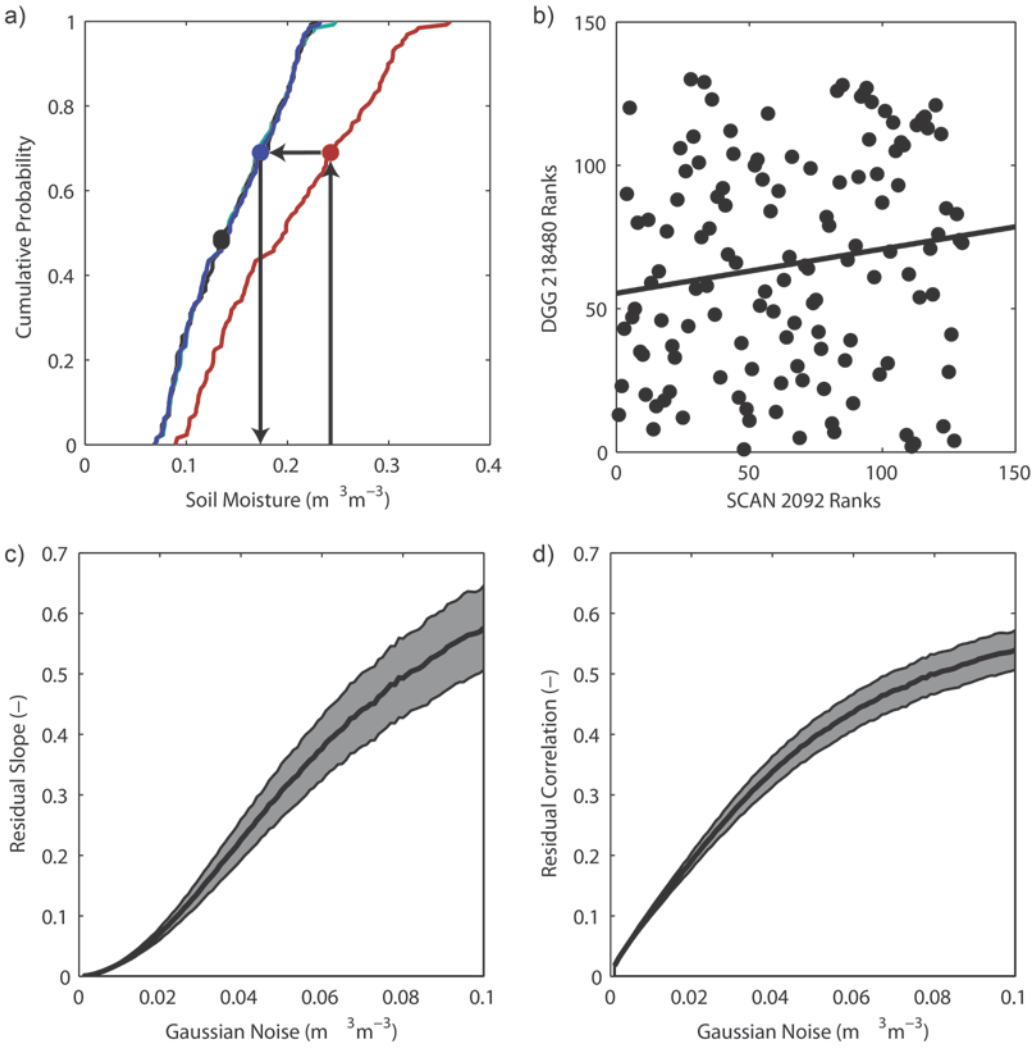

Figure 6-2: a) The concept of CDF matching a SMOS retrieval for DGG 218480 (red)-SCAN 2092(black) where the CDF matched (blue) and linear rescaled(turquoise) soil moisture are shown. b) DGG-SCAN rank correlation. c) Mean residual slope ±1σ and d) Mean residual correlation ±1σ from 1000 repetitions of adding mean zero Gaussian noise.

Figure 6-2 a) demonstrates the general principal of CDF matching where the SMOS soil moisture representative of a particular cumulative probability was scaled to match that of the SCAN soil moisture. A perfect match would result in the two datasets sharing the same statistical moments, eliminate the bias, decrease RMSE and increase linear correlation between the two

datasets. A shortcoming of the process comes from the necessity to rank (sort) the two datasets. When the two datasets are temporally consistent it can be expected that CDF matching will be more effective, however, since the satellite soil moisture product is known to be contaminated with errors due at least to the sensors radiometric uncertainty (~2 K for SMOS; Kerr et al.2010), and possibly systematic errors, the quality of the CDF match can be expected to suffer. This can be seen in Fig. 6-2 b) which is a scatter plot of the soil moisture ranks for SCAN 2092 and DGG 218480. It is clear from Fig. 6-2 b) that the ranks are correlated although a large degree of variability is also evident. The result is that the cumulative probability of any individual satellite retrieval at a particular time t may have a considerably different cumulative probability from the *in situ* measurement at t. This can be seen in Fig. 6-2 a) where the retrieved ($0.24 m^3 m^{-3}$ (P = 0.69)) and observed ($0.14 m^3 m^{-3}$ (P = 0.48)) soil moisture of 17 Apr 2010 are shown. To demonstrate the impact of differing rank values on CDF matching a numerical simulation is conducted. The soil moisture values from SCAN 2092 that match the SMOS retrieval times (n=130) for the year 2010 were taken as a reference dataset. The dataset was corrupted with additive Gaussian noise and the CDF of the noisy dataset was matched to that of the original dataset. This experiment was repeated 1000 times for each σ = [0, …, $0.1 m^3 m^{-3}$]. Fig. 6-2 c) and d) shows that as the amount of noise in the dataset increased, the process of CDF matching resulted in higher residual slope (multiplicative bias) and higher correlation between the residuals and the reference soil moisture dataset. Note that in this experiment, the initial conditions were such the c and b of Eq. (6-1) were zero such that any linear trend or correlation in the residuals could only have been induced by matching the CDFs of noisy data.

6.4.3 Linear Rescaling

Linear rescaling (LR) is a second approach that will be analyzed for matching the climatology of retrieved and observed soil moisture. Linear rescaling matches the mean and standard deviation of the retrieved soil moisture to the reference dataset following Draper et al. (2009):

$$\theta' = \mu_R + \frac{\sigma_R}{\sigma_\theta}\left(\theta - \mu_\theta\right) \tag{7-6}$$

where θ' is the rescaled retrieved soil moisture θ that is matched to the mean μ and standard deviation σ of the reference dataset R. For reference, the CDF of the linear rescaled soil moisture for SCAN 2092 was presented in Fig. 6-2 a). It is clear that linear rescaling matches the mean

and variance of the *in situ* soil moisture, but that the shape of the CDF of the linear rescaled soil moisture better resembles that of the SMOS retrieved soil moisture, particularly at the tails. This represents a fundamental difference of the two methods.

6.4.4 Correction of Multiplicative Bias

Figure 6-2 shows that multiplicative bias can be induced by the temporal mismatch between *in situ* and retrieved soil moisture when matching CDFs. A similar issue arises with linear rescaling since the SMOS retrievals are known to be corrupted, thereby inflating the observed variance. Bootstrap resampling of the SMOS retrievals and *in situ* data was tested as a method to overcome these shortcomings. By resampling with replacement for 1000 replicates, an ensemble of datasets with similar statistical properties, but no temporal dependence was generated. Both bias correction methods were applied to each observed (SMOS) and reference (SCAN) replicate. The parameters for the final bias correction are taken as mean parameters from the bootstrap ensemble (BT) or the single realization that produced the lowest residual slope (multiplicative bias) (LS). The resampling therefore results in four additional bias correction methods: CDFM-BT, CDFM-LS, LR-BT and LR-LS.

6.5 Results

6.5.1 SMOS Validation

Several updates have been made to the SMOS L1 processor and SMOS L2 soil moisture processor since the original analysis of Al Bitar et al. (2012) using version 4.00 of the SMOS L2 operational processor. The main updates to version 5.51 presented in this study were an improvement in RFI detection and the change in the dielectric

Table 6-2: SMOS Validation for 2010

DGG Number	SCAN ID	Bias	R	RMSE	N
172276	2160	0.063	0.451	0.080	63
186675	2018	0.048	0.755	0.063	88
187758	2002	-0.006	0.664	0.050	26
203609	2093	-0.006	0.723	0.049	120
203626	2001	0.004	0.685	0.070	68
218480	2092	0.058	0.741	0.076	130
219434	2168	0.062	0.614	0.072	113
231355	2079	-0.007	0.737	0.065	109
235420	2030	0.032	0.869	0.061	44
237475	2084	-0.031	0.737	0.078	100
240043	2024	-0.042	0.610	0.091	121
241088	2053	-0.073	0.827	0.083	113
5023541	2051	0.002	0.327	0.070	67

constant model from the model of Dobson Model (Dobson et al. 1985; Hallikainen et al. 1985) to that of Mironov et al. (Mironov et al. 2013). Considering the same study period of 2010, Table 6-2 shows statistical performance of the SMOS soil moisture retrieval at the SCAN sites. Caution should be used in making a direct comparison to the results of Al Bitar et al. (2012) as there are several key differences between the two presented data sets, mainly the use of only ascending overpasses in this study and less potential retrievals due to better RFI flagging. Generally, the performance of the soil moisture retrievals has improved with the updated processor version with the largest improvements being a large reduction in bias at some SCAN sites. Comparing the retrieval performances found in Table 6-2 with the performance metrics of the same sites presented in Table 6-3 of Al Bitar et al. (2012) the mean absolute bias has decreased by $0.02m^3m^{-3}$, the mean correlation has increased by 0.08 and mean RMSE has decreased by $0.01m^3m^{-3}$. As previously mentioned, a direct comparison of results should be made with caution, however these differences for mean absolute bias and mean correlation are greater than the respective $0.01m^3m^{-3}$ and 0.04 differences between 2010 and 2011 retrievals found when the same data processing was considered herein.

The performance of the ascending soil moisture retrievals for 2011 at the SCAN sites is presented in Table 6-3. The values of Table 6-3 are highlighted for those sites where the absolute differences in performance between 2010 and 2011 exceed a threshold of $0.02m^3m^{-3}$, 0.1 and $0.02m^3m^{-3}$ for bias, correlation and RMSE respectively. Consideration of these differences is important when applying bias correction such as CDF matching or linear rescaling as larger differences in retrieval performance between years can result in matching parameters from a calibration time period that are not valid in later

Table 6-3: SMOS Validation for 2011

DGG Number	SCAN ID	Bias	R	RMSE	N
172276	2160	0.066	**0.724**	0.075	85
186675	2018	0.040	0.715	0.061	122
187758	2002	**0.044**	**0.788**	0.065	38
203609	2093	0.009	0.645	0.053	139
203626	2001	**-0.022**	0.588	0.070	100
218480	2092	0.049	0.757	0.083	88
219434	2168	0.052	**0.507**	0.063	132
231355	2079	-0.022	0.720	0.067	109
235420	2030	**0.009**	0.808	0.070	96
237475	2084	-0.034	0.723	0.087	102
240043	2024	**-0.081**	0.605	0.108	147
241088	2053	**-0.007**	**0.645**	**0.105**	120
5023541	2051	**0.109**	**-0.043**	**0.146**	22

periods. In most cases these large differences between 2010 and 2011 retrieval performances are in sites where there are fewer than 70 retrievals in at least one of the two years. Therefore, these differences may result from the smaller sample size or the influence of factors such as RFI, flooding or snow cover which have resulted in a lower number of successful retrievals. However, small sample size does not account for the differences in annual performance at SCAN 2168, SCAN 2024 or SCAN 2053. Inspection of temporal plots of these sites provides insights to the differences in the retrieval performance (See supplementary material). The soil moisture observed by SCAN 2168 was very dry during the winter of 2011 and is over-predicted by SMOS observations. The temporal discrepancy from SCAN 2024 arises from the larger number of winter observations in 2011 for which SMOS retrievals had a dry bias. The soil moisture at SCAN 2053 appears to be abnormally wet in 2010 which was expected to be a combination of differences in meteorological conditions between the two years (T. Toldsdorf, personal communication, 25 July, 2014). In all three cases the SMOS retrievals appear consistent from year to year despite differences in observed soil moisture.

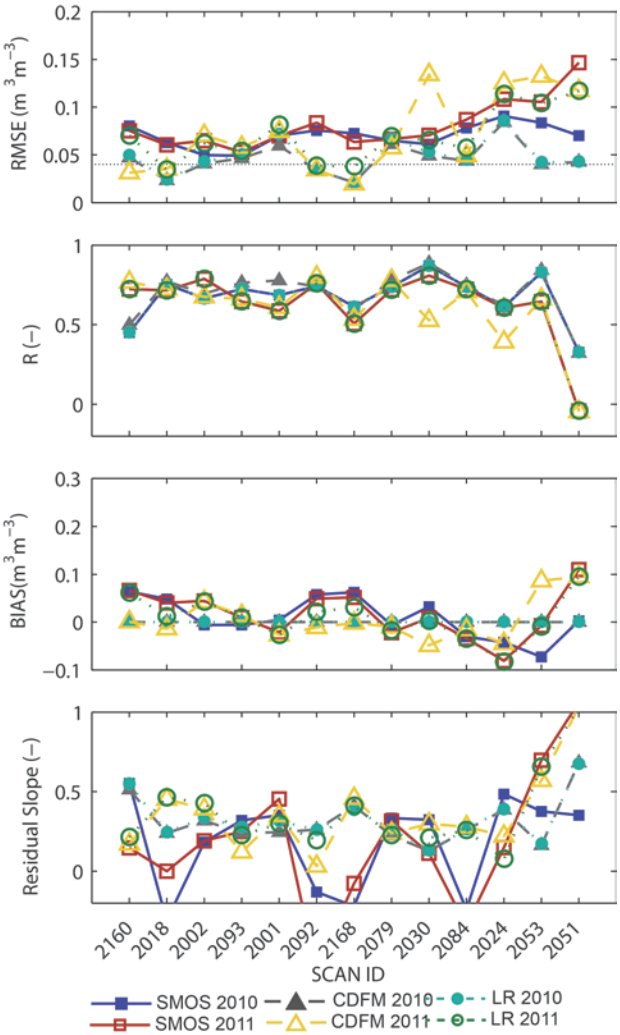

Figure 6-3: Performance of SMOS, CDF matched and linear rescaled soil moisture compared to SCAN sites for the calibration year 2010 and validation year 2011. The dashed line in the RMSE figure is the SMSO target of $0.04 m^3m^{-3}$.

6.5.2 Bias Correction Results

The performance metrics for each site, bias correction technique and year can be seen in Fig. 6-3. Both CDF matching and linear rescaling reduce the bias to zero ($<1.0 \times 10^{-16} m^3m^{-3}$) for all study sites during the calibration period and RMSE and correlation are comparable between the two methods. A similar result was found in Australia by Su et

al. (2013). In all but two stations (SCAN 2092; 2168), CDF matching produced lower RMSE and higher correlation than linear rescaling. For the data sets where linear rescaling produced better results than CDF matching the difference in RMSE was approximately $0.001 m^3 m^{-3}$ and the difference in correlation was less than 0.01, and therefore these differences are considered negligible. The mean RMSE improved from $0.07 m^3 m^{-3}$ at all stations to $0.046 m^3 m^{-3}$, which is slightly above the SMOS target of $0.04 m^3 m^{-3}$ (Kerr et al. 2010), when the CDF of the SMOS data was matched to the *in situ* measurements. The multiplicative bias of the SMOS retrievals varies greatly between sites and sometimes years. CDFM constrains the multiplicative bias to slope of ~0.25 - 0.5 which was an improvement at most sites, but decreased performance in terms of multiplicative bias at SCAN 2002, 2092 and 2051. The residual slope of the LR bias correction technique is similar to that of the SMOS data, since the manipulation of data is less with the LR technique.

The purpose of bias correction is often to determine a set of parameters which will be valid for continued use. Therefore, the validation of the CDF matching and linear rescaling techniques is an important step which has often been neglected. Here, the fitted parameters for CDF matching and linear rescaling that were fitted to the 2010 datasets were applied to the SMOS and SCAN data for the year 2011. The results in Fig. 6-3 demonstrate that the performance of these methods decreases significantly if the distributions of the soil moisture or SMOS retrievals during the two years differ significantly. This is most evident when looking at the bias values. By definition the bias is related to the difference in means, so when the soil moisture climatology is different between two time periods, the result is both bias correction methods fail to adequately remove the bias. At 8 of 13 sites, the CDF matching technique performed better at removing the bias during the validation period, but was only superior to linear rescaling at 6 and 5 sites for RMSE and *r* respectively. The differences in terms of RMSE were found to be ~$0.01 m^3 m^{-3}$ between the two methods for the validation period and ~0.05 for correlation performance.

6.5.3 Multiplicative Bias

For the purposes of this study, multiplicative bias is defined as those errors (residuals) which are correlated to the *in situ* soil moisture. The presence of multiplicative bias has also been noted by Su et al. (2013) following renormalization of SMOS, ASCAT and AMSR-E data, where errors

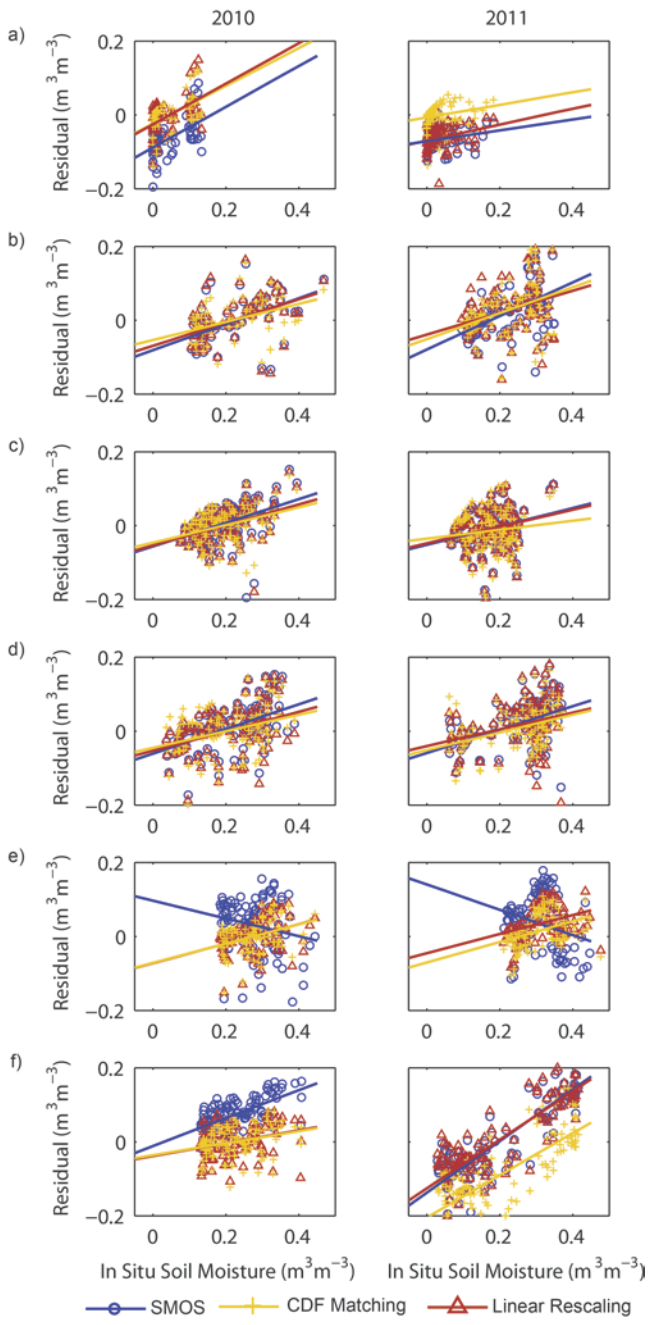

Figure 6-4: Residual (SCAN-SMOS[blue], CDFM[yellow], LR[red] showing multiplicative bias represented as the slope of the line of best fit.

were found to be negatively correlated to soil moisture. The presence of multiplicative bias in the SMOS retrieved and bias corrected soil moisture for selected sites can be seen in Fig. 6-4. At 4 of the 13 sites that were studied, the error between SMOS retrieved soil moisture and SCAN soil moisture observations were negatively correlated to SCAN soil moisture. Following the correction of the SMOS data to that of the *in situ* soil moisture all sites were found to have errors which were positively correlated to *in situ* soil moisture.

Correcting the bias of the SMOS data can be seen to provide some improvement in the magnitude (slope) of the systematic errors and causes a slight reduction in the correlation of the errors with respect to *in situ* soil moisture. The residual plots in Fig. 6-4, and those for sites not shown, revealed that the variance of the errors about the line of best fit was relatively constant, that is, the residuals had uniform variance about the line of best fit, not about the theoretical zero value. Since the variance is and the mean value following matching will produce there exists a condition where the variance of the residuals can become inflated if there magnitude (slope) of the multiplicative bias is high. As an example, the standard deviation of the residuals for SCAN 2053-DGG 241088 (Fig 6-4. (f)) was 0.041m^{3}m^{-3} for the year 2010 prior to CDF matching and was 0.040m^{3}m^{-3} following CDF matching. The standard deviation of the residuals

about the line of best fit was 0.030m^3m^{-3} and 0.039m^3m^{-3} prior and post CDF matching respectively. This concept is illustrated in Fig. 6-5 using the original data (prior to CDF matching) from Fig 6-4. (f). It is clear from Fig. 6-5 that the standard deviation of the residuals is overestimated for values below the mean soil moisture and underestimated for values above the mean soil moisture and that the standard deviation about the linear trend line is a better representation of the random component of the residuals.

6.5.4 Correction of Multiplicative Bias

Figure 6-6 shows the results of bootstrap resampling to reduce the presence of multiplicative bias. The performance of CDFM-BT for the calibration year was found to be similar to that of standard CDFM. A two sample Kolmogorov-Smirnov test (α = 0.05) determined that all distributions were the same between the CDFM and CDF-BT corrected SMOS retrievals. During the validation period, slight differences in the CDFM polynomial between the standard and BT method resulted in minor differences in performance at most sites with SCAN 2030 and 2024 being notable exceptions. By selecting the replicate with the lowest multiplicative bias, the CDFM-LS bias correction produced a CDF that was the same (α = 0.05) as the *in situ* data at all sites except SCAN 2093 and 2030. For further applications of the LS method, the Kolmogorov-Smirnov test should be used to reject parameter sets that produce the lowest multiplicative bias, but do not correctly match the distribution. Neglecting the results from these two sites, the CDFM-LS method results in a trade-off where multiplicative bias is reduced at the cost of a slight increase in RMSE and absolute bias. The mean differences between CDFM and CDFM-LS were about 0.009m^3m^{-3}, 0.008, 0.003m^3m^{-3}, and 0.2 for RMSE, correlation, bias, and residual slope respectively. This shift represents negligible differences in the traditional performance metrics, but a considerable

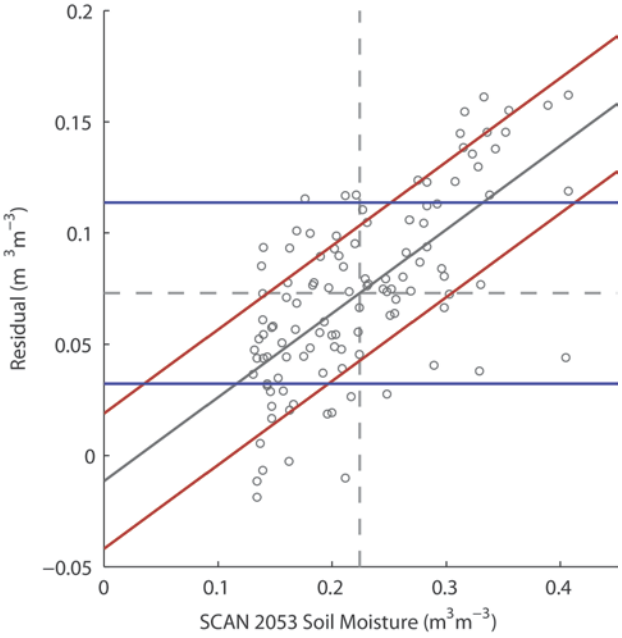

Figure 6-5: Residual for SCAN 2053-DGG 241088 showing the impact of multiplicative bias on variance (standard deviation) estimation. There is a noticeable difference when standard deviation is calculated about the mean residual (blue) or the line of best fit (red). The mean soil moisture and mean residual are plotted for reference (grey dash).

improvement in multiplicative bias. During the validation year, the CDFM-BT and CDFM-LS bias corrections performed similarly or poorer than CDFM with the greatest discrepancies in performance existing during those periods where the statistical properties of the *in situ* soil moisture for 2010 were not representative of those for 2011. The results from the bootstrapping

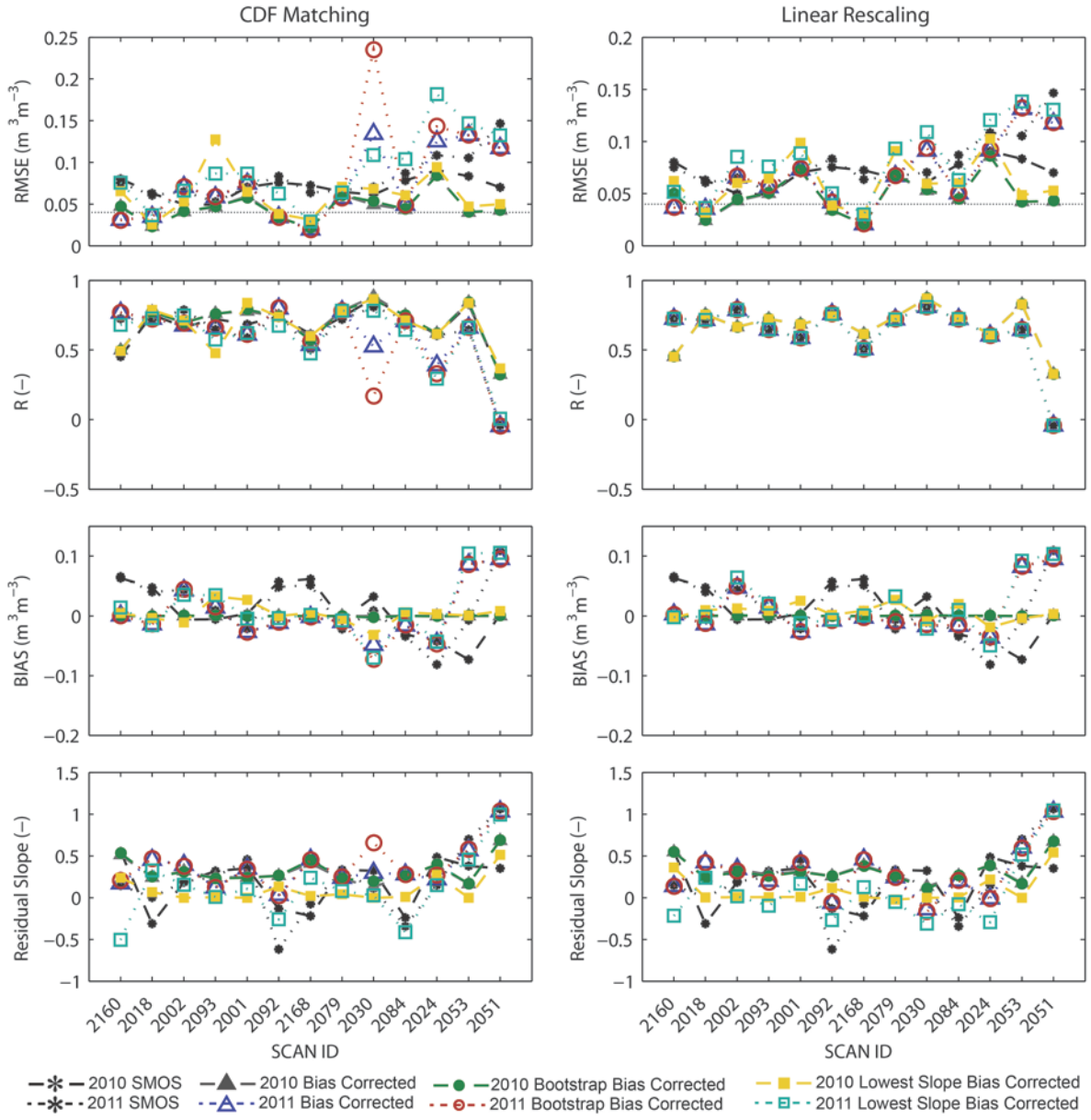

Figure 6-6: Performance of CDFM and LR bias correction methods in combination with the BT and LS resampling methods for the calibration year 2010 and validation year 2011. The black dotted line in the RMSE plot is the SMOS target of 0.04m^3m^{-3}.

treatments of LR were similar to those of CDFM. Due to the nature of the LR method, the differences in correlation between each version of the LR method was approximately zero with differences in performance being seen in the other metrics. LR-BT during the calibration period was similar to that of LR ($\alpha = 0.05$). With the exception of only one site, SCAN 2160, the LR-LS had the same CDF ($\alpha = 0.05$) as the standard LR method. As with CDFM-LS, LR-LS had lower multiplicative bias than standard LR at a slight cost to the other performance metrics. The mean differences between the performance of LR-LS and LR were $0.01 m^3 m^{-3}$, 0.00, $0.006 m^3 m^{-3}$ and 0.23 for RMSE, correlation, bias and residual slope respectively. Application of the LS method to the LR bias correction resulted in lower residual slope than CDFM-LS, although the RMSE and bias were also slightly higher. Considering the validation period, LR-BT performance was similar to LR and LR-LS was found to be similar or poorer than LR.

Figure 6-7 demonstrates the impact of the bias correction methods on the CDFs of the data at three selected sites. CDFM and CDFM-BT provide the most accurate match of the distribution functions, where the mismatch of the CDF of the CDFM-LS bias corrected data at SCAN 2093 is evident. All methods had difficulty capturing the shape of the CDF at SCAN 2001, although none of the bias corrected data were significantly different from the SCAN observations ($\alpha =$

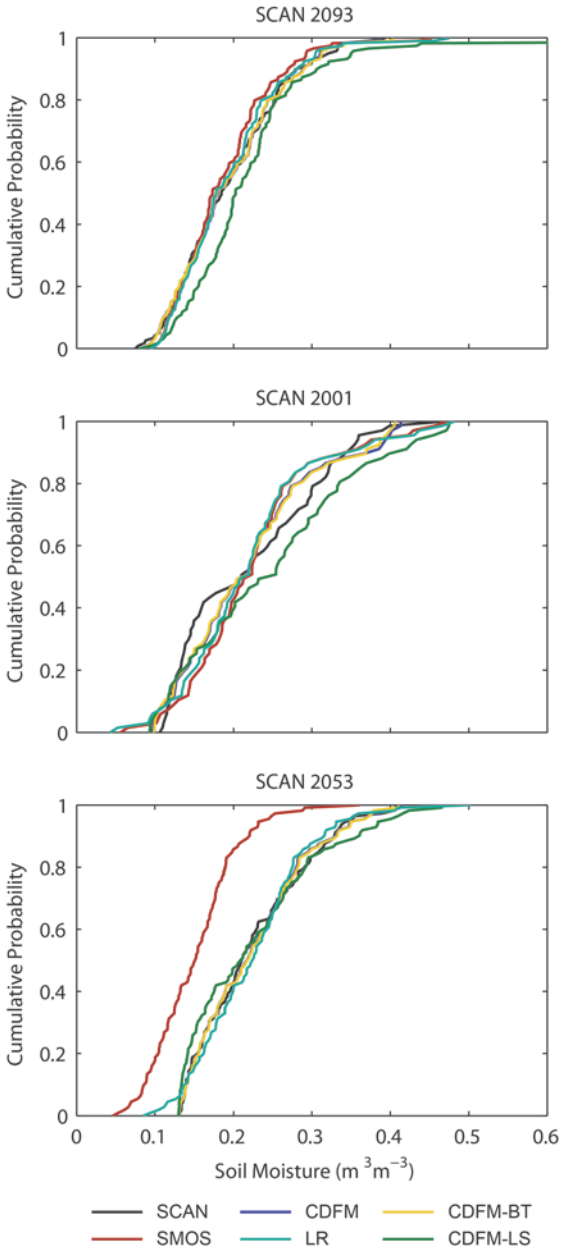

Figure 6-7: Impact of bias correction on the cumulative distribution functiosn for three select sites during the 2010 calibration year.

0.05). From a visual comparison of the shapes of the CDFs from CDFM and LR methods reveals how the shape of the CDF from the CDFM correction methods more closely resemble that of the SCAN data, whereas the shape of the LR CDF resembles that of the SMOS observations. Since the overall performance of the methods was similar in terms of performance metrics, preference of one method over the other should be based on the needs of the application.

6.6 Discussion and Conclusions

There are many applications of retrieved satellite soil moisture for which the use of raw retrievals is undesirable due to the mismatch of statistical properties with *in situ* observations or model output. With respect to *in situ* measurements from the USDA SCAN network, SMOS soil moisture was found to contain not only bias and random errors, but also multiplicative bias resulting in the under or over-estimation of soil moisture values. The multiplicative bias may result from errors in the SMOS retrieval algorithm, the spatial mismatch between the retrieval node and the point scale *in situ* measurement (Al Bitar et al. 2012; Juglea et al. 2010; Gebremichael and Vivoni, 2008) or differences in the representation of the land surface and soil properties (Grant et al. 2010; Mialon et al. 2012). If not properly accounted for, the multiplicative bias may result in the misrepresentation of statistical properties as shown in Fig. 6-5. Since most synthetic LDAS simulations assume mean zero Gaussian noise (i.e. Kumar et al. 2008; Reichle et al. 2008), the impact of multiplicative errors on soil moisture applications requires further research to determine the degree of correction required.

To extract information about the temporal properties of the satellite observations, bias correction techniques are used (Draper et al. 2011; Liu et al. 2012; Su et al. 2013). This study demonstrates that these temporal properties may be skewed by the presence of multiplicative biases which are either inherent in the observations themselves, or are induced by the bias correction techniques. In either case, the presence of multiplicative bias is undesirable for many applications of satellite observations. It was demonstrated that the multiplicative bias could be significantly reduced by resampling both the satellite observations and the reference data to create new datasets with similar statistical properties to bias correct. The resampling technique resulted in a slight increase in RMSE and bias, which was considered negligible compared to the reduction in multiplicative biases. The benefit of this new bias correction method to reduce

multiplicative bias in LDAS applications will be the subject of future research, as it is presently unclear if the added computational cost of resampling results in a realized benefit for applications.

The methods of CDF matching (Drusch et al. 2005) and linear rescaling (Draper et al. 2009) were evaluated at each SCAN site and were found to produce comparable results during the calibration period. During the validation year, linear rescaling produced more consistent results and slightly better performance than the CDFM technique. The significant decrease in performance, particularly in terms of bias, at some sites demonstrates the need for the development of temporally changing bias correction methods (Drusch et al. 2005). In this study the two-sample Kolmogorov-Smirnov test was used to evaluate the distributions of the bias corrected data. This test could serve as a filter to determine when the reference dataset is no longer represented by the parameters of the bias correction method and new parameters should be determined.

To ensure the representativeness of the reference dataset, Reichle and Koster (2004) recommended spatial sampling within a spatial window to increase the length of soil moisture records. Similarly, Drusch et al. (2005) apply CDFM to regions in the US presumed to exhibit similar soil moisture properties and found that in the western US temporal sampling was more reliable than spatial sampling and the opposite was true for the eastern US. The findings herein were similar, where the transferability of both bias correction methods from one year to the next was not necessarily dependent on the number of observations, but their representativeness. Since geophysical variables are well known to vary temporally (Seneviratne et al. 2013; Peel and Bloschl, 2011; Coulibaly, 2006) and surface heterogeneity can cause considerable spatial variability in satellite observations, even in relatively small regions (Kornelsen and Coulibaly, submitted; Rosenbaum et al. 2012; Kornelsen and Coulibaly, 2013), statistical methods, such as the Kolmogorov-Smirnov test, should be used in bias correction to ensure that the established parameters for the bias correction are applicable to the current spatial and temporal conditions.

6.7 References

Albergel, C., Zakharova, E., Calvet, J-C. Zribi, M., Pardé, M., Wigneron, J-P., et al. (2011). A first assessment of the SMOS data in southwestern France using in situ and airborne soil

moisture estimates: The CAROLS airborne campaign. Remote Sensing of Environment, 115, 2718-2728.

Al Bitar, A., Leroux, D., Kerr, Y. H., Merlin, O., Richaume, P., Sahoo, A., et al. (2012). Evaluation of SMOS soil moisture products over continental U.S. using the SCAN/SNOTEL network. IEEE Transactions on Geoscience and Remote Sensing, 50, 1572–1586.

Al-Yaari, A., Wigneron, J-P., Ducharne, A., Kerr, Y., de Rosnay, P., de Jeu, R., et al. (2014). Global-scale evaluation of two satellite-based passive microwave soil moisture datasets (SMOS and AMSR-E) with respect to Land Data Assimilation System estimates. Remote Sensing of Environment, 147, 181-195.

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., et al. (2007). Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). Geophysical Research Letters, 34, L20401.

Berthet, L., Andréassian, V., Perrin, C., & Javelle, P. (2009). How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. Hydrology and Earth Systems Science, 13, 819-831.

Bircher, S., Skou, N., Jensen, K.H., Walker, J.P., & Rasmussen, L. (2012) A soil moisture and temperature network for SMOS validation in Western Denmark. Hydrology and Earth Systems Science, 16, 1445-1463.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., & Hasenauer, S. (2010). ASCAT soil wetness index validation through in situ and modeled soil moisture data in central Italy. Remote Sensing of Environment, 114, 2745–2755.

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., et al. (2011). Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. Remote Sensing of Environment, 115, 3390–3408.

Bronstert, A., Creutzfeldt, B., Graeff, T., Hajnsek, I., Heistermann, M., Itzerott, S., et al. (2012). Potentials and constraints of different types of soil moisture observations for flood simulations in headwater catchments. Natural Hazards, 60, 879-914.

Brubaker, K.L., & Entekhabi, D. (1996). Analysis of feedback mechanisms in land-atmosphere interaction. Water Resources Research, 32, 1343–1357.

Collow, T.W., Robock, A., Basara, J.B., & Illsont, B.G. (2012). Evaluation of SMOS retrievals of soil moisture over the central United States with currently available in situ observations. Journal of Geophysical Research, 117, D09113.

Coulibaly, P. (2006). Spatial and temporal variability of Canadian seasonal precipitation (1900-2000). Advances in Water Resources, 29, 1846-1865.

Crow, W.T., & 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, W12519.

Dall'Amico, J. T., Schlenz, F., Loew, A., & Mauser, W. (2012). First results of SMOS soil moisture validation in the upper Danube catchment. IEEE Transactions on Geoscience and Remote Sensing, 50, 1507–1516.

Das, N.N., Mohanty, B.P., Cosh, M.H., & 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, 415-429.

de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., & Isaksen, L. (2013). A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. Quarterly Journal of the Royal Meteorological Society, 139, 1199–1213.

Dobson, M.C., Ulaby, F.T., Hallikainen, M.T., & Elrayes, M.A. (1985). Microwave dielectric behavior of wet soil .2. Dielectric mixing models. IEEE Transactions on Geoscience and Remote Sensing 23, 35–46.

Draper, C.S., Walker, J.P., Steinle, P.J., de Jeu, R.A.M., & Holmes, T.R.H. (2009). An evaluation of AMSR-E derived soil moisture over Australia. Remote Sensing of Environment, 113, 703–710.

Draper, C.S, Mahfouf, J. F., Calvet, J. C., Martin, E., & Wagner, W. (2011). Assimilation of ASCAT near-surface soil moisture into the SIM hydrological model over France. Hydrology and Earth System Sciences, 15, 3829–3841.

Draper, C.S., Reichle, R.H., De Lannoy, G.J.M., & Liu, Q. (2012). Assimilation of passive and active microwave soil moisture retrievals. Geophysical Research Letters, 39, L04401.

Drusch, M., Wood, E.F., & Gao, H. (2005). Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture. Geophysical Research Letters, 32, L15403.

Drusch, M. (2007). Initializing numerical weather prediction models with satellite-derived soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and the TMI soil moisture data set. Journal of Geophysical Research, 112, D03102.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., et al. (2010). The Soil Moisture Active Passive (SMAP) Mission. Proceedings of the IEEE, 98, 704–716.

Gebremichael, M., & Vivoni, E.R. (2008). Spatial sampling uncertainty in SMEX04 soil moisture fields: A data-based resampling experiment. Remote Sensing of Environment, 112, 326-336.

Gherboudj, I., Magagi, R., Goïta, K., Berg, A.A., Toth, B., & Walker, A. (2012). Validation of SMOS data over agricultural and Boreal forest areas in Canada. IEEE Transactions on Geoscience and Remote Sensing, 50, 1623-1635.

Grant, J.P., Van de Griend, A.A., Wigneron, J-P. Saleh, K., Panciera, R., & Walker, J.P. (2010). Influence of forest cover fraction on L-band soil moisture retrievals from heterogeneous pixels using multi-angular observations. Remote Sensing of Environment, 114, 1026-1037.

Hallikainen, M.T., Ulaby, F.T., Dobson, M.C., Elrayes, M.A., & Wu, L.K. (1985). Microwave dielectric behavior of wet soil .1. Empirical models and experimental-observations. IEEE Transactions on Geoscience and Remote Sensing, 23, 25–34.

Jackson, T. J., Bindlish, R., Cosh, M. H., Tianjie, Z., Starks, P. J., Bosch, D.D., et al. (2012). Validation of Soil Moisture and Ocean Salinity (SMOS) soil moisture over watershed networks in the U.S. IEEE Transactions on Geoscience and Remote Sensing, 50, 1530–1543.

Juglea, S., Kerr, Y.H., Mialon, A., Wigneron, J-P., Lopez-Baeza, E., Cano, A. et al. (2010). Modelling soil moisture at SMOS scale by use of a SVAT model over the Valencia Anchor Station. Hydrology and Earth Systems Science, 14, 831-846.

Kerr, Y. H., Waldteufel, P.,Wigneron, J. P., Delwart, S., Cabot, F., Boutin, J., et al. (2010). The SMOS mission: New tool for monitoring key elements of the global water cycle. Proceedings of the IEEE, 98, 666–687.

Kerr, Y. H.,Waldteufel, P., Richaume, P.,Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., et al. (2012). The SMOS soil moisture retrieval algorithm. IEEE Transactions on Geoscience and Remote Sensing, 50, 1384–1403.

Kornelsen, K. C., & Coulibaly, P. (submitted) Design of an optimal soil moisture monitoring network using SMOS retrieved soil moisture. IEEE Transactions on Geoscience and Remote Sensing.

Kornelsen, K. C., & Coulibaly, P. (2013). McMaster Mesonet soil moisture dataset: description and spatio-temporal variability analysis. Hydrology and Earth Systems Science, 17, 1589-1606.

Kumar, S.V., Reichle, R.H., Peters-Lidard, C.D., Koster, R.D., Zhan, X., Crow, W.T., et al. (2008). A land surface data assimilation framework using the land information system: Description and applications. Advances in Water Resources, 31, 1419-1432.

Lacava, T., Matgen, P., Brocca, L., Bittelli, M., Pergola, N., Moramarco, T., et al. (2012). A first assessment of the SMOS soil moisture product with in situ and modeled data in Italy and Luxembourg. IEEE Transactions on Geoscience and Remote Sensing, 50, 1612–1622.

Leroux, D.J., Kerr, Y.H., Richaume, P., & Fieuzal, R. (2013). Spatial distribution and possible sources of SMOS errors at the global scale. Remote Sensing of Environment, 133, 240-250.

Liu, Y.Y., Parinussa, R.M., Dorigo, W.A., De Jeu, R.A.M., Wagner, W., et al. (2011). Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. Hydrology and Earth Systems Science, 15, 425-436.

Liu, Y.Y., Dorigo, W.A., Parinussa, R.M., de Jeu, R.A.M., Wagner, W., McCabe, M.F., et al. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals. Remote Sensing of Environment, 123, 280-297.

Mialon, A., Wigneron, J-P., de Rosnay, P., Escorihuela, M.J., Kerr, Y.H. (2012). Evaluating the L-MEB model from long-term microwave measurements over a rough field, SMOSREX 2006. IEEE Transactions on Geoscience and Remote Sensing, 50, 1458-1467.

Mironov, V., Kerr, Y.,Wigneron, J. P., Kosolapova, L., & Demontoux, F. (2012). Temperature and texture-dependent dielectric model for moist soils at 1.4 GHz. IEEE Geoscience and Remote Sensing Letters, 10, 419–423.

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

Owe, M., de Jeu, R., & Walker, J. (2001). A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index. IEEE Transactions on Geoscience and Remote Sensing, 39, 1643–1654.

Owe, M., de Jeu, R., & Holmes, T. (2008). Multisensor historical climatology of satellite derived global land surface moisture. Journal of Geophysical Research — Earth Surface, 113, F01002.

Pan, M., Sahoo, A. K., Wood, E.F., Al Bitar, A., Leroux, D., & Kerr, Y.H. (2012). An initial assessment of SMOS derived soil moisture over the continental United States. IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing, 5, 1448-1457.

Parrens, M., Mahouf, J.-F., Barbu, A.L., & Calvet, J.-C. (2014). Assimilation of surface soil moisture into a multilayer soil model: design and evaluation at local scale. Hydrology and Earth Systems Science, 18, 673-689.

Peel, M. C., & Blöschl, G. (2011). Hydrological modelling in a changing world. Progress in Physical Geography, 35, 249-261.

Reichle, R.H., & Koster, R.D. (2004). Bias reduction in short records of satellite soil moisture. Geophysical Research Letters, 31, L19501.

Reichle, R.H., & Koster, R.D. (2005). Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. Geophysical Research Letters, 32, L02404.

Reichle, R. H., Crow, W. T., & Keppenne, C. L. (2008). An adaptive ensemble Kalman filter for soil moisture data assimilation. Water Resources Research, 44, W03423.

Rosenbaum, U., Bogena, H. R., Herbst, M., Huisman, J.A., Peterson, T. J., Weuthen, A., et al. (2012). Seasonal and event dynamics of spatial soil moisture patterns at the small catchment scale. Water Resources Research, 48, W10544.

Rowlandson, T.L., Hornbuckle, B.K., Bramer, L.M., Patton, J.C., & Logsdon, S.D. (2012). Comparison of evening and morning SMOS passes over the Midwest United States. IEEE Transactions on Geoscience and Remote Sensing, 50, 1544-1555.

Sánchez, N., Martínez-Fernández, J., Scaini, A., & Pérez-Gutiérrez, C. (2012). Validation of the SMOS L2 soil moisture data in the REMEDHUS network (Spain). IEEE Transactions on Geoscience and Remote Sensing, 50, 1602–1611.

Schaefer, G.L., Cosh, M.H., & Jackson, T.J. (2007). The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). Journal of Atmospheric and Oceanic Technology, 24, 2073-2077.

Seneviratne, S.I., Wilhelm, M., Stanelle, T., van den Hurk, B., Hagemann, S., Berg, A., et al. (2013). Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment. Geophysical Research Letters, 40, 5212-5217.

Su, C.-H., Ryu, D., Young, R.I., Western, A.W., & Wagner, W. (2013). Inter-comparison of microwave satellite soil moisture retrievals over the Murrumbidgee Basin, southeast Australia. Remote Sensing of Environment, 134, 1-11.

Wigneron, J. P., Kerr, Y., Waldteufel, P., Saleh, K., Escorihuela, M. J., Richaume, P., et al. (2007). L-band microwave emission of the biosphere (L-MEB) model: Description and calibration against experimental data sets over crop fields. Remote Sensing of Environment, 107, 639–655.

Yilmaz, M.T., Crow, W.T., Anderson, M.C., & Hain, C. (2012). An objective methodology for merging satellite and model-based soil moisture products. Water Resources Research, 48, W11502.

Chapter 7: Conclusions and Recommendations

7.1 Conclusions

The research presented in this Ph.D. thesis focuses on a practical method for deriving high resolution soil moisture information that can be assimilated into land surface and hydrological models. Research using high resolution airborne or *in situ* observations of soil moisture has demonstrated that the assimilation of soil moisture can improve model skill, particularly for generating forecasts (Pauwels et al. 2002; Crow and Wood, 2003). Similar results have been obtained for coarse resolution soil moisture information that had been disaggregated (Merlin et al. 2006). The chapters of this thesis demonstrated the necessity of high resolution soil moisture and its potential sources, the efficacy of the bias correction approach based on the concept of temporal persistence, the demonstration of the proposed method using *in situ* data and a distributed hydrological land surface scheme, improvements in the bias correction method and a computationally tractable approach to determine root-zone soil moisture from surface soil moisture measurements. It is expected that the findings in this research will benefit future land data assimilation applications, numerical weather prediction and flood and drought forecasting.

The main conclusions of the thesis are summarized as follows:

7.1.1 Requirement for High Resolution Soil Moisture and its Potential Sources

- High resolution soil moisture will be beneficial to many applications, however, the current capacity of synthetic aperture radar to deliver this data with the consistency and quality that is required for operational applications is currently limited.
- The launch of future SAR constellations should improve the potential of SAR for soil moisture retrieval because the same geographic area will be able to be observed from multiple satellites at different incidence angles. This will provide a more reliable retrieval of soil moisture and limit the influence of an under-determined retrieval problem.
- The dissagreggation of radiometer TB with SAR, which is proposed for SMAP, may be a viable method for generating high resolutions soil moisture data but is limited by the number of incidence angles viewed.
- Methods which dissagregate soil moisture based on ancillary satellite observations such as MODIS suffer from long revisit periods and cloud cover. Methods which use surface information such as topography are of limited quality. Methods requiring model data are

accurate but have high data requirements. This is alleviated by data assimilation since the model is already in use.

7.1.2 The Generation of Root-Zone Soil Moisture from Surface Observations

- Root-zone soil moisture is of greater importance for flood and drought prediction than surface observations but is based on simple index approaches (Wagner et al. 1999) or data assimilation (Reichle et al. 2014). There is therefore a need for a root-zone soil moisture product that is reliable, accounts for non-linear soil moisture dynamics but is computationally tractable.
- Artifical neural networks could reliably predict soil moisture at depths that were ≤ 30 cm if they were trained using adequate data. The trained neural network could be applied at many locations within a given region, such as the Great Lakes Basin.
- The use of an ensemble of neural networks increased the quality and robustness of the root-zone soil moisture product.

7.1.3 Theoretical Considerations of Bias Correction for Downscaling

- Temporal persistence, or the concept that the relative difference between soil moisture at a particular point and that of the watershed average was stable in time, was demonstrated in the humid-continental Hamilton-Halton watershed using models and *in situ* data, the humid subtropical Little Washita and Little River watersheds using *in situ* data.
- The concept of temporal persistence was extended to brightness temperature, which is a function of soil moisture, soil roughness, vegetation thickness, vegetation water content and other influences.
- Decomposition of the mean squared difference (MSD) into the bias, conditional bias and variance/timing errors demonstrated that bias and variance errors were major contributors to the difference between coarse (watershed/satellite) scale observations and those at higher resolution.
- Using data from the McMaster Mesonet, it was demonstrated that moderate to large rainfalls rarely impacted the relative spatial distribution of soil moisture. Under wet conditons or following large rainstorms that did homogenize the relative spatial

distribution of soil moisture, the temporally persistent rank pattern re-emerged within 18-24 hours even at poorly drained sites.

7.1.4 MESH-CMEM Coupling

- MESH was able to simulate soil moisture in the Hamilton-Halton watershed with a high degree of skill.
- The coupling of MESH and CMEM was used to simulate TB at multiple incidence angles that was of comparable quality to other CMEM validation experiments.
- Ascending overpasses of SMOS were more consistent with MESH-CMEM simulations in the Hamilton-Halton watershed. Possible reasons for the differences between orbits may be RFI, heterogeneity of the land surface temperature during the evening overpass, dew and yet to be modelled influences of vegetation water content.
- SMOS TB observations were more consistent with MESH-CMEM than SMOS soil moisture retrievals. This finding likely results from unrelated ancillary data used in the SMOS L2 soil moisture retrieval algorithm and provides practical support for theoretical considerations that assimilation systems should use TB rather than soil moisture retrievals (Lahoz and De Lannoy, 2014).

7.1.5 Practical Results from Bias Correction Downscaling Experiments

- In both bias correction downscaling experiments with *in situ* data from the USDA watersheds and MESH-CMEM it was found that coarse scale data could be corrected to match local scale observations with high skill ($R \approx 0.9$ USDA; $R > 0.95$ MESH-CMEM).
- Bias correction was capable of correcting SMOS scale data to MESH-CMEM grid scale (~800 m) despite heterogeneity of soil texture, vegetation cover and topography. The main limitation to the approach was grid cells that had high fractions of impervious surfaces, open water or wetlands.
- Statistical analysis demonstrated that the skill of downscaled SMOS observations compared to individual MESH-CMEM grid cells was consistent with SMOS observations compared to MESH-CMEM simulations that were scaled to SMOS resolution. This indicates very little information was lost during the downscaling process.

7.1.6 Enhancement of Bias Correction Methods

- Following both linear rescaling and CDF matching, systematic multiplicative biases were found in the bias corrected data. In the case of CDF matching, this could be directly attributed to the bias correction technique.
- Bias correction parameters derived by resampling of the orginal satellite and *in situ* soil moisture data resulted in a bias corrected dataset that had slightly improved correlation ($R_{Diff} \approx 0.01$) and a significant decrease in residual correlation ($R_{\varepsilon Diff} \approx 0.2$) at a cost of a negligible increase in RMSE ($RMSE_{Diff} < 0.01 m^3 m^{-3}$).
- The selection of parameters for bias correction was identified as an important topic as differences in retrieval performance, climatological conditions, etc. could result in inappropriate bias correction parameters. The naïve application of these parameters could produce worse soil moisture data than uncorrected data.

7.2 Recommendations for Future Research

The research presented in this thesis represents the first steps towards establishing the proposed method as a viable approach to data assimilation. The work presented herein demonstrated in a few watersheds that bias correction was a valid method and more work is required in differing watersheds with different climate and surface conditions to determine the limitations of the proposed method. Similarly, bias correction was evaluated herein using CMEM simulations of brightness temperature. It was decided to use CMEM because high resolution airborne obervations of brightness temperature typically only exist for a few days during any given study (i.e. dall' Amico et al. 2013; Magagi et al. 2013) and it is therefore difficult to reliably establish statistical parameters for the bias correction methods.

In contrast to the benefits of natively high resolution soil moisture data, which has been well established (Pauwels et al. 2009; Koren et al. 2008; Parada and Liang, 2008); to date little research has been done on the assimilation of downscaled soil moisture data (i.e. Merlin et al. 2006). While many theoretical and practical considerations were addressed herein, the method has yet to be applied in an actual data assimilation system. In doing so, it is also necessary that the bias correction downscaling approach be tested against other accepted downscaling methods such as the dissagregation of soil moisture using MODIS data (Merlin et al. 2006; 2012) or the

downscaling of satellite observations in the data assimilation system itself using the Kalman Gain to account for the differences in scale (De Lannoy et al. 2010). Unlike Merlin et al. (2012), the method proposed herein can provide a downscaled data product for every satellite observation. Also, it is expected that the bias correction technique will be more computationally efficient than the Jacobian calculation required to determine the Kalman Gain in De Lannoy et al. (2010). This method is currently being investigated for inclusion in the Canadian Land Data Assimilation System (CaLDAS; Carrera et al. 2014).

Similarly, the findings in Chapter 6 need to be evaluated within an actual data assimilation system. Figure 6-5 highlights the potential problem of selecting error covariances in a data assimilation system with residual multiplicative bias' present. Future research should focus on automatic error parameter calibration (Crow and van den Berg, 2010) to determine if the LDAS is selecting the biased or un-biased error covariance and if the computation cost of resampling will result in tangible benefits to the assimilation analysis.

Another topic of considerable interest and requiring ongoing research is when and how often bias correction parameters should be determined. Such a decision has yet to be made for CaLDAS (Carrera, M. Personal communication, July, 2014) and a moving window approach has been selected for the assimilation of the ASCAT soil moisture index (Bartalis et al. 2008) into the ECWMF weather prediction system (De Rosnay et al. 2013). The recommendation to use the Kalmogorov-Smirnov test in Chapter 6 is a potential alternative to a constantly moving window that could be evaluated less often.

There will soon be an abundance of coarse resolution soil moisture and brightness temperature observations from experimental satellites such as SMOS (Kerr et al. 2010) and SMAP (Entekhabi et al. 2010) operational meteorological satellites such as the MetOp ASCAT (Bartalis et al. 2008) and the future operational version of SMOS in the planning stages (Zurita et al. 2013). Passive microwave observations are more consistent and accurate than active microwave alternatives (Appendix B) but at a resolution which is coarser than many hydrological processes. There is therefore an urgent need for practical approaches to downscale these satellite observations for applications in systems like CaLDAS, forthcoming high resolution numerical weather prediction systems, and operational drought and flood forecasting systems such as the Canadian Adaptive Flood Forecasting and Early Warning System (CAFFEWS, Coulibaly, P. personal communication, December 17, 2013).

7.3 References

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J. and Anderson, C. (2008) Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT), *Geophysical Research Letters* 34, L20401.

Carrera, M.L., Charpentier, D., Bélair, S., Zhang, S. and Bilodeau, B. (2014) Downscaling SMOS brightness temperatures to produce higher-resolution soil moisture analyses. *Proceedings of IGARRS, the IEEE Geoscience and Remote Sensing Symposium,* Quebec City, Canada.

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, 137-149.

Dall'Amico, J.T., Schlenz, F., Loew, A., and Mauser, W. (2012) First results of SMOS soil moisture validation in the upper Danube catchment. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 1507–1516.

De Lannoy, G.J.M., Reichle, R.H., Houser, P.A., Arsenault, K.R., Verhoest, N.E.C., and Pauwels, V.R.N. (2010) Satellite-scale snow water equivalent assimilation into a high-resolution land surface model, *Journal of Hydrometeorology*, 11, 352-369.

de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., and Isaksen, L. (2013) A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, 139, 1199–1213.

Entekhabi, D., Njoku, E.G. , O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, J.K, Moghaddam, Moran, M. S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L. and van Zyl, J. (2010) The Soil Moisture Active Passive (SMAP) Mission. *Proceedings of the IEEE*, 98, 704–716.

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.-J., Font, J. , Reul, N., Gruhier, C., Juglea, S.E., Drinkwater, M.R., Hahne, A., Martin-Neira, M. and Mecklenburg, S. (2010) The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proceedings of the IEEE*, 98, 666–687.

Koren, V., Moreda, F., & Smith, M. (2008) Use of soil moisture observations to improve parameter consistency in watershed calibration. *Physics and Chemistry of the Earth*, 33, 1068-1080.

Lahoz, W.A., and De Lannoy, G.J.M. (2014) Closing the gaps in our knowledge of the hydrological cycle over land: Conceptual problems, *Surveys in Geophysics*, 35, 623-660.

Magagi, R., Berg, A., Goïta, K., Bélair, S., Jackson, T.J., Toth, B., Walker, A., McNairn, H., O'Neill, P.E., Moghaddam, M., Gherboudj, I., Colliander, A., Cosh, M.H., Burgin, M., Fisher, J.B., Kim, S.-B., Mladenova, I., Djamaï, N., Rouseeau, L.-P.B., Belanger, J., Shang, J., and Merzouki, A. (2013) Canadian Experiment for Soil Moisture in 2010 (CanEx-SM10): Overview and Preliminary Results, *IEEE Transactions on Geoscience and Remote Sensing*, 51, 347-363.

Merlin, O., Chehbouni, A., Boulet, G., & Kerr, Y. (2006). Assimilation of disaggregated microwave soil moisture into a hydrologic model using coarse-scale meteorological data. *Journal of Hydrometeorology,* 7, 1308-1322.

Merlin, O., Rüdiger, C., Al Bitar, A., Richaume, P., Walker, J. P., and Kerr, Y. H. (2012). Disaggregation of SMOS soil moisture in southeastern Australia. *IEEE Transactions on Geoscience and Remote Sensing,* 50, 1556-1571.

Parada, L.M., and Liang, X. (2008) Impacts of spatial resolutions and data quality on soil moisture data assimilation. *Journal of Geophysical Research-Atmospheres*, 113, D10101.

Pauwels, V.R.N., Hoeben, R., Verhoest, N.E.C., De Troch, F.P., and Troch, P.A. (2002) Improvement of TOPLATS-based discharge predictions through assimilation of ERS-based remotely sensed soil moisture values. *Hydrological Processes,* 16, 995-1013.

Pauwels, V.R.N., Balenzano, A., Satalino, G., Skriver, H., Verhoest, N., and Mattia, F. (2009) Optimization of soil hydraulic model parameters using synthetic aperture radar data: An integrated multidisciplinary approach. *IEEE Transactions on Geoscience and Remote Sensing*, 47, 455-457.

Reichle, R.H., De Lannoy, G.J.M., Forman, B.A., Draper, C.S., and Liu, Q. (2014) Connecting satellite observations with water cycle variables through land data assimilation: Examples using the NASA GEOS-5 LDAS. *Surveys in Geophysics*, 35, 577-606.

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

Zurita, A.M., Corbell, I., Martin-Neira, M., Plaza, M.A., Torres, F., and Benito, F.J. (2013) Towards a SMOS Operational Mission: SMOSOps-Hexagonal. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6, 1769-1780.

Appendix A: Comparison of Interpolation, Statistical and Data-Driven Methods for Imputation of Missing Values in a Distributed Soil Moisture Dataset

Published as Kornelsen, K.C. and Coulibaly, P. (2014) Comparison of Interpolation, Statistical and Data-Driven Method for Imputation of Missing Values in a Distributed Soil Moisture Dataset, *Journal of Hydrologic Engineering*, 19, 26-43, doi: 10.1061/(ASCE)HE.1943-5584.0000767.

A.1 Abstract

Missing values in in-situ monitoring data is a problem often encountered in hydrologic research and applications. Values in a dataset may be missing due to sensor error or failure of data recording devices. While various imputation techniques have focused on hydro-meteorological data, very few studies have investigated gap filling methods for soil moisture data. This study aims to fill that gap by investigating well established statistical and data-driven methods for infilling missing values in high resolution soil moisture time series. Hourly soil moisture data have been collected in the Hamilton-Halton Watershed, Southern Ontario since 2006 at four research sites. Each site contains nine stations with TDR soil sensors at six soil depths. From these distributed datasets, values were randomly (~5%) and systematically (~20%) removed from the data in order to evaluate the effectiveness of the monthly average replacement (MAR), soil layer relative difference (SLRD), linear and cubic interpolation, artificial neural networks (ANN) and evolutionary polynomial regression (EPR) infilling methods. When values were randomly removed, interpolation, ANN and EPR were able to infill the missing values with similar efficiency, while MAR and SLRD were the least effective. Similarly, when large systematic gaps were present in the data, interpolation and ANN were the most effective methods of infilling, respectively. However, the effectiveness of both infilling methods is limited as serial gaps become larger than 72-100 hours.

A.2 Introduction

Soil moisture is an important variable that influences many processes in the hydrologic cycle. It is the driving factor in the partitioning of infiltration and runoff (Fernández-Gálvez et al. 2007; Rollenbeck and Anhuf 2007), and serves as long term water storage as well as operating as a feedback mechanism for atmospheric processes (Seneviratne et al. 2010; Entekhabi et al. 1996). While the benefits of understanding the role of soil moisture in the hydrologic cycle is now well documented, in practice, soil moisture data often suffer from missing values as a result of instrumentation failure or errors, data transmission errors, etc. The presence of missing values may compromise the integrity of a set of data and therefore limit the use of the data for various

applications and research work. It is therefore, necessary to impute (infill) missing values in order to obtain continuous time series needed for research and engineering applications.

Infilling missing data is traditionally done with statistical or geostatistical methods. Methods such as inverse distance weighting and multiple regression, both linear and non-linear, have been used for infilling precipitation data (Mair and Fares, 2009), streamflow (Ng et al. 2009) and suspended sediments (Ulke et al. 2011). These methods are recommended by ASCE (1996) and rely on the relationship of data between locations, but also suffer when influences such as topography or significant distance decrease the strength of the relationship between multiple stations (Mair and Fares, 2009). Multiple regression particularly has problems of multicollinearity, heteroscedasticity and data normality assumptions (Mair and Fares 2009). Dumedah and Coulibaly (2011) evaluated statistical infilling methods for soil moisture and found that simple methods such as monthly average replacement and rank stability methods outperformed regression-based techniques. Rank stability was introduced by Vachaud et al. (1985), and is based on the observation that, while the mean soil moisture state may change, the rank of a particular sampling location with respect to nearby sampling sites remained constant. Similarly, the organized influences on the soil moisture distribution result in spatial autocorrelation (see Western et al. 2002 for a review), although more samples than are available are required for accurate estimation of the variograms used in many geostatistical methods (Western et al. 1998). For this reason, many of the techniques presented herein will focus on infilling based on time series methods, which show particular promise due to the strong temporal autocorrelation in the hourly soil moisture data.

Recent studies have shown the benefits of multivariate machine learning techniques such as artificial neural networks (ANN) and evolutionary polynomial regression (EPR) because they can better account for the stochastic components (Bermudez et al. 2009), higher order interactions (Gheyas and Smith 2010) and hysteresis in the data (Jain et al. 2004). The advantage of using data-driven techniques is that they are derived from the data and therefore make few a priori assumptions, yet they retain the ability to represent physical processes well (Murtagh et al. 2000).

ANN techniques have been applied to data infilling mostly in atmospheric sciences (e.g. Junninen et al. 2004; Hopke et al. 2001), although some studies have found that data-driven techniques are effective at infilling data relevant to the field of hydrology (Ng et al. 2009;

Coulibaly and Evora 2007; Khalil et al. 2001; Elshorbagy et al. 2000; Murtagh et al. 2000). Simple feed-forward networks such as Multilayer Perceptrons (MLP) were demonstrated to be capable of reproducing missing values of daily precipitation and temperature (Coulibaly and Evora, 2007) and suspended sediments (Ulke et al. 2011) with a high degree of accuracy, although results can sometimes be improved through the use of more complex, higher order neural networks (Elshorbagy et al. 2008; Coulibaly and Evora, 2007) or through data pre-processing (Wu et al. 2010). EPR has also proven to be effective in modeling and infilling hydrologic data (Elshorbagy et al. 2010a; Elshorbagy et al. 2010b; Liu et al. 2008; Giustolisi et al. 2007) and has successfully been applied to the characterization and modeling of soil moisture (Elshorbagy et al. 2010a; Elshorbagy et al. 2010b; Elshorbagy and El-Baroudy, 2009).

This research evaluated the effectiveness of data-driven techniques (ANN and EPR) for the purpose of infilling missing values in hourly soil moisture data in Southern Ontario, Canada. Data-driven techniques were compared to standard interpolation techniques and to the statistical methods identified by Dumedah and Coulibaly (2011) as being the most effective methods for infilling random missing values in high resolution soil moisture data. Each method was evaluated using Pearson Correlation Coefficient (PCC) and Root Mean Square Error (RMSE) as measures of the effectiveness of the various techniques.

A.3 Study Area and Data

The four study sites of the McMaster MESONET soil moisture time domain reflectometry (TDR) array are located in the Hamilton-Halton Watershed in Southern Ontario (Fig. A-1). The watershed is approximately 75% rural-agricultural with the majority of urban development occurring along the Lake Ontario shoreline. Two TDR sites are located in the Dundas Valley Conservation Area, which are Orchard (a) and Governor Road (d). Both sites have silty loam soils with good drainage and moderate infiltration patterns. Orchard is reclaimed agricultural land which is on a gently north facing slope. It is predominantly meadow with sparse apple trees. Governor Road consists of mixed forest which is predominantly pine and has steep and gentle slopes. The two Kelso sites (b, c) have clay loam soils resulting in poor drainage and slow infiltration. The terrain is generally flat and ground cover consists of meadow grasses and some recently planted coniferous trees.

Throughout this paper, a site refers to a particular geographic area in which an entire soil moisture array exists. Each site has nine TDR stations, which are consecutively numbered and consist of six TDR probes (CS616) at approximate depths of 10 cm, 20 cm, 30 cm, 50 cm, 70 cm and 100 cm below the surface. All of the soil moisture data at a site are recorded onto a single data logger (CR10X) which is situated at station 1 (see Fig A-1). Such a network is cost effective; however given that all the site data are collected using a unique data-logger, a main drawback is a systemic gap that can occur when the data logger experiences a problem or a failure. Data can therefore be missing for the entire site for a given period. Here, soil moisture data have been collected at hourly intervals since the end of 2006. In this study, data from 2006 to 2010 at 10-70 cm depths are used. Each site has an associated meteorological station

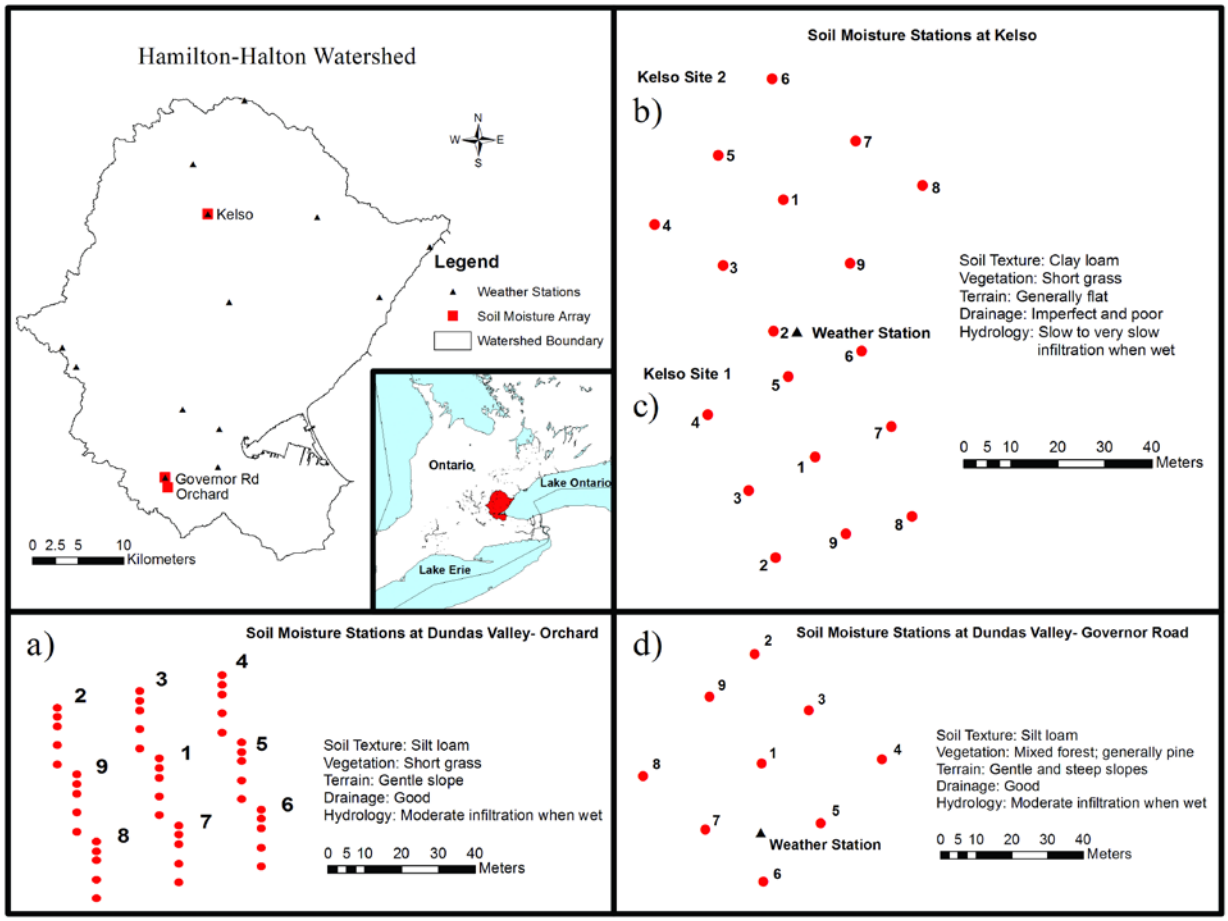

Figure A-1: The Hamilton-Halton Watershed, Ontario (Canada). Site a) is shown in three dimensions to highlight the architecture of the TDR array (depth not to scale); site (b-d) are represented as a top down plan view.

collecting temperature, relative humidity, precipitation, wind speed and incident solar radiation at half hourly intervals. Reference evapotranspiration ET_o was calculated for each hour based on the Penman-Monteith method (Monteith, 1965).

A.4 Methodology

A.4.1 Data Pre-processing

The soil moisture data were preprocessed to remove any actual missing values from the observed data. The resulting complete data set was non-continuous in time but contained no missing values. From the complete data set, two artificial missing value treatments were used. The first treatment involved randomly removing values from a site, station and depth. This resulted in many missing values, but it did not necessarily represent all conditions where values may be missing in actual field research conditions. To further test the various methods ability to infill missing values, which may result in long spans of systematic missing values for an entire site, a second experiment was used which removed data at increasing intervals (10-1200 serial values) as well as the randomly removed values. The systematically created gaps were persistent across all stations and depths for the given time periods. This treatment was used to simulate the temporary failure of the data recording device for various lengths of time. The relative amount of missing values from each study site is summarized in Table A-1. The systematic gap condition is somewhat unique to these high spatial resolution soil moisture arrays as all soil moisture stations at a particular site are connected to a common data logger, which is not the case in many other hydrological data sets as each monitoring station often has its own data logger. This pre-condition limits the effectiveness of geostatistical techniques since intra-site comparison is not possible and inter-site comparison is difficult because of the complexity of the soil moisture relationship with the topography, meteorology and physiography at the different sites.

Table A-1: Relative Amount of Missing Values at Each Site

Site	Random	Systematic	Total	Data Points (n)
a) Orchard	5%	14%	19%	22127
b) Kelso 02	6%	15%	21%	19475
c) Kelso 01	5%	20%	26%	15138
d) Governor Rd	6%	16%	21%	18700

Each infilling method was trained using data from the respective missing value treatments or cases to ensure infilling methods were robust in the face of actual data gaps. A training set was established for each method, except interpolation, by removing or ignoring the artificial missing values in the data set. The remaining time periods with missing values provided an independent validation data set. This procedure maximized the size of the training set (i.e. all years could be considered) without compromising cross-validation.

A.4.2 Monthly Average Replacement Method (MAR)

The MAR method was recently selected by Dumedah and Coulibaly (2011) as being effective in infilling missing soil moisture values. Therefore, the method is retained as a benchmark for this investigation. Complete data were used to generate the monthly average and variance for each month at every depth for a given station. It is represented as

$$\overline{\theta}_m = \frac{1}{n}\sum_{m=1}^{n}\theta_{im} \tag{A-1}$$

where θ is the soil moisture in cm^3cm^{-3} during month m at station i. When there was a missing value the average soil moisture and its associated variance error for only that month, site, station and depth (i.e. January, Orchard, Stn 1, 10 cm) was used as a replacement. The method used the data set with missing values and calculated the mean for each month ignoring missing values, and stored the mean and variance in a lookup table. When data were missing the associated values were selected from the lookup table.

A.4.3 Soil Layer Relative Difference (SLRD)

SLRD was the other selected statistical method identified by Dumedah and Coulibaly (2011) as being highly effective for infilling missing soil moisture values. It relies on the stability of rank (Vachaud et al. 1985) of the soil moisture stations at each site with respect to one another and assumes a degree of homogeneity within a given layer. The site mean was calculated similar to equation (A-1). The rank relative difference, δ_{ij}, is determined in equation (A-2) by taking the difference between soil moisture θ_{ij} at station i for time j and the mean value of the total layer, $\overline{\theta_j}$.

$$\delta_{ij} = \frac{\theta_{ij} - \overline{\theta_j}}{\overline{\theta_j}} \tag{A-2}$$

The missing value was estimated θ_{est} as the sum of the layer mean θ_i and the mean relative difference $\overline{\delta_j}$ between the station and the layer mean. The soil moisture and its variance error for the given time period were therefore calculated according to equations A-3 and A-4.

$$\theta_{est} = \theta_i + \theta_i \overline{\delta_j} \tag{A-3}$$

$$\frac{1}{\sigma^2} = \frac{1}{\sigma_i^2} + \frac{1}{\sigma_{(\delta_j)}^2} \tag{A-4}$$

The SLRD method was also applied to the data sets with missing values present. Ignoring missing values, the mean relative difference between the station mean and the layer mean was calculated and stored in a lookup table. When a value was missing at a discrete point in time the layer mean was calculated and the mean relative difference from the lookup table was multiplied by the layer mean soil moisture. This produced a unique value at each time step. Unfortunately, for this method to be effective there must be some values present in a given layer at the time step being considered. If no data are present, as is the case with the systematic gap treatment, then SLRD cannot be applied.

A.4.4 Interpolation

Linear interpolation was conducted using piecewise polynomial functions (Kahaner et al. 1989). The linear piecewise polynomial $\theta(j)$ is solved by:

$$\theta(j) = \theta_k \frac{j - j_{k+1}}{j_k - j_{k+1}} + \theta_{k+1} \frac{j - j_k}{j_{k+1} - j_k} \tag{A-5}$$

if $j_k \le j \le j_{k+1}$ where the soil moisture at time j was interpolated for the length of the gap to be interpolated k=1,2,…,n-1. This allowed missing values at each time j to be interpolated by solving for the unique value of $\theta(j)$ in equation A-5.

Cubic Interpolation was conducted using the Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) (Kahaner et al. 1989; Fritsch and Carlson, 1980). PCHIP produces a polynomial similar to linear interpolation, however, PCHIP also accounts for an extra data point

on either side of the missing data so that the derivative of the cubic polynomial function is continuous. This creates smooth interpolants $\theta(j)$ on the subinterval $j_k \leq j \leq j_{k+1}$.

Both Linear and Cubic Interpolation were conducted with missing values present. The respective polynomials were calculated for every gap in the data and used to interpolate missing values. It is important to note that there was no periodicity of the polynomial equations; rather a separate polynomial was used for every gap regardless of its size. Linear interpolation filled the data gap with a linear function, whereas PCHIP produces a similar non-linear (smoothed) function that does not overshoot the points to be interpolated (Kahaner et al. 1989).

A.4.5 Artificial Neural Networks

The Multilayer Perceptron (MLP) was the neural network selected for investigating the infilling of missing data and is the standard feed-forward neural network used in many applications (Ulke et al. 2011; Coulibaly and Evora 2007, Jain et al. 2004; Elshorbagy et al. 2000). It is represented by equation A-6, which consists of an input vector, a hidden layer and an output layer. The input layer is connected to the hidden layer via the weight matrix which connects all of the input variables to each neuron in the hidden layer. The data is processed through a sum function and transferred to the output layer neuron (Haykin 1999). In this case, a single output neuron was used. The output of the network is given by:

$$y_k = f\left(\sum_{j=1}^{h} \mathbf{w}_j \mathbf{G}(\mathbf{s}_i) + b_k\right) \tag{A-6}$$

where $f(.)$ is the linear activation function of the output neuron k, w_j is the weight matrix connecting the hidden and output units, b_k is the output neurons bias, and si is the weighted sum of the input data of the hyperbolic tangent sigmoid function $\mathbf{G}(.)$ used as the activation function in the hidden neurons. The advantages of the MLP network is that it is simple in design and has proven to be both robust and capable of being applied in a variety of applications in both hydrology and data infilling (Coulibaly and Evora 2007; Junninen et al. 2004; Hopke et al. 2001; Murtagh et al. 2000). The ANNs were trained with the Bayesian Regularization Algorithm (MacKay 1992). Bayesian Regularization was chosen because it incorporates the training of the Levenberg-Marquardt algorithm while optimizing the generalization capabilities and providing consistent training results.

Five meteorological variables were selected from amongst the measured variables using a combination of sensitivity analysis and correlation analysis as well as selection to remove redundancy from the data. Sensitivity analysis was conducted using the NeuroSolutions software where an ensemble of ANN's were trained with the same architecture, shifting each input by 2σ to determine the sensitivity of the output to a particular input. Correlation analysis was used to remove inputs that were highly correlated, for example, relative humidity (RH) and ET_o were selected over various measures of vapour pressure. This procedure resulted in the selection of inputs which were uncorrelated and for which the network was highly sensitive. Coulibaly and Evora (2007) suggested that simple MLP networks are more effective than various forms of time lag and recurrent neural networks for infilling daily precipitation and temperature data and also offers the advantage of parsimony. Therefore, the MLP network was trained and simulated as a static network with static meteorological variables and antecedent soil moisture data lagged in the training data as an input prior to being presented to the network.

Table A-2: MLP-BR Model Structure and Parameters along with Inputs.

Variable infilled	Inputs	Number of hidden units	Hidden layer transfer function	Training method	Output layer transfer function
θ_{10cm}	$\theta_{10,j-1}$, T,RH,SR,ET,API	40	Tansig	Bayesian Regularization	Linear
θ_{20cm}	$\theta_{20,j-1}$, θ_{10}, T,RH,SR,ET,API	40	Tansig	Bayesian Regularization	Linear
θ_{30cm}	$\theta_{30,j-1}$, θ_{20}	40	Tansig	Bayesian Regularization	Linear
θ_{50cm}	$\theta_{50,j-1}$, θ_{30}	40	Tansig	Bayesian Regularization	Linear
θ_{70cm}	$\theta_{70,j-1}$, θ_{50}	40	Tansig	Bayesian Regularization	Linear

Note: API is antecedent precipitation index; ET is evapotranspiration (mm); RH is relative humidity (%); SR is incoming solar radiation (Wm^{-2}); T is temperature (°C); and Tansig is the hyperbolic tangent sigmoid function

The number of hidden neurons was selected by trial-and-error approach. The hidden neurons were incrementally increased and the MSE of the model calculated at each increment. The number of hidden neurons was selected by the determination of the point at which increasing the number of hidden neurons did not produce significant improvement in the ensemble MSE. The MLP applied at all sites had the same architecture and gave similar results, and 40 neurons were selected as the optimal number of hidden units. The MLP-BR model parameters are shown in Table A-2.

The MLP networks were trained for every site, station and depth using data from 2009. The random missing value treatment was used for each site to train the network, where periods

with artificial missing values were removed from the training set. The partial autocorrelation function applied to the soil moisture data determined that highly significant autocorrelation was present between lag five and lag ten, therefore removing random values (with a maximum of three sequential values), did not negatively impact the training of the MLP. This also ensured that the method may be applicable if a complete training set is not available.

A.4.6 Evolutionary Polynomial Regression (EPR)

Evolutionary Polynomial Regression was first proposed by Giustolisi and Savic (2003) and is based on a hybrid evolutionary paradigm. The general symbolic expression is

$$\hat{y} = \sum_{j=1}^{m} F\left(\mathbf{X}, f(\mathbf{X}), a_j\right) + a_0 \tag{A-7}$$

where $\hat{y}$ is the model output, a_j is the constant in the polynomial expression, F is the constructed function, $\mathbf{X}$ is the input matrix, f is the user defined function and m is the number of terms in the polynomial equation.

EPR uses a genetic algorithm to determine which inputs to include in the polynomial equation, $f(\mathbf{X})$, as well as their exponents and a least squares regression method to determine the adjustable parameters, a_j, which is fitted to the data against a user defined cost function. The number of generations is determined by the number of terms in the function and the number of input variables. EPR starts by searching through the space of exponent values determined by the user, and then determining the value of the adjustable parameter (Guistolisi et al. 2004a). EPR often returns a result which is parsimonious and symbolic (Guistolisi et al. 2004a; Giustolisi and Savic 2003).

A separate EPR model was generated for each sensor. From the ensemble of candidate models, the most common structures were selected and the mean of the coefficients used. This method resulted in one EPR model per depth which was both parsimonious and mitigated the potential of EPR to overfit the training data (Savic et al. 2009). The models applied are shown in Table A-3. All of the resulting models were heavily dependent on antecedent soil moisture as can be seen by the high value of coefficient *b*.

Table A-3: Soil Moisture EPR Equations

Model Structure	***a***	***b***	***c***
$\theta_{10cm} = \theta_{jk} = aT_k RH_k^2 API_k^2 + b\theta_{jk-1} + c$	7.41e-008	0.998	0.0001
$\theta_{20cm} = \theta_{jk} = aT_k RH_k^2 API_k^2 \theta_{j-1k}^2 + b\theta_{jk-1} + c$	4.54e-007	0.997	0.0005
$\theta_{30cm} = \theta_{jk} = a\theta_{jk-1}^2 + b\theta_{jk-1} + c$	0.024	0.975	0.006
$\theta_{50cm} = \theta_{jk} = a\theta_{j-1k} + b\theta_{jk-1} + c$	0.060	0.926	0.006
$\theta_{70cm} = \theta_{jk} = a\theta_{j-1k}^2 + b\theta_{jk-1} + c$	0.130	0.929	0.135

Note: the value of *j* is the depth at time *k*.

A.5 Inputs for Data-driven Methods

Soil moisture is heavily dependent on the presence of precipitation to increase the amount of moisture in the soil. However, this became a significant problem with the data-driven techniques when implemented at an hourly time interval. In 93% of hourly measurements, precipitation was 0 mm. Since there was no precipitation in most instances, the data mining techniques did not become properly trained to account for those instances where precipitation was present. As a result, vast overestimations occurred whenever precipitation was present. To overcome this problem, a modified version of the Antecedent Precipitation Index (API) (Casenave 1982) was used. The API represents soil water storage in equation A-8 as:

$$IK_j = (IK_{j-1} + N_{h,j-1}) - ET_0 \quad \text{(A-8)}$$

where the value of the index *IK* at time *j* in mm before rain is the sum of the previous index value IK_{j-1} and the height of the previous rainfall $N_{h,n-1}$ in mm, less soil water loss. Because of the presence of known ET_o, this value was substituted for the static soil drying factor ($e^{\alpha t}$) used by Casenave (1982).

Values of the API were constrained to the highest and lowest soil moisture values for each station which was assumed to be the saturation point and permanent wilting point, respectively. The result was an index which could account for both precipitation and evapotranspiration and would provide continuous values representing the rate of precipitation and evapotranspiration to train the data-driven methods.

For each station, simulations were run incrementally moving to a deeper depth. For the 10 and 20 cm depth the MLP-BR and EPR were trained for each site, station and depth using Temperature (°C), RH (%), incident solar radiation (Wm^{-2}), ET_o ($mmhr^{-1}$), API and antecedent

soil moisture (cm^3cm^{-3}) at a one hour lag. The 20 cm depth additionally had soil moisture from the 10 cm layer as input. It was assumed that the significant serial correlation at multiple lags would result in the over-reliance of antecedent conditions by the data-driven methods. By limiting these methods to soil moisture at lag 1 only, it was hoped to minimize the sensitivity to antecedent soil moisture from multiple lags and maximizes the influence of meteorological data.

It has been demonstrated that the effects of meteorological variables on soil moisture are most pronounced in the upper soil layers (Markewitz et al. 2010; Schwarzel et al. 2010). Therefore, the networks for the 30-70 cm depths were trained using antecedent soil moisture and soil moisture from the layer immediately above the layer being modeled. In this way, soil moisture from the layer above contributed to increases or decreases in soil moisture as precipitation and ET affect the upper soil layers.

A.6 Results and Discussion

The following section presents the results of the various statistical and data-driven methods that were used to fill in the missing soil moisture values at each of the four sites. The modeled soil moisture values were compared to the validation data. Tables A-4 and A-5 will be used to compare the various methods, where the data at each site have been amalgamated by layer. A more rigorous analysis will be presented for the two methods which demonstrated the highest level of performance.

The MAR method had the poorest performance of any method for the random gap treatment. MAR poorly represented wet and dry conditions and adequately represented moderate soil moisture conditions, due to the central tendency nature of the method. As the number of serial missing values increased in the systematic gap treatment, the accuracy of MAR did not decrease to the same proportion as many of the other methods. The advantage of MAR over SLRD lies in the fact that it does not require data from within the same layer. In an instance where no values were present during a given time period, MAR was still capable of infilling missing values, whereas SLRD did not provide a value. Despite this drawback, SLRD

Table A-4: Infilling Methods Accuracy When Values are Randomly Missing

Orchard

a)

Depth	MAR		SLRD		Linear Interp		Cubic Interp		MLP-BR		EPR	
	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE
10 cm	0.326	0.099	0.855	0.044	0.999	0.002	1.000	0.003	0.999	0.003	0.999	0.004
20 cm	0.404	0.053	0.910	0.023	0.999	0.002	0.999	0.002	0.996	0.005	0.993	0.007
30 cm	0.447	0.038	0.917	0.017	0.994	0.004	0.995	0.004	0.937	0.016	0.939	0.016
50 cm	0.497	0.041	0.867	0.023	0.990	0.006	0.990	0.007	0.982	0.009	0.982	0.017
70 cm	0.549	0.036	0.913	0.017	0.998	0.002	0.998	0.002	0.936	0.015	0.996	0.006

Kelso 02

b)

Depth	MAR		SLRD		Linear Interp		Cubic Interp		MLP-BR		EPR	
	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE
10 cm	0.775	0.085	0.961	0.036	0.997	0.009	0.997	0.009	0.996	0.010	0.991	0.017
20 cm	0.752	0.086	0.945	0.034	0.989	0.015	0.988	0.016	0.985	0.019	0.981	0.020
30 cm	0.721	0.057	0.904	0.032	0.957	0.021	0.957	0.022	0.953	0.024	0.935	0.028
50 cm	0.671	0.053	0.855	0.034	0.929	0.024	0.927	0.024	0.889	0.030	0.883	0.031
70 cm	0.606	0.050	0.872	0.027	0.966	0.014	0.968	0.014	0.941	0.019	0.924	0.023

Kelso 01

c)

Depth	MAR		SLRD		Linear Interp		Cubic Interp		MLP-BR		EPR	
	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE
10 cm	0.550	0.086	0.896	0.046	0.977	0.021	0.976	0.022	0.978	0.021	0.967	0.026
20 cm	0.604	0.058	0.918	0.030	0.954	0.022	0.961	0.020	0.955	0.022	0.926	0.029
30 cm	0.587	0.052	0.884	0.030	0.921	0.024	0.916	0.025	0.934	0.022	0.893	0.029
50 cm	0.568	0.050	0.906	0.025	0.957	0.017	0.950	0.018	0.931	0.022	0.936	0.025
70 cm	0.473	0.048	0.825	0.031	0.912	0.023	0.905	0.024	0.895	0.025	0.882	0.027

Governor Road

d)

Depth	MAR		SLRD		Linear Interp		Cubic Interp		MLP-BR		EPR	
	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE
10 cm	0.316	0.071	0.847	0.040	0.996	0.005	0.997	0.005	0.994	0.006	0.945	0.024
20 cm	0.159	0.058	0.877	0.026	0.994	0.006	0.994	0.006	0.996	0.005	0.984	0.010
30 cm	0.264	0.050	0.907	0.022	0.984	0.009	0.986	0.009	0.983	0.009	0.979	0.011
50 cm	0.429	0.043	0.855	0.025	0.992	0.006	0.990	0.007	0.991	0.007	0.980	0.100
70 cm	0.348	0.063	0.858	0.024	0.979	0.009	0.981	0.009	0.971	0.011	0.973	0.011

Table A-5 Infilling Methods Accuracy When Values are Randomly and Systematically Missing

Orchard

a)

Depth	MAR		Linear Interp		Cubic Interp		MLP-BR		EPR	
	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE
10 cm	0.394	0.089	0.760	0.053	0.747	0.056	0.749	0.053	0.631	0.065
20 cm	0.294	0.057	0.682	0.040	0.650	0.042	0.815	0.027	0.190	0.364
30 cm	0.323	0.035	0.723	0.031	0.691	0.042	0.761	0.025	0.133	0.355
50 cm	0.318	0.039	0.793	0.028	0.744	0.032	0.426	0.059	0.113	0.284
70 cm	0.391	0.035	0.864	0.021	0.838	0.024	0.768	0.024	0.164	0.333

Kelso 2

b)

Depth	MAR		Linear Interp		Cubic Interp		MLP-BR		EPR	
	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE
10 cm	0.429	0.135	0.926	0.053	0.923	0.054	0.671	0.103	0.535	0.124
20 cm	0.414	0.121	0.919	0.046	0.903	0.051	0.466	0.117	-0.207	0.281
30 cm	0.512	0.080	0.868	0.040	0.868	0.040	0.483	0.082	-0.239	0.253
50 cm	0.428	0.065	0.913	0.026	0.891	0.029	0.104	0.112	-0.151	0.216
70 cm	0.299	0.053	0.954	0.016	0.913	0.024	0.363	0.070	-0.151	0.229

Kelso 1

c)

Depth	MAR		Linear Interp		Cubic Interp		MLP-BR		EPR	
	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE
10 cm	0.452	0.090	0.882	0.050	0.871	0.051	0.699	0.072	0.412	0.116
20 cm	0.452	0.074	0.893	0.040	0.870	0.042	0.685	0.060	0.235	0.395
30 cm	0.440	0.062	0.893	0.034	0.848	0.038	0.667	0.052	0.218	0.392
50 cm	0.360	0.061	0.906	0.030	0.870	0.033	0.489	0.058	0.216	0.302
70 cm	0.171	0.063	0.882	0.030	0.859	0.033	0.407	0.060	0.224	0.379

Governor Road

d)

Depth	MAR		Linear Interp		Cubic Interp		MLP-BR		EPR	
	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE
10 cm	0.234	0.065	0.801	0.050	0.791	0.051	0.853	0.032	0.500	0.063
20 cm	0.285	0.055	0.812	0.040	0.801	0.041	0.820	0.034	0.023	0.397
30 cm	0.387	0.047	0.875	0.032	0.856	0.035	0.839	0.030	-0.011	0.393
50 cm	0.501	0.044	0.869	0.031	0.845	0.033	0.791	0.033	-0.126	0.312
70 cm	0.426	0.057	0.918	0.021	0.911	0.022	0.895	0.022	-0.101	0.369

outperformed MAR in most instances for the random gap treatment, which is consistent with findings reported by Dumedah and Coulibaly (2011). Overall, both MAR and SLRD were less efficient than the interpolation, EPR and ANN based methods.

Both linear and cubic interpolation were among the best methods for infilling missing values. This was particularly true when values were randomly missing. The main reason for the high level of performance of both interpolation methods when gaps were random is the high serial correlation in the hourly soil moisture data. Since the vast majority of gaps in the data were less than the size of the significant autocorrelation lags, simply connecting the gaps with either a linear or cubic polynomial allowed for accurate representation of the soil drying characteristics. This is similar to the EPR polynomials, where in most cases the resulting EPR equation represented a standard drying curve for hourly soil moisture (See Table A-3). In the same way interpolation represented a unique rate of moisture loss/retention under given atmospheric and soil conditions.

When the amount of serial missing values was extended beyond that of significant autocorrelation lags the accuracy of interpolation methods greatly decreased. Despite this, interpolation remained one of the best methods for infilling data as the length of the systematic gaps increased. In most instances linear interpolation showed slightly better performance than cubic interpolation.

Overall the data-driven methods were able to infill data with a higher degree of accuracy than the statistical methods (ΔRMSE ~ 0.001-0.4), although often less accurately than interpolation (ΔRMSE ~ 0.001-0.01). When gaps were random both MLP-BR and EPR were able to infill the data with similar accuracy and were comparable to interpolation. RMSE for both methods was often less than ~0.01 when only random gaps were considered. When systematic gaps were considered the EPR equations resulted in the poorest infilling performance. This was the result of the inability of the EPR method to properly account for the increase in soil moisture following rainfall events. This may also suggest that the ability of EPR to model long term soil moisture at hourly resolution in a single equation is limited. If EPR is considered as a knowledge discovery tool (Giustolisi et al. 2004b), the rate at which moisture is lost from the soil per hour can be inferred from the 10 cm and 20 cm equations in Table A-3.

$$\Delta\theta \approx 1 - b \tag{A-8}$$

The rate of loss of soil moisture is approximately 0.002 cm^3cm^{-3} and the value of *b* could be substituted into equation A-7 for ET_o to produce a linear hourly API if local data is unavailable.

The MLP-BR method, while also performing worse when missing values were systematic as opposed to random, was able to better represent soil moisture and the variability in soil moisture brought on by changes in meteorological conditions. This was particularly true of the Governor Rd and Orchard sites where the PCC is consistently above 0.7 during the systematic missing value treatment.

In order to isolate the impact of surface conditions on the different infilling methods, the skill of SLRD in Table A-4 and MAR, interpolation and MLP-BR in Table A-5 are compared among sites. The second missing value treatment (Table A-5) was selected because the increased length of data gaps caused higher variability in the skill of the infilling methods. SLRD is unique in this research, as it estimates soil moisture based on a spatial comparison of soil moisture within the site making it sensitive to intra-site variability. The Kelso and Orchard sites are relatively flat and homogeneous compared to Governor Rd, which has more steeply sloped topography and greater spatial variability. The result is higher correlation between SLRD modeled data and observed values at Kelso and Orchard as compared to Governor Rd, as the spatial homogeneity of the prior sites caused SLRD to be more effective. An inter-site comparison of the individual infilling methods presented in Table A-5 reveals that both interpolation methods and MAR show the lowest infilling efficiency at Orchard and Governor Rd where the better drainage and sloping terrain results in a more dynamic soil moisture state. Conversely, an inter-site comparison of MLP-BR displays the highest skill at Governor Rd followed by Orchard and Kelso, in decreasing order.

In order to contrast the infilling methods with similar surface conditions an intra-site comparison is conducted with focus on data from Table A-5 and SLRD data from Table A-4. At the Kelso sites linear interpolation slightly outperforms cubic interpolation and both methods have significantly greater accuracy than MLP-BR, SLRD and MAR, in decreasing order. Particularly evident is the gap in skill between interpolation and MLP-BR for systematic missing values, where the moderated response of soil moisture to meteorological forcing was particularly detrimental to MLP-BR. At Orchard and Governor Rd the performance of MAR was weak compared to the other methods and SLRD performance is good but not as strong as interpolation and MLP-BR. At both Orchard and Governor Rd, the MLP-BR and interpolation techniques are

comparable. The greater drainage and slight to steep slopes increase the temporal heterogeneity of soil moisture, which slightly hampers the infilling ability of interpolation. Similarly, the coarser texture increases the soils responsiveness to meteorological forcing resulting in greater suitability of MLP-BR. Therefore, in conditions of moderate complexity interpolation and MLP-BR have similar performance.

These results indicate that the aptitude of simple infilling methods, such as MAR and interpolation, was highest when the soil moisture state was stable and homogeneous. Similarly, the dynamic MLP-BR infilled soil moisture best when the soil was responsive to external forcing, whereas the high clay content at Kelso resulted in slow soil moisture response to varying meteorological conditions, diminishing the ANN infilling skill. EPR is not considered in these comparisons as the results in Table A-5 are poor and inconsistent for the systematic missing value treatment making interpretation of EPR's skill with respect to surface conditions difficult. When missing values were random only, EPR, MLP-BR and interpolation were all competitive and performed better than SLRD and MAR.

Closer scrutiny was given to both linear interpolation and MLP-BR as the chosen best methods for infilling missing data. Figures A2-A5 show scatter and data plots for the 10 cm and 30 cm depths at each site for the random missing value treatment. These depths represent the most hydrologically active soil layer considered (10 cm) and the transition to those layers in the soil system which tend to be more hydrologically stable (30 cm).

For both depths, when missing values were random the correlation between observed and predicted values was high. Both interpolation and MLP-BR performed equally well with PCC and RMSE values being within ±0.01 of the other method. The scatter plots show more data points further from the unity line at the Kelso sites. Greater variability is particularly noticeable in the scatter plot for Kelso 01 as a result of a large number of outliers generated by the data recording device, which slightly compromised the infilling procedure. The compromise only occurred when outliers or periods of drastic changes in soil conditions were presented to the infilling algorithms.

Interestingly, interpolation and MLP-BR both performed better at the 10 cm than at the 30 cm depth. This result was contrary to what would be expected as soil moisture should be more stable and therefore easier to predict at the 30 cm depth. This lower accuracy can be easily explained for MLP-BR, because soil moisture was sequentially infilled. Since the MLP-BR

requires soil moisture from the layer above it as input, errors in infilling were propagated to lower depths. Since the missing values are assumed to be randomly distributed in this treatment, the error was minimized as missing values for the same station at multiple depths rarely occurred. MLP-BR also has a tendency to slightly overestimate missing values at 30 cm which was not present at the 10 cm depth. For linear interpolation, the decrease of infilling accuracy with depth was predominantly the result of outliers in the data, which again proved to compromise the accuracy of the infilling results.

Figures A6-A9 show the scatter and data plots when linear interpolation and MLP-BR were used to infill missing values at 10 cm and 30 cm when missing values were both random and systematic. Greater variability can be seen in the results when values were systematically removed. Systematic errors were sometimes evidenced by distinct lines in the scatter plots. In general, linear interpolation resulted in the data cloud being more closely grouped around unity than MLP-BR, although it also showed a greater tendency towards banding patterns in the scatter plots. The data plots reveal that linear interpolation tended to produce centrist biased values, especially at the 30 cm depth. In contrast, MLP-BR values more closely represent the variance of soil moisture and tend to follow wetting and drying trends in the soil moisture values, however this is not always the case (see Kelso-02 10 cm). Of the 36 TDR sensors considered at the 10 cm depth, linear interpolation produced significant mean and variance ($\alpha = 0.05$) at 5 and 6 sites respectively when systematic missing values were introduced. In comparison MLP-BR only had a significant mean and variance at 5 and 4 sites respectively. At the 30 cm depth, these statistics decreased to 5 and 2 sites for linear interpolation and 5 and 0 sites for MLP-BR ($\alpha = 0.05$). The P-value for those stations where the mean and variance of infilled data were similar to the observed soil moisture based on the Wilcoxon-Rank Sum and Levene's tests are reported in Table A-6.

Figures A-7 and A-9 show data plots of the first 1000 missing values in the systematic treatment, where the length of the gap in the values was progressively increased, seen as progressively larger gaps towards the right of the sub-figures. This methodology was selected in order to estimate the maximum size of the gap for which it was appropriate to infill missing values, and as such visual interpretation of these plots is of far greater importance than the measures of fitness, which may be misleading. The figures show that as the length of the gap increased, both of the considered infilling methods became progressively worse at infilling.

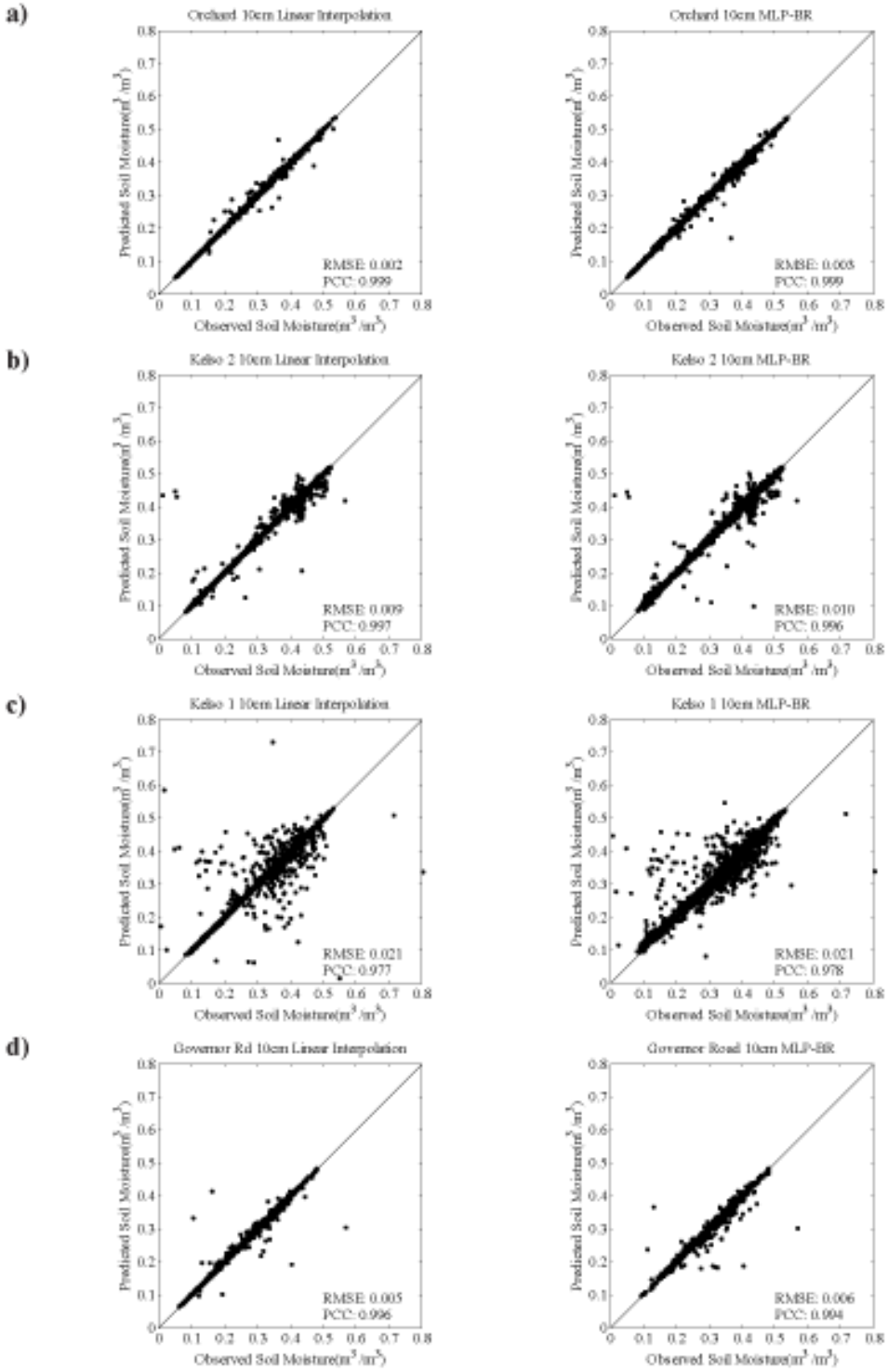

Figure A-2: Scatter plot of observed versus predicted values at 10 cm depth at (a) Orchard; (b) Kelso 02; (c) Kelso 01; (d) Governor Rd when values are randomly missing.

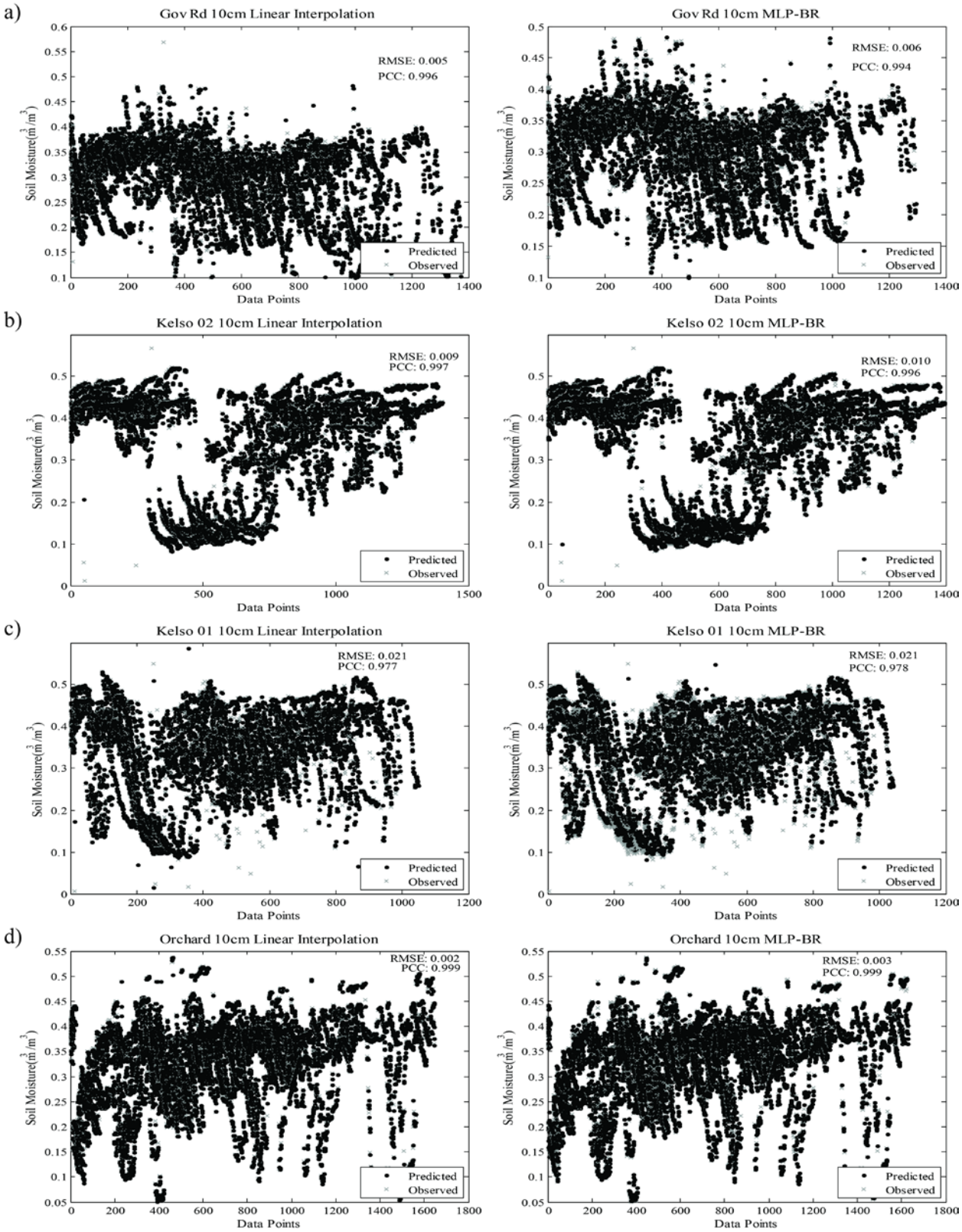

Figure A-3: Data plot of observed and predicted values at 10-cm depth at (a) Orchard; (b) Kelso 02; (c) Kelso 01; (d) Governor Rd when values are randomly missing.

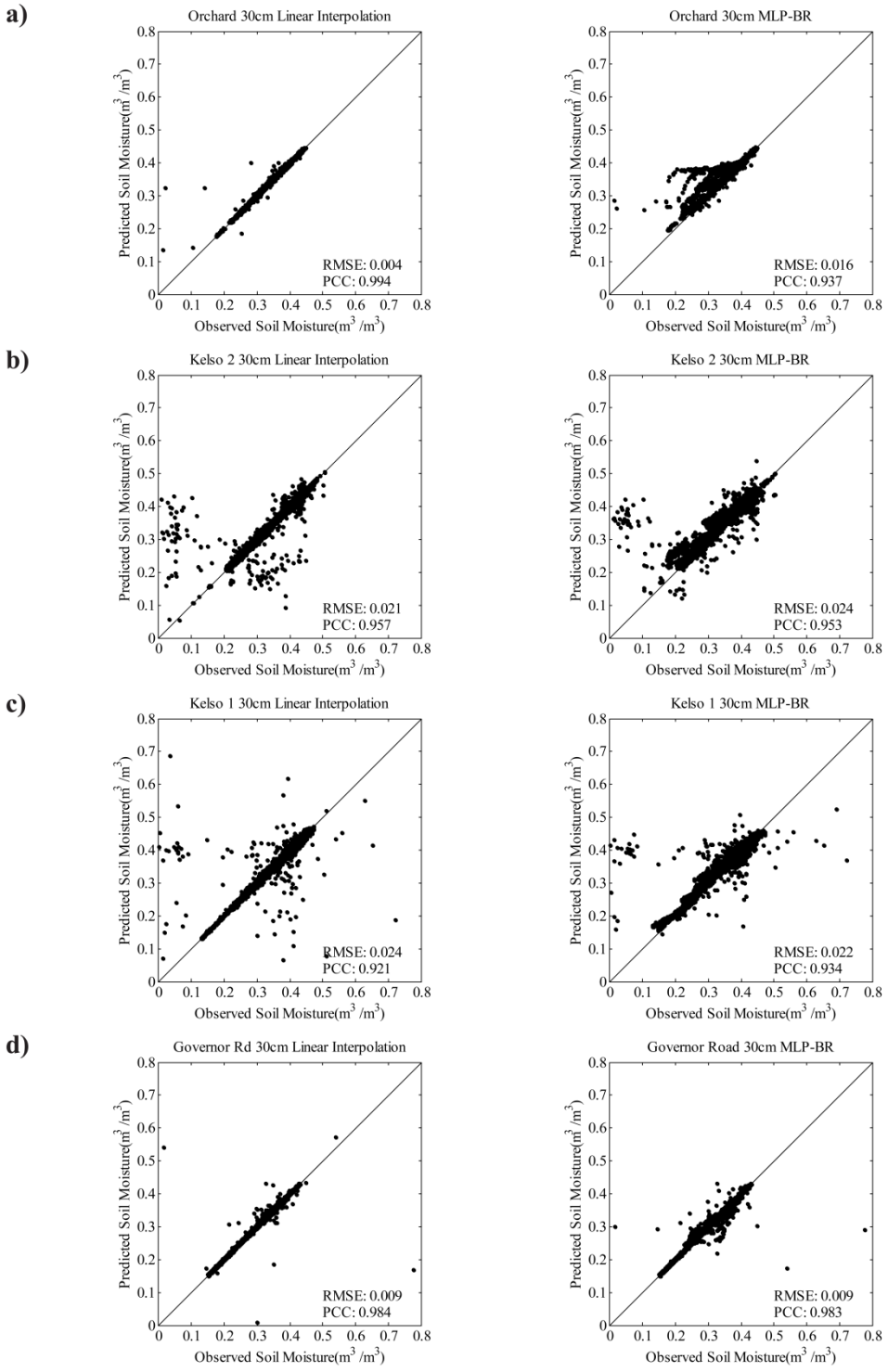

Figure A-4: Scatter plot of observed versus predicted values at 30 cm depth at (a) Orchard; (b) Kelso 02; (c) Kelso 01; (d) Governor Rd when values are randomly missing.

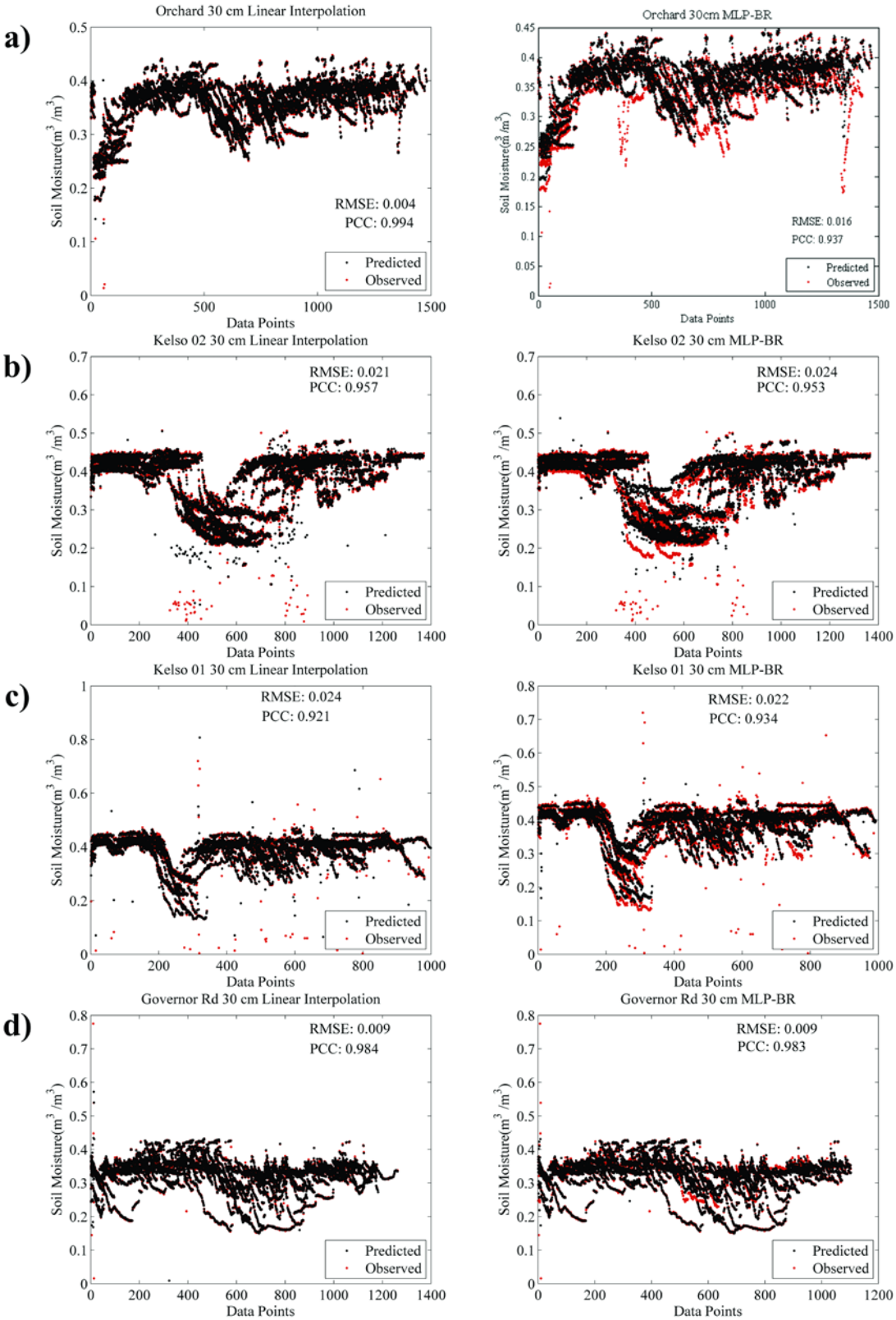

Figure A-5: Data plot of observed and predicted values at 30 cm depth at (a) Orchard; (b) Kelso 02; (c) Kelso 01; (d) Governor Rd when values were randomly missing.

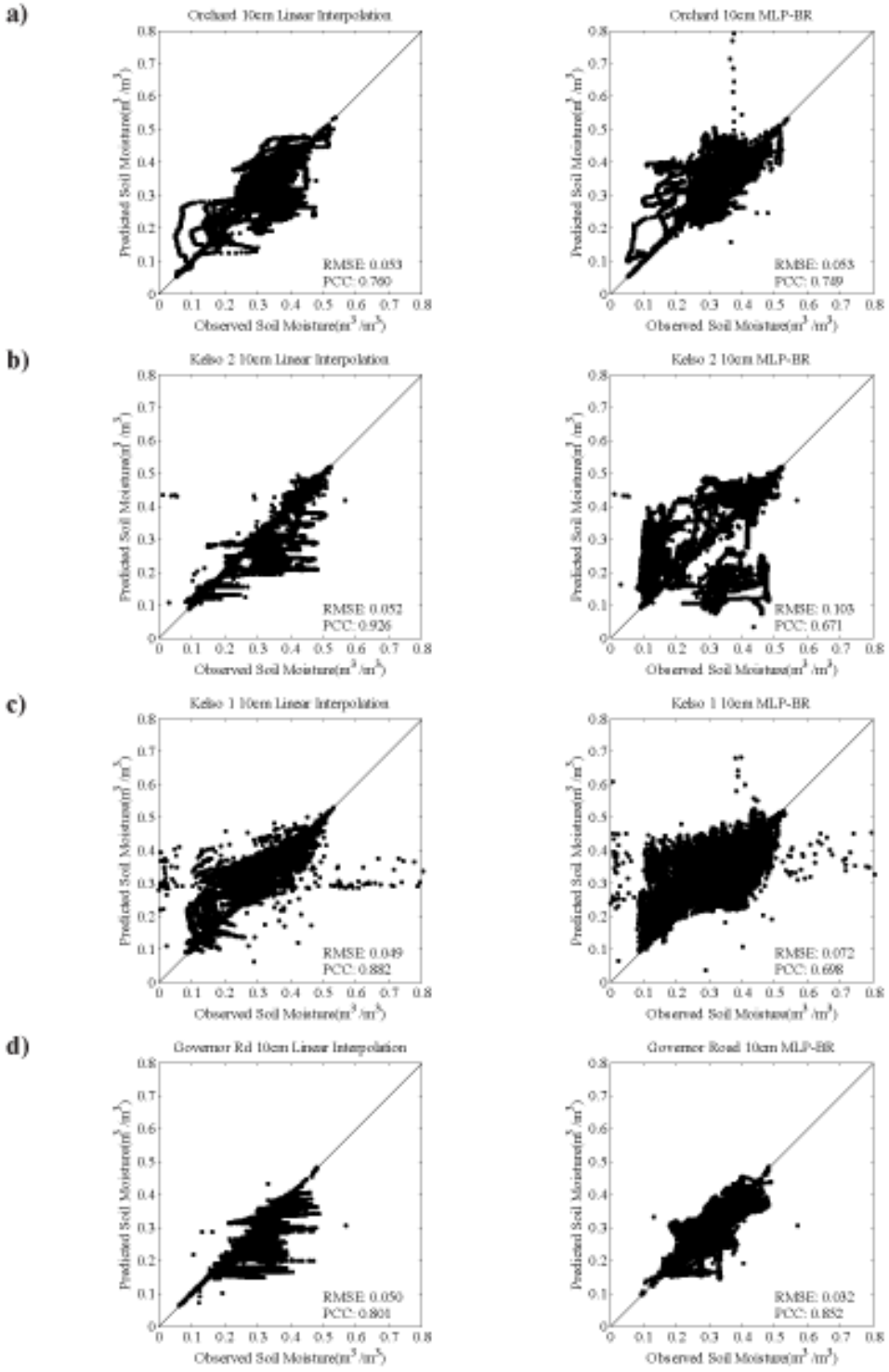

Figure A-6: Scatter plot of observed versus predicted values at 10 cm depth at (a) Orchard; (b) Kelso 02; (c) Kelso 01; (d) Governor Rd when values wre randomly and systematically missing.

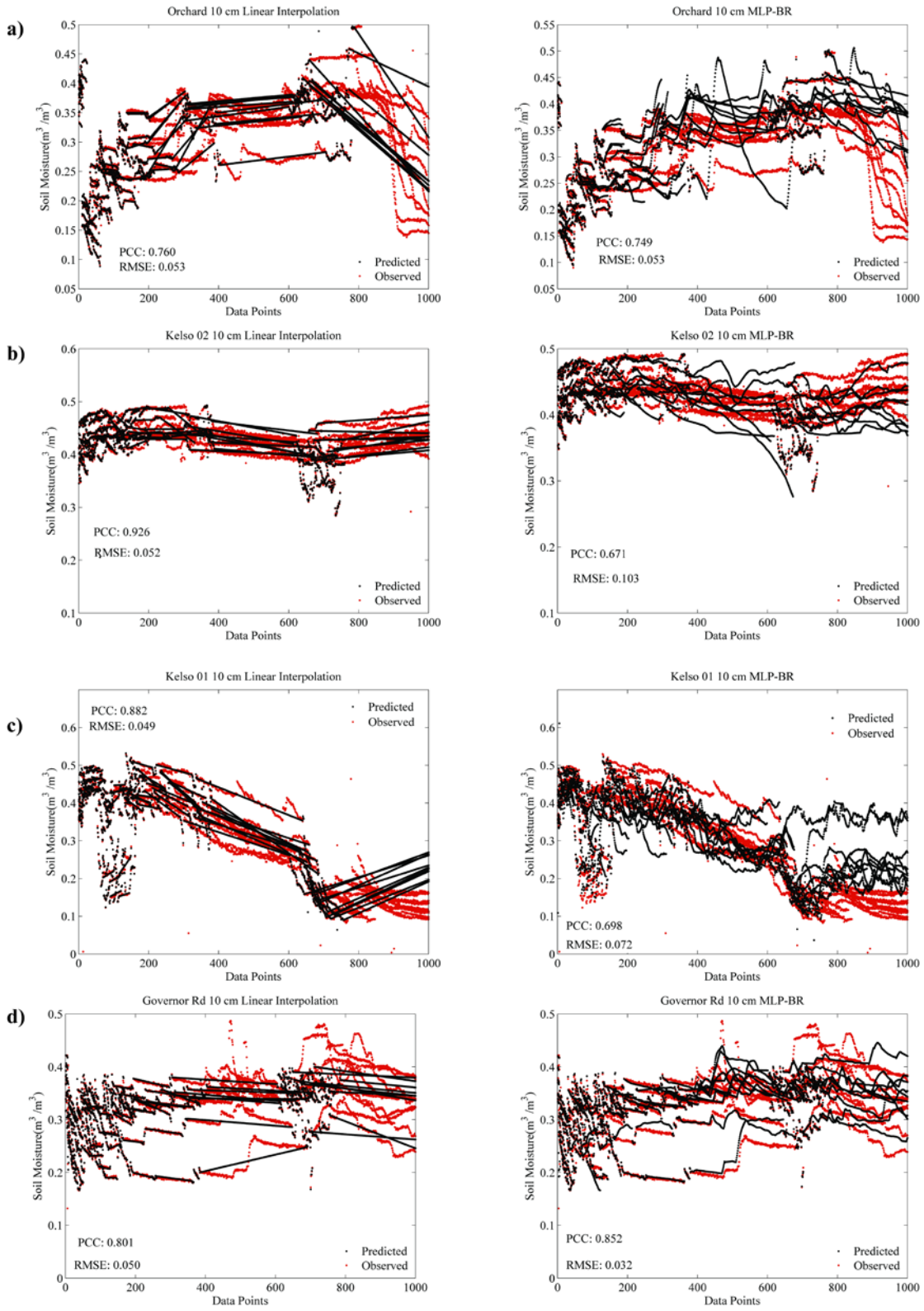

Figure A-7: Data plot of observed and predicted values at 10 cm depth at (a) Orchard; (b) Kelso 02; (c) Kelso 01; and (d) Governor Rd when values are randomly and systematically missing.

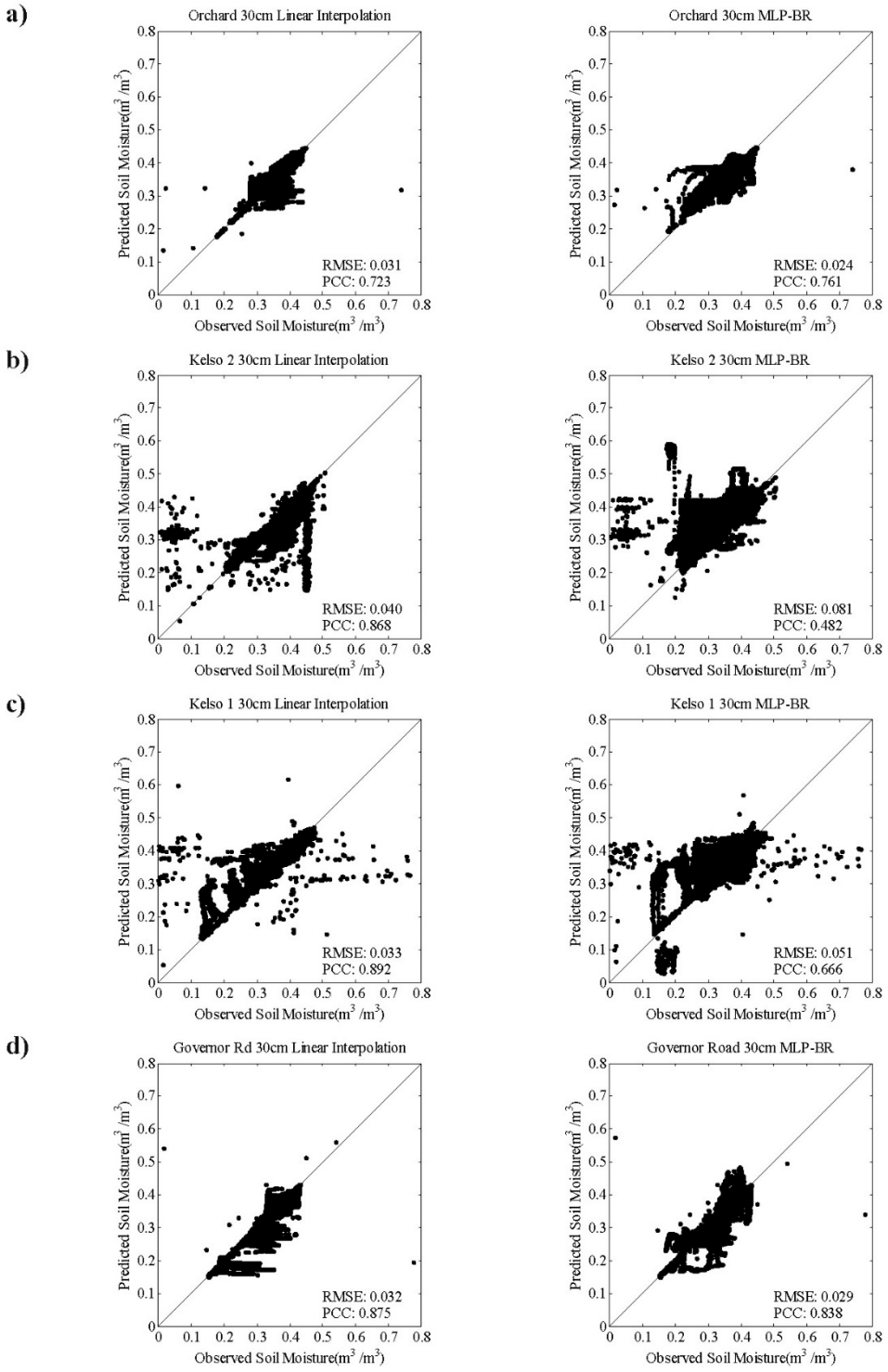

Figure A-8: Scatter plot of observed versus predicted values at 30 cm depth at (a) Orchard; (b) Kelso 02; (c) Kelso 01; (d) Governor Rd when values are randomly and systematically missing.

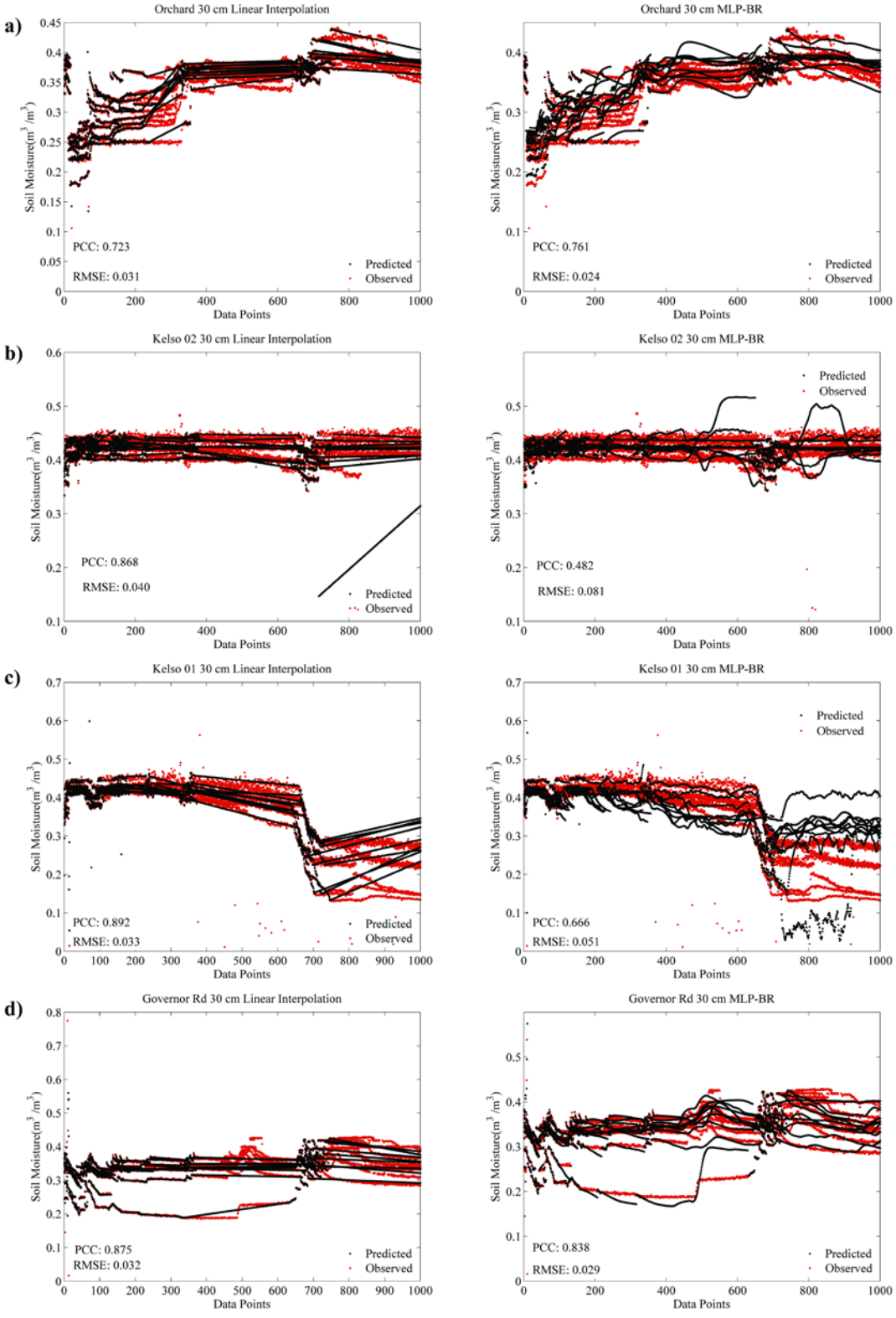

Figure A-9: Data plot of observed and predicted values at 30 cm depth at (a) Orchard; (b) Kelso 02; (c) Kelso 01; (d) Governor Rd when values are randomly and systematically missing.

Table A-6: P-values of stations where the mean and/or variance of the infilled soil moisture from MLP-BR and Linear Interpolation were similar (α = 0.05) to that of the observed soil moisture using the Wilcoxon-Rank Sum and Levene's tests.

Site	Station	Mean P-Value	Variance P-Value
MLP-BR 10 cm			
Orchard	1	-	0.72
	4	-	0.24
	9	0.86	-
Kelso 1	1	0.13	-
Kelso 2	3	-	0.30
	4	0.08	-
	5	-	0.30
Gov Rd	5	0.15	-
	8	0.59	-
Linear Interpolation 10 cm			
Orchard	2	-	0.35
	3	-	0.56
Kelso 1	4	0.13	-
	7	0.38	-
Kelso 2	1	-	0.39
	3	0.12	-
	4	0.96	-
	5	-	0.20
	6	0.66	-
	8	-	0.47
Gov Rd	5	-	0.86
MLP-BR 30 cm			
Orchard	3	0.07	-
	7	0.10	-
Kelso 1	8	0.13	-
	9	0.15	-
Kelso 2	3	0.30	-
Linear Interpolation 30 cm			
Kelso 1	3	0.19	-
	4	0.46	-
	7	0.39	-
Kelso 2	3	-	0.30
	4	0.08	-
	5	-	0.30
Gov Rd	1	0.25	-

For the purposes of this study, it can be shown that as more than approximately 72-100 hours are missing, neither of the selected best methods will accurately infill data. The approximate length that should be infilled without compromising data applicability also changes as the dynamics of the soil changes. When considering linear interpolation, it can be seen that a less dynamic 30 cm soil layer was more accurately represented for longer gaps than the more dynamic 10 cm layer. Similarly, linear interpolation at 10 cm was more appropriate for the less dynamic Kelso sites than for Orchard or Governor Rd. This trend was, however, not applicable when considering MLP-BR, which did not necessarily represent trends in the data as the length of gaps increased.

Generally, as the number of serial missing values increased the various methods infilling ability decreased proportionately, especially with respect to data variance. While API proved to be a suitable alternative to precipitation, the errors caused when the data mining techniques were trained with precipitation is believed to be acute because of the hourly timescale of the data, and would likely not be as prevalent at a daily timescale (Elshorbagy et al. 2010b; Elshorbagy and El-Baroudy 2009; Liu et al. 2008). This demonstrates a difficulty of data mining with too fine a temporal resolution, particularly when considering meteorological variables. At fine temporal

resolutions, both ANN and EPR show heavy reliance on antecedent soil moisture and a low ability to capture the influence of precipitation (API). These results suggest that both the MLP and EPR may not be appropriate for modeling high resolution soil moisture data over long time periods if the modeling interval is short (hours).

A.7 Conclusions

Data-driven methods such as artificial neural networks and evolutionary polynomial regression were tested against conventional infilling methods such as statistical and interpolation infilling methods. Data-driven methods offer the advantage of being derived from the data, and therefore have the potential to better represent underlying physical processes than either statistical methods or interpolation.

Results suggest that data-driven methods may be a suitable alternative to conventional methods of data infilling. Both data-driven and interpolation methods produced similar results (PCC ~0.98, RMSE ~0.01) in terms of infilling randomly missing values and both methods were capable of infilling random missing values with a higher degree of accuracy than statistical methods (MAR PCC ~0.50, SLRD PCC ~0.90). Because of the high serial correlation in hourly soil moisture data, both can be considered to adequately represent physical processes if large serial missing values are not present, although interpolation offers the benefit of parsimony. For the systematic gap treatment, interpolation and ANN proved the best methods for infilling missing soil moisture data. In this study interpolation was always comparable or better than MLP-BR, and displayed the highest skill when the soil was poorly drained with simple topography. The ANN methods skill was greatest when the topography was complex and had good drainage, indicating future work should assess the advantage of ANN with well drained and sandy soils. EPR was an effective method when the length of gaps in the data was short, but provided poor and inconsistent results as the length of gaps increased. The use of the API became necessary when using both data-driven methods in order to produce continuous values for a precipitation related variable. The separation of soil moisture data into those time periods with and without precipitation may improve data-driven model results as it could remove the necessity of API and the influence of a large number of zero values in the training datasets. Overall, the comparative model results showed that for both random and systematic gaps,

classical interpolation techniques and simple MLP models were the most effective infilling techniques for hourly soil moisture data. However, both methods are limited when the serial gaps exceed 72-100 hours.

The McMaster MESONET was designed to fill a knowledge gap between periodic very high spatial resolution soil moisture monitoring (i.e. Famiglietti et al. 2008; Cosh et al. 2004; Western & Grayson, 1998) and long term sparsely distributed monitoring as found in Dorigo et al. (2011). The infilling methods presented herein would be most applicable to the latter, as there is a stronger focus on time-series methods. A particular advantage of interpolation and ANN, as presented, is that given appropriate temporal sampling, missing data can be infilled when there is little spatial connection between monitoring sites. However, these methods require relatively high temporal auto-correlation and may not be effective for very steep slopes, sandy soils, or where influences such as macro-pores result in inconsistent drainage. Similarly, when data are temporally inconsistent, geospatial techniques such as SLRD are more appropriate, but will likely suffer if the soil moisture pattern is highly heterogeneous, as is the case for Governor Rd, or poorly organized.

Acknowledgements

This work was partially funded by the Canadian Foundation for Innovation (CFI), the Natural Science and Engineering Research Council (NSERC) and the Ontario Ministry of Research and Innovation (OMRI). The authors are grateful to Dr. G. Dumedah for assistance in data processing and Dr. O. Giustolisi for technical advice on the use of EPR. This work used EPR software developed by Orazio Giustolisi (Technical University of Bari) and Dragan Savic (University of Exeter). The authors are also grateful to three anonymous reviewers for their comments that helped to improve the manuscript.

A.8 References

ASCE. (1996). Hydrology handbook, American Society of Civil Engineers, New York.

Beale, M., Hagan, M., and Demuth, H. (2010). MATLAB Neural Network Toolbox 7 Users Guide. The Math Works Inc., Natick MA.

Bermudez, J. D., Corberan-Vallet, A., and Vercher, E. (2009). "Forecasting time series with missing data using Holt's model." J.Stat.Plan.Infer., 139(8), 2791-2799.

Casenave, A. (1982). "The small rain-simulator: conditions of use and guidelines for data analysis, Ivory Coast." Cahiers-ORSTOM, Serie Hydrologie, 19(4), 207-227.

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 Sens. Environ., 92, 427-435.

Coulibaly, P., and Evora, N. (2007). "Comparison of neural network methods for infilling missing daily weather records." J. Hydrol., 341(1), 27-41.

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A. and Jackson, T. (2011). "The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements." Hydrol. Earth Syst. Sci. 15, 1675-1698.

Dumedah, G., and Coulibaly, P. (2011). "Evaluation of statistical methods for infilling missing values in high-resolution soil moisture data." J.Hydrol., 400(1-2), 95-102.

Elshorbagy, A., Corzo, G., Srinivasulu, S., and Solomatine, D. P. (2010a). "Experimental investigation of the predictive capabilities of data-driven modeling techniques in hydrology - Part 1: Concepts and methodology." Hydrol.Earth Syst.Sci., 14(10), 1931-1941.

Elshorbagy, A., Corzo, G., Srinivasulu, S., and Solomatine, D. P. (2010b). "Experimental investigation of the predictive capabilities of data-driven modeling techniques in hydrology - Part 2: Application." Hydrol.Earth Syst.Sci., 14(10), 1943-1961.

Elshorbagy, A., and El-Baroudy, I. (2009). "Investigating the capabilities of evolutionary data-driven techniques using the challenging estimation of soil moisture content." J.Hydroinf., 11(3-4), 237-251.

Elshorbagy, A., and Parasuraman, K. (2008). "On the relevance of using artificial neural networks for estimating soil moisture content." J.Hydrol., 362(1-2), 1-18.

Elshorbagy, A., Simonovic, S., and Panu, U. (2000). "Peroformance evaluation of artificial neural networks for runoff prediction." J. Hyrologic Eng., 5(4), 424-427

Entekhabi, D., RodriguezIturbe, I., and Castelli, F. (1996). "Mutual interaction of soil moisture state and atmospheric processes." J.Hydrol., 184(1-2), 3-17.

Famiglietti, J. S., Ryu, D., Berg, A. A., Rodell, M. and Jackson, T. J. (2008). "Field observations of soil moisture variability across scales." Water Resour. Res. 44, W01423.

Fernandez-Galvez, J., Verhoef, A., and Barahona, E. (2007). "Estimating soil water fluxes from soil water records obtained using dielectric sensors." Hydrol.Process., 21(20), 2785-2793.

Gheyas, I. A., and Smith, L. S. (2010). "A neural network-based framework for the reconstruction of incomplete data sets." Neurocomputing, 73(16-18), 3039-3065.

Giustolisi, O., Doglioni, A., and Savic, D. A. (2004a). "Data reconstruction and forecasting by evolutionary polynomial regression." Proceedings of the 6th International Conference on Hydroinformatics, World Scientific Publishing Company, Toh Tuck Link, Singapore, 787-794.

Giustolisi, O., Doglioni, A., and Savic, D. A. (2004b). "Knowledge Discovery by Evolutionary Polynomial Regression." Proceedings of the 6th International Conference on Hydroinformatics, World Scientific Publishing Company, Toh Tuck Link, Singapore, 1647-1654.

Giustolisi, O., Doglioni, A., Savic, D. A., and Webb, B. W. (2007). "A multi-model approach to analysis of environmental phenomena." Environ.Modell.Softw., 22(5), 674-682.

Giustolisi, O., and Savic, D. A. (2003). "Evolutionary Polynomial Regression (EPR): Development and applications." Rep. No. 2003/10, School of Engineering, Computer Science and Mathematics, Centre for Water Systems, University of Exeter, England.

Haykin, S. (1999). Neural Networks: A Comprehensive Foundation. Prentice Hall, Upper Saddle River, NJ.

Hagan, M. T., and Menhaj, M. B. (1994). "Training Feedforward Networks with the Marquardt Algorithm." IEEE Trans.Neural Networks, 5(6), 989-993.

Hopke, P. K., Liu, C. H., and Rubin, D. B. (2001). "Multiple imputation for multivariate data with missing and below-threshold measurements: Time-series concentrations of pollutants in the Arctic." Biometrics, 57(1), 22-33.

Jain, S., Singh, V., and van Genuchten, M. (2004). “Analysis of soil water retention data using artificial neural networks.” J. Hydrologic Eng. 9(5), 415-420.

Junninen, H., Niska, H., Tuppurainen, K., Ruuskanen, J., and Kolehmainen, M. (2004). "Methods for imputation of missing values in air quality data sets." Atmos.Environ., 38(18), 2895-2907.

Khalil, M., Panu, U. S., and Lennox, W. C. (2001). "Groups and neural networks based streamflow data infilling procedures." J.Hydrol., 241(3-4), 153-176.

Köhne, J. M., Wöhling, T., Pot, V., Benoit, P., Leguédois, S., Bissonnais, Y. L., and Šimůnek, J. (2011). "Coupled simulation of surface runoff and soil water flow using multi-objective parameter estimation." J. Hydrology, 403(1-2), 141-156.

Liu, X., Coulibaly, P., and Evora, N. (2008). "Comparison of data-driven methods for downscaling ensemble weather forecasts." Hydrol.Earth Syst.Sci., 12(2), 615-624.

Mackay, D. J. C. (1992). "A Practical Bayesian Framework for Backpropagation Networks." Neural Comput., 4(3), 448-472.

Mair, A., and Fair, A. (2009). "Assessing rainfall data homogeneity and estimating missing records in Mākaha Valley, O'ahu, Hawai'i." J. Hydrologic Eng., 15(1), 61-66.

Markewitz, D., Devine, S., Davidson, E. A., Brando, P., and Nepstad, D. C. (2010). "Soil moisture depletion under simulated drought in the Amazon: impacts on deep root uptake." New Phytol., 187(3), 592-607.

Monteith, J., L., (1965). "Evaporation and environment". Symp. Soc. Exp. Biol.19, 205-34

Murtagh, F., Zheng, G., Campbell, J. G., and Aussem, A. (2000). "Neural network modelling for environmental prediction." Neurocomputing, 30(1-4), 65-70.

Ng, W., Panu, U., and Lennox, W. (2009). "Comparative studies in problems of missing extreme daily streamflow records." J. Hydrologic Eng. 14(1), 91-100.

Rollenbeck, R., and Anhuf, D. (2007). "Characteristics of the water and energy balance in an Amazonian lowland rainforest in Venezuela and the impact of the ENSO-cycle." J. Hydrol., 337(3-4), 377-390.

Savic, D. A., Giustolisi, O., and Laucelli, D. (2009). "Asset deterioration analysis using multi-utility data and multi-objective data mining." J.Hydroinf., 11(3-4), 211-224.

Schwarzel, K., Menzer, A., Clausnitzer, F., Spank, U., Hantzschel, J., Grunwald, T., Kostner, B., Bernhofer, C., and Feger, K. H. (2009). "Soil water content measurements deliver reliable estimates of water fluxes: A comparative study in a beech and a spruce stand in the Tharandt forest (Saxony, Germany)." Agric.for.Meteorol., 149(11), 1994-2006.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J. (2010). "Investigating soil moisture-climate interactions in a changing climate: A review." Earth-Sci.Rev., 99(3-4), 125-161.

Ulke, A., Tayfur, G., and Ozkul, S. (2011). "Predicting suspended sediment loads and missing data for Gediz River, Turkey." J. Hydrologic Eng. 14(9), 954-965.

Vachaud, G.A., Passerat de Slians, A. Balabanis, P., and Vauclin, M. (1985). "Temporal stability of spatially measured soil water probability density function." Soil Sci. Soc. Am. J., 49(4), 822-828.

Western, A., W., Grayson, R., B., and Blöschl, G. (2002). "Scaling of soil moisture: A hydrological perspective". Annu. Rev. Earth Planet. Sci. 30, 149-180.

Western, A., W., and Grayson, R., B. (1998). “The Tarrawarra data set: soil moisture patterns, soil characteristics and hydrological flux measurements. Water Resour. Res. 34, 2765-2768.

Western, A., W., Blöschl, G., and Grayson, R., B. (1998). “Geostatistical characterization of soil moisture patterns in the Tarrawarra catchment.” J. Hydrol. 205, 20-37.

Wu, C. L., Chau, K. W., and Fan, C. (2010). "Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques." J. Hydrol., 389(1-2), 146-167

Appendix B: Advances in Soil Moisture Retrieval from Synthetic Aperture Radar and Hydrological Applications

Summary of Appendix B: Kornelsen, K.C. and Coulibaly, P. (2013) Advances in Soil Moisture Retrieval from Synthetic Aperture Radar and Hydrological Applications, Journal of Hydrology, 476, 460-489, doi: 10.1016/j.jhydrol.2012.10.044.

Summary:

This paper is a comprehensive review of the state of the art of high resolution soil moisture retrievals from synthetic aperture radar (SAR) satellites with a particular emphasis on the operational potential of SAR retrieved soil moisture. The review consists of two parts. The first surveys the technical aspects of soil moisture retrieval, including the classification of surface roughness, radiative transfer modelling of soil and vegetation, quantification of uncertainty and the use of multi-configuration SAR data. The second part focused on the hydrological applications of high resolution soil moisture information, particularly with respect to data quality from SAR, retrieval errors, issues surrounding the scaling properties of soil moisture and the use of multiple sources of data.

The key findings were:

- Retrieval of soil moisture from SAR is a difficult task because more parameters/variables are required to characterize the surface conditions than information available from a single observation.
- This difficulty may be overcome by forthcoming SAR constellations, although long revisit times may still limit SAR retrieved soil moisture in terms of operational applications.
- The noise inherent in SAR images requires filtering which reduces the spatial resolution to a similar order of magnitude as passive microwave observations.
- High resolution soil moisture is an important hydrological variable and will benefit hydrological applications more than coarse resolution soil moisture.
- The skill of the soil moisture retrieval is less limiting to hydrological applications, particularly data assimilation, than the spatial / temporal availability of the retrievals.
- The best application of SAR is the determination of the distribution of soil moisture inside a radiometer footprint as will be done with SMAP.

B.1 Abstract

The sensitivity of Synthetic Aperture Radar (SAR) to soil moisture is well established, however, the retrieval of soil moisture from SAR is confounded by the effects of surface roughness and vegetation. This difficulty has resulted in limited applications of SAR as an operational source of soil moisture in hydrology despite the demonstrated benefits of high-resolution distributed soil moisture. Technical and methodological advances such as multi-configuration radar and forthcoming SAR constellations are increasingly mitigating the shortcomings of SAR with respect to soil moisture estimation at the field and catchment scale. At the same time, progress in data assimilation and a better understanding of the impact of phenomena, such as climate change, are revealing the hydrological importance of soil moisture spatial distribution. Thus, despite the currently modest retrieval accuracy, SAR is an important source of soil moisture state information for the hydrological community. Towards the end of increasing the hydrological utilization of SAR soil moisture, a comprehensive literature review was conducted to provide the state-of-the-art of SAR soil moisture retrieval methodology, its limitations and potential. Following the methodology review, a discussion of the benefits and limitations of soil moisture data retrieved from SAR is used to outline the scope of SAR derived soil moisture for hydrological applications.

B.2 Introduction

The spatial and temporal distribution of soil moisture is a key state variable in various hydrological and meteorological applications. From the perspective of climate science, soil moisture and the associated flux between the soil and the atmosphere play an important part in the Earth's climate regimes. The coupling that takes place between soil moisture and the atmosphere can have profound impacts on the planet's climate systems, especially when the role of vegetation is considered (see Seneviratne et al. 2010 for a review). In hydrologic studies, soil moisture is a critical component as it controls the partitioning between infiltration and runoff, where infiltration determines the amount of water available for vegetation growth and runoff has a strong impact on the rate of surface erosion and river processes. Combined, the hydrological and climatological processes influenced by soil moisture can impact many environmental

phenomena from extreme events like droughts and flooding to state patterns such as the ecological distribution of homogenous vegetation zones. In each of these cases knowledge of the distribution and amount of water in the soil can aid developing better modelling and decision support tools.

Despite the many advantages that can be derived from the knowledge of soil moisture distribution, measurement of soil moisture at the field or watershed scale remains difficult. This is largely due to the difficulty and cost associated with obtaining spatially representative *in situ* soil moisture measurements by point scale sampling. Such methods are labour intensive and impractical to be carried out in more than a few watersheds worldwide.

Remote sensing offers a potential alternative for characterizing the distribution and quantity of soil moisture at a variety of scales, without expensive *in situ* monitoring networks. Specifically, microwave remote sensing, both active and passive, is influenced by the dielectric properties of the soil, and therefore soil moisture. The relationship between the radar signal and soil moisture must be separated from that of other influences that affect radar signals.

The recent launch of the Soil Moisture and Ocean Salinity (SMOS) satellite has made possible operational monitoring of soil moisture using passive remote sensing (Kerr et al. 2010). The emission behaviour of the surface is such that the typical resolution of a passive satellite sensor is on the order of tens of kilometers and even modern sensors such as SMOS have a spatial resolution of 40 km (Kerr et al. 2010). This coarse spatial resolution limits applicability of passive radar sensors at the watershed scale and does not account for the spatial dynamics of soil moisture.

In contrast Synthetic Aperture Radar (SAR) uses a self-contained source of microwave radiation to illuminate the surface and measures the amount of radiation returned to the sensor (Fung, 1994; Ulaby et al. 1982). This allows SAR to monitor surface characteristics, including soil moisture, at a spatial resolution of meters to tens of meters under almost all weather conditions (Ulaby et al. 1982). Despite providing remote sensing capability at high spatial resolutions, retrieval of soil moisture from SAR sensors is a challenging problem due to the inherent complexity of radar backscattered signal and the inverse problem of backscattering models (Loew and Mauser, 2006; Mattia et al. 2006; Shi et al. 1997), and therefore has been the topic of much research over the past twenty years. This difficulty, combined with the prevalence

of passive soil moisture products, has restricted research activities to identify the potential benefits of high resolution surface soil moisture using actual SAR data.

This review analyzes the potential for retrieval of soil moisture from SAR satellites to enhance research and operational hydrology. It is broken down into two parts. The first highlights the influence of soil and radar properties on radar backscattering and discusses the most commonly used backscatter models, their advancements and limitations, followed by an overview of techniques used to invert those models for separating soil moisture from other influences. This will provide important background for the hydrologist to identify appropriate methodologies for incorporation of SAR as a data source. The second section focuses on the impact of soil moisture for hydrological applications. It discusses the variability of soil moisture and its distribution at the spatial resolution of SAR sensors, the potential benefits of high resolution distributed soil moisture in hydrological modelling and makes practical recommendations to overcome some of the shortcomings associated with SAR soil moisture retrieval.

B.3 Soil Moisture Retrieval Methods

B.3.1 Background of Soil and Radar Properties

The radio portion of the electromagnetic spectrum are well suited for remote sensing purposes, as longer wavelengths are known to penetrate clouds and light rain, making radar an ideal choice for all-weather sensing (Fung, 1994; Ulaby et al. 1981). The microwave spectrum extends from 0.3 GHz or 100 cm to 100 GHz or 0.3 cm. For remote sensing applications, the most common bands are the longer wavelength L band (0.39-1.55 GHz), the compromise C band (3.9-5.75 GHz) and the short wavelength X band (5.75-10.9 GHz) (Ulaby et al. 1981).

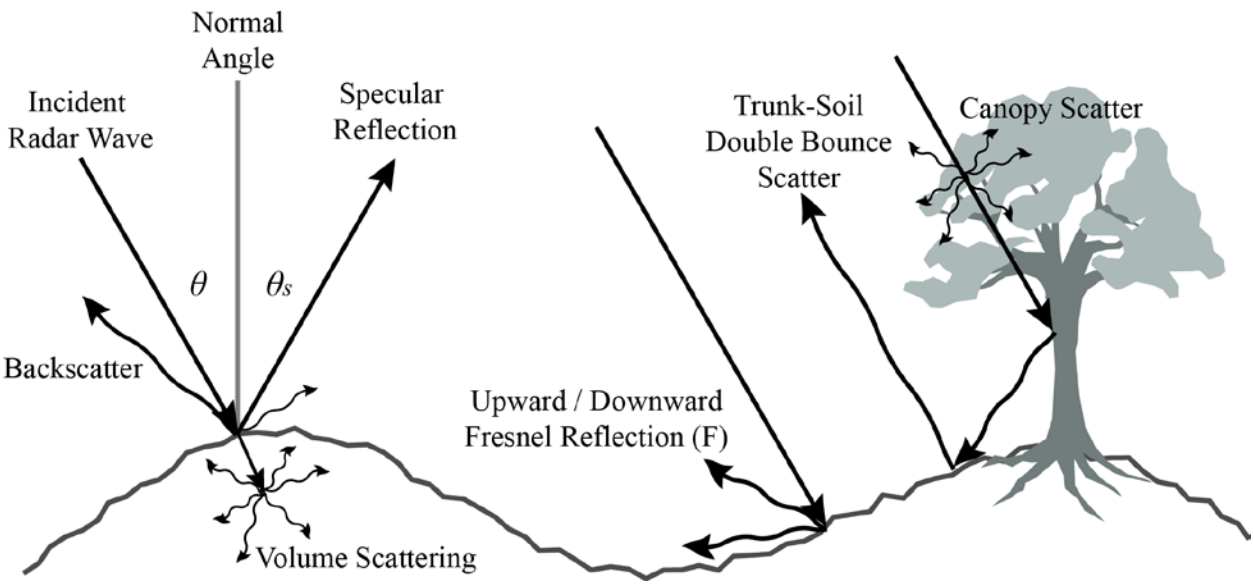

Figure B-1: Conceptual scatting of an incident radar signal by a rough and vegetated surface. The Fresnet reflection shown is important in the formulation of the AIEM.

Upon contact with the surface, the incident energy may be absorbed (attenuated), scattered, reflected or penetrate into the medium and be further absorbed or scattered as in Figure B-1. The amount of scattering which returns to the sensor (backscatter) are collected for interpretation as an image. The incident radar signal is polarized in either horizontal (H) or vertical (V) fashion and is observed as backscatter in the same way. The result is either co-polarized (HH or VV) or cross-polarized (HV or VH) backscatter. The amount of energy backscattered from a natural surface is dependent on the radar configuration, soil characteristics and vegetation characteristics. The backscattering coefficient σ^0 at incoming and outgoing polarization p is function of the radar incidence angle θ, polarization and wavelength λ as well as the surface root mean square height σ (also denoted as s, RMS and RMS height), the surface correlation length l (also denoted as L), the dielectric properties of the soil ε^r and vegetation ε^v, as well as the shape, structure and angle of vegetation as a scatterer and reflector θ_{veg}. Since surface roughness parameters are sensitive to the wavelength at which they are perceived, they are often normalized by the wave number k. It is the influence and interaction of these various components, combined with the ability to penetrate the vegetation and surface, which makes interpretation of radar remote sensing difficult. Despite the difficulties, with proper treatment the dielectric constant ε^r can be isolated and soil moisture determined from the radar signal (Fung, 1994; Ulaby et al. 1986). Table B-1 provides a list of past, present and proposed spaceborne SAR sensors as well as their specific attributes for reference, and Table B-2 provides a summary

of selected advances in soil moisture retrieval. The authors acknowledge that Table B-2 is by no means exhaustive, but hope it will serve to direct readers to appropriate references.

Table B-1: Spaceborne synthetic aperture radar (SAR) platforms and their respective attributes.

Platform/Sensor	Operational Dates	Radar Configuration
SEASAT	Launch: June 1978 Completion: October 1978	Band: L Resolution: 25m Incidence Angle: 20-26° Polarization: HH
SIR-A **(Shuttle Imaging Radar)**	Launch: November 1981 Completion: November 1981	Band: L Resolution: 40m Incidence Angle: 50° Polarization: HH
SIR-B **(Shuttle Imaging Radar)**	Launch: October 1984 Completion: October 1984	Band: L Resolution: 20-30m Incidence Angle: 15-65° Polarization: HH
ERS-1 **(European Resource Satellite -1)**	Launch: July 1991 Completion: March 2000	Band: C Resolution: 30m Incidence Angle:23° Polarization: VV
JERS-1 **(Japanese Earth Resources Satellite)**	Launch: February 1992 Completion: October 1998	Band: L Resolution:18m Incidence Angle: Polarization: HH
SIR-C/X-SAR **Space Shuttle** **(Shuttle Imaging Radar)**	Launch: April/October 1994 Completion: April/October 1994	Band: L, C, X Resolution:30m Incidence Angle:17 to 63° Polarization: HH, VV, HV, VH
ERS-2 SAR **(European Resource Satellite -2)**	Launch: April 1995 Completion: July 2011	Band: C Resolution: 30m Incidence Angle:23° Polarization: VV
RADARSAT 1	Launch: November 1995	Band: C Resolution 8-100m Incidence Angle: 10-59° Polarization: HH
ENVISAT ASAR **(Advanced Synthetic Aperture Radar)**	Launch: March 2002 Completion: April 2012	Band: C Resolution:30-1000m Incidence Angle:14-45° Polarization: HH,VV,HV,VH Two per acquisition
ALOS –PALSAR **(Phased Array L-band Synthetic Aperture Radar)**	Launch: January 2006 Completion: May 2011	Band: L Resolution:7-100m Incidence Angle: 8-60° Polarization:VV,HH,VH,HV
Terra SAR-X	Launch: 2007	Band: X Resolution: 1-16m Incidence Angle: 20-60° Polarization: VV, HH, VH, HV
RADARSAT 2	Launch: December 2007	Band: C Resolution: 3-100m

		Incidence Angle: 20-50° Polarization: HH,VV,HV,VH Simultaneously
COSMO-SkyMed (4 Satellites) (Constellation of Small Satellites for Mediterranean basin Observation)	Launch: 2007-2010	Band: X Resolution: 1-100m Incidence Angle: 20-50° Polarization: HH,VV,HV,VH
Huan Jing (HJ) Constellation (4 Satellites)	Launch: 2008-2012	Band: S Resolution: 20m Incidence Angle: 25-47° Polarization: VV
Sentinel-1 (2 Satellites)	Launch: 2013-2015	Band: C Resolution: 4 x 5m – 25 x 80m Incidence Angle: 20-45° Polarization: VV,HH,HV,VH
ALOS 2- PALSAR-2	Launch: 2013	Band: L Resolution: ~1m – 100m Incidence Angle:8-70° Polarization: VV,HH,HV,VH
SMAP (Soil Moisture Active Passive)	Launch:2014-2015	Band: L Resolution: 3-10km Incidence Angle: 35-50° Constant Polarization: VV,HH,HV,VH
RADARSAT Constellation (3-6 Satellites)	Launch: 2014-2015	Band: C Resolution: 3-100m Incidence Angle: Polarization: HH,VV,HV,VH Simultaneously
SAOCOM (4 Satellites) Satélite Argentino de Observación Con Microondas	Launch: 2014-2017	Band: L Resolution: 10 – 100m Incidence Angle: 15-50° Polarization: HH,VV or HH/VV,HH/HV,VV/VH
COSMO-SkyMed (2nd Gen) (2 Satellites)	Launch: 2015-2016	Band: X Resolution: 1-35m Incidence Angle: 20-50° Polarization: HH,VV,HV,VH

Table B-2: Summary of select advances in soil moisture retrieval.

Reference	Study Area	Data	Methods	Major Findings
Davidson *et al*. 2000	Marestaing, France; Basilicata, Italy	CESBIO- ESA Laser Profiler	- Determined ACF's based on long profiles	- Gaussian ACF represents rough soils at short intervals - Exponential ACF represents rough soils at long intervals - Fractal processes are introduced at longer profile lengths
Moran *et al*. 2000	Upper San Pedro Basin, AZ, USA	ERS 2 - C Band - 23° incidence - VV polarization	- Registered SAR image to a dry date to remove the effects of roughness - Retrieved soil moisture with least squares fitting	- Registering to a dry date removed the effects of roughness allowing for a simple linear model to be effective - The relationship between backscatter and soil moisture was weak - The relationship between dry registered backscatter and soil moisture was strong
Zribi *et al*. 2000	Orgeval, France	ERASME SIR-C - C band - 25-50° incidence - Multiple polarizations	- Defined empirical fractal correlation function - Compared IEM and Moment Method simulations using fractal ACF to observed backscatter	- Introduced fractal dimension to ACF - Fractal ACF can represent Gaussian and exponential ACF - Retrieval improved with fractal ACF
Bindlish and Barros, 2001	Washita 94 Experiment, Little Washita, USA	SIR-C and X-SAR - C and X Band - Multiple configurations	- Combined watercloud model and IEM - Watercloud parameterized with NDVI	- Cross polarizations increase the relationship between NDVI and backscatter - Separating watercloud model by vegetation type improves soil moisture retrieval
Baghdadi *et al*. 2002c	Pays de Caux, France	ERS & RADARSAT 1 - C Band - 23-47° incidence - HH polarization	- Inverted IEM with least squares fitting	- At high incidence angles surface roughness dominates - Row direction influence stronger at high angles - Proposed optimal correlation length
Zribi and	Orgeval, France;	SIR-C	- Introduced new parameter to invert surface roughness	- Introduced Zs parameter - Defined an empirical relationship between Zs and

Deschambre, 2002	Pays de Caux, France; Alpilles-Reseda, France	RADARSAT ERASME - C Band - Multiple incidence - Multiple polarization		the change in backscatter between high and low incidence angles
Srivastava *et al*. 2003	Agra, Mathura and Bharatpur, India	RADARSAT-1 - C Band - 10-23° and 41-46° incidence - HH polarization	- Retrieved soil moisture with an empirical model - Used low and high incidence to correct for surface roughness	- Combination of low and high incidence angles removes the effect of surface roughness - Empirical model correlation increased significantly when roughness term used - The effect of soil moisture at different angles is negligible compared to surface roughness
Baghdadi *et al*. 2004	Pays de Caux, France; Rhone Valley, France; Orgeval, France; Chateauguay, Canada; Riviere aux Brochets, Canada	ERS RADARSAT-1 SIR-C X-SAR,ERASME - Multiple configurations	- Performed a semi-empirical calibration of IEM - Replaced measured correlation length with calibration function	- With the calibration method the surface roughness can be characterized with only the RMS height - Fractal ACF provided best results
Manninen *et al*. 2005	Suonenjoki, Finland	ENVISAT ASAR - C Band - Multiple incidence - Multiple polarization	- Compared SPOT NDVI and Depolarization ratio for estimating LAI	- LAI from ASAR better than LAI from SPOT - HH/VV polarization ratio is sensitive to LAI
Alvarez-Mozos *et al*. 2006	La Tejeria Watershed, Navarre, Spain	RADARSAT 1 - C Band - Low incidence - HH polarization	- Watercloud model parameterized with NDVI from Landsat-7 - Coupled watercloud and IEM - IEM inverted through Newton-	- There is great dispersion between IEM simulations and observations at the field scale, owing to the spatial variability of surface roughness - At the catchment scale IEM/watercloud model agreed with observations

			Raphson iterative method	
Baghdadi *et al*. 2006	Villamblain, France; Toulouse, France;	ENVISAT ASAR - C Band - Multiple incidence - HH, HV, VV polarization	- Inverted IEM with a least squares fitting - Use multi-polarization, multi-angle data for inversion - Used calibrated correlation length	- Multi-polarization does not result in significant improvement in soil moisture retrieval - Using high and low incidence angles does result in improved soil moisture retrieval - Combined high and low angles mitigates the effects of roughness
Callens *et al*. 2006	Ghent University Experimental Farm, Melle, Belgium	CESBIO- ESA Laser Profiler	- Decomposed 25m transects into 4m transects - Detrended profile to remove topographic effects	- Change in surface roughness from precipitation only significant if recently tilled - For long profiles RMS height is stable while correlation length varies - Correct correlation length can be determined from short profiles if fitted to a exponential relationship
Lakhankar *et al*. 2006	SGP97 Oklahoma, USA	RADARSAT-1 - C Band - HH polarization	- Compare Neural Networks and Fuzzy Logic for IEM inversion - NDVI to parameterize vegetation	- Neural Networks provide better results, but required multiple training - Fuzzy Logic produced less accurate but more consistent results
Mattia *et al*. 2006	Matera, Italy	ENVISAT ASAR - C Band - 15-31° incidence - HH and VV polarization	- Soil moisture was retrieved using a Variational assimilation method with IEM and GO model - A Thornthwaite-type model is used for *a priori* information	- Low quality *a priori* information can reduce retrieval error by approx 4% whereas high quality *a priori* information can reduce error by 8% - *A priori* guess of surface roughness also improved retrieval
Notarnicola *et al*. 2006	SMEX 02 Walnut Gulch, USA	AirSAR - C and L Bands - 40° incidence - HH and VV polarization	- Invert IEM with a Bayesian method - Remove vegetation effect with NDVI and NDWI	- Vegetation removal is more effective at C rather than L band - Double bounce is more pronounced at L band decreasing retrieval capacity - NDWI removes the vegetation effect better than NDVI
Rahman *et al*. 2007	Walnut Gulch, USA	RADARSAT-1 - C Band - 47° incidence	- Determine the relationship between surface roughness and correlation length by using dry images - Inverted IEM with look-up	- Correlation length varies with surface roughness - Correlation length can be accurately inverted with only RMS height if the surface is dry

		- HH polarization	table	
Srivastava *et al*. 2006	Agra, Mathura and Bharatpur, India	RADARSAT-1 - C Band - 10-23° incidence - HH polarization	- Retrieve soil moisture with a least squares empirical model	- Retrieved soil moisture had the lowest correlation with gravimetric and volumetric water content and the highest correlation with free water
Verhoest *et al*. 2007	Randomly generated synthetic data	ERS Configuration adopted - C Band - 15-30° incidence - VV polarization	- Trained Takagi-Sugeno Fuzzy Model using the possibility distributions to retrieve soil moisture	- Fuzzy approach was comparable to possibilistic approach results - Fuzzy approach is computationally efficient - Soil moisture could be retrieved with uncertainty bounds
Notarnicola *et al*. 2008		U of Bern Truck-mounted scatterometer - Multiple configurations to simulate ERS and ASAR data	- Compared Bayesian and Artificial Neural Networks from retrieving soil moisture	- The ANN inversion approach is slightly better than the Bayesian approach - Errors are reduced when multiple angles are used - Increasing the number of parameters increases the ANN retrieval accuracy while decreasing the Bayesian retrieval accuracy
Rahman *et al*. 2008	Walnut Gulch, USA	ENVISAT ASAR - C Band - 24-41° incidence - VV polarization	- Use low/high incidence images to invert surface roughness when the surface is dry	- RMS height and correlation length can be derived separately from Zs - Produced map of surface roughness and correlation length - Soil moisture could be retrieved with good accuracy at the watershed scale, but results were poor at the field scale
Said *et al*. 2008	Roorkee, India	ERS-2 - C Band - VV	- Used artificial neural networks and multiple inputs to invert soil moisture	- Neural network inversion was better than conventional or multiple linear regression - LAI better represents vegetation effects than

		polarization		vegetation water content and crop height - Combining data from multiple seasons limited the effectiveness of soil moisture retrieval
Thoma *et al.* 2008	Walnut Gulch, AZ; Little Washita, OK; Little River, GA, USA	RADARSAT-1 - C Band - 46° incidence - HH polarization	- To reduce speckle in soil moisture retrieval using median filtering and spatial averaging prior to retrieval	- Spatial filtering increased the accuracy of retrieval at the cost of precision - Optimum ground area is 25 to 160 times the SAR spatial resolution - Smallest effective ground resolution was 162m to 1131m depending on variability and roughness
Lievens *et al.* 2009	Synthetic soil profiles	Simulated ENVISAT ASAR and ALOS PALSAR - C and L Band - Multiple incidence - Multiple polarization	- Simulated surface data to test the influence of various surface conditions on soil moisture retrieval	- Error in RMS height has a greater impact than error in correlation length - C band retrieval is less sensitive to inaccuracy than L band retrieval - Impact of surface roughness errors increase with soil moisture content - Polynomial removal of roughness trends is best solution
Pathe *et al.* 2009	Oklahoma, USA	ENVISAT ASAR - GM Mode - C Band - HH polarization	- Validate application of ERS Radiometer change detection algorithm to ASAR Global Mode (GM)	- Correlation with soil moisture is weaker than with radiometer because of noise in the data - Spatial averaging improves correlation results
Huang *et al.* 2010	Agricultural Fields Ypsilanti, Michigan	Michigan POLAR-SCATTER	- Used Numerical Maxwell Model in 3D (NMM3D) which is a MoM model to create a LUT for soil moisture inversion	- Inversion of the MoM produced soil moisture retrieval results that were comparable or better than AIEM inversion - With new algorithms MoM can be used as a forward model
Joseph *et al.* 2010	OPE3 Experiment Beltsville, MA, USA	NASA/GWU Truck mounted Scatterometer - C and L Band	- Compare ratio method and watercloud model - Invert IEM for soil moisture	- Soil moisture retrieval with ratio method was more accurate than with watercloud model - As vegetation water content increases beyond 0.2 kgm^{-1} measured backscatter exceeds IEM simulation - There is a greater difference in backscatter from vegetation for VV than HH polarization

		- Multiple incidence - HH, VV, HV, VH polarization		
Lawrence *et al.* 2011	Synthetic Data		- Test a FEM Model against MoM model for rough surface scattering and emission	- FEM is not a accurate as MoM but offers but alternative for multi-layered structures - More research on numerical models is needed

B.3.1.1Soil Dielectric Properties

If other factors are ignored, the backscattering coefficient increases with an increase in soil moisture until the moisture content reaches ~35 % vol. when the radar signal becomes insensitive to soil moisture (Dobson and Ulaby, 1981; Ulaby et al. 1986). The dielectric behaviour of the soil is also influenced by the distribution of the grain sizes through the amount of free water as shown in Figure B-2 (Hallikainen et al. 1985; Mironov et al. 2004; Srivastava et al. 2006). Sandy soils have a higher amount of free water than clay soils which results in the correlation between backscatter and soil moisture being higher in sandy soils (Blumberg et al. 2000; Kong and Dorling, 2008; Srivastava et al. 2006; Walker et al. 2004). This occurs because, in clay soils the water molecules are adsorbed onto the soil particles effectively immobilizing their dipoles, disallowing bound water to interact with the radar signal (Jackson and Schmugge, 1989). With this consideration in mind, Srivastava et al. (2006) found that soil moisture retrieval had a correlation of 0.96 when soil moisture was converted to free water, but was only 0.88 when gravimetric water content was considered.

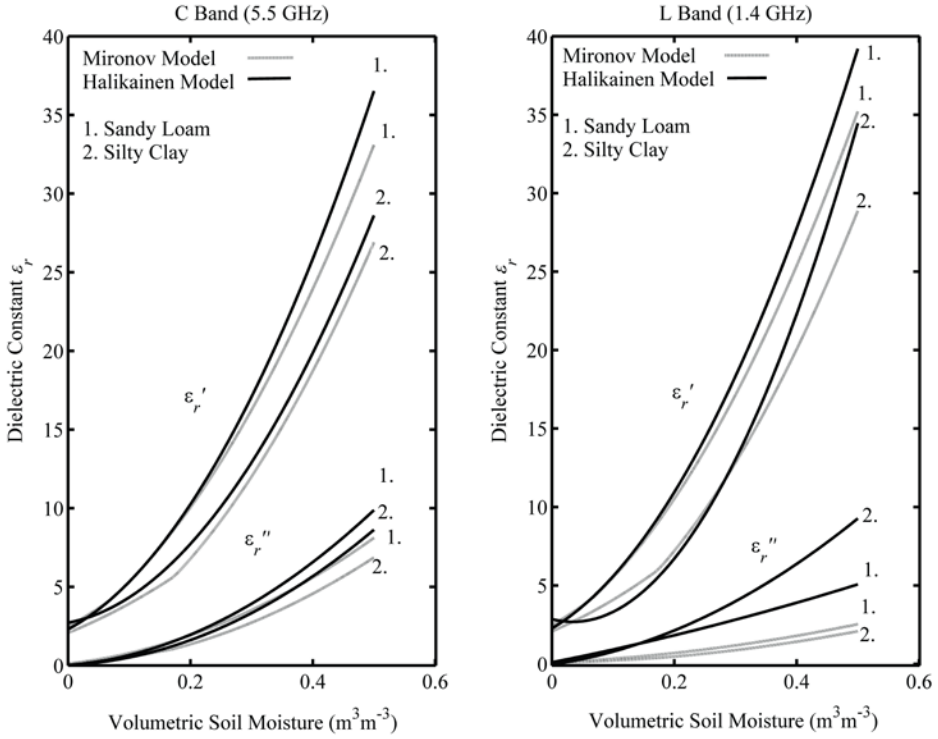

Figure B-2: Comparing the behaviour of the real (e_r') and imaginary (e_r'') part of the dielectric constant as a result of soil moisture using the empirical model of Hallikainen et al. (1985) and the MBSDM of Mironov et al. (2009).

The most common methods found in the literature to relate the soil moisture and the dielectric constant are the empirical curves of Hallikainen et al. (1985), Dobson et al. (1985) and Peplinski et al. (1995). While these models are simple to use, they do not account for the changes in dielectric properties with bound and free water (Mironov et al. 2004). In order to address this shortcoming the Generalized Refractive Mixing Dielectric Model (GRMDM) was developed, which models the soil complex dielectric constant as a function of frequency and both free and bound soil water (Mironov et al. 2002; Mironov et al. 2004). To improve model convenience, the GRMDM was numerically fitted using clay content as the sole soil parameter, resulting in the Mineralogically Based Spectroscopic Dielectric Model (MBSDM), which was found to be more accurate and have a greater validity range than comparable models (Mironov et al. 2009).

Radar signal is also known to penetrate the soil surface based on its wavelength and the amount of soil moisture. Long wavelengths penetrate further into the soil than short wavelengths, and the amount of penetration increases as the soil dries. To account for this, sampling should be done at the penetration depth, which is usually assumed to be represented by the mean to fixed depth of 5cm for a vertically homogenous profile (Boisvert et al. 1997). However, if a vertical soil moisture gradient is present mean sampling may result in backscatter overestimation (Boisvert et al. 1997). The assumption of vertically homogeneous soil moisture distribution may be a factor limiting the development of operational SAR soil moisture retrieval (Rabus et al. 2010), as vegetation and poor thermal conductivity cause evapotranspiration from the surface to be greater than at the subsurface, leading to profile gradients and decoupling of surface soil moisture from that of lower depths (Calvet and Noilhan, 2000).

B.3.2 Characterization of Surface Roughness

When a radar wave impinges on a smooth surface at a specific incident angle, some energy is scattered in all directions, some of which returns to the sensor as backscatter, and the rest is reflected in the specular direction. As the roughness of the surface increases, the specular reflection decreases and the surface-scattering increases, which increases the amount of backscatter returned to the sensor as conceptualized in Figure B-1 (Ulaby et al. 1982). The relationship between SAR backscatter and surface roughness has an exponential dependence (Baghdadi et al, 2008; Zribi and Deschambre, 2002), which tends to be stronger at higher angles of incidence when the soil surface is relatively smooth (Altese et al. 1996; Baghdadi et al. 2002c;

Fung, 1994; Zribi and Deschambre, 2002).This effect is so strong that Bourgeau-Chavez et al. (2007) suggest the variability in surface roughness may cause a more dynamic range of backscatter values than the variability of moisture in the soil, which has led to its accurate characterization being a primary area of interest in soil moisture retrieval (See Verhoest et al. 2008 for a topical review).

Most models are known to perform well when tested against measured backscatter values from a surface with known roughness properties in laboratory settings (Davidson et al. 2001; Fung, 1994). It is this good agreement that gives credence to the use of backscatter models for soil moisture retrieval. Despite this good agreement with theoretical surfaces, application and inversion of surface scattering models under natural conditions has proven to be a difficult task (Altese et al. 1996; Baghdadi et al. 2002b). The underlying difficulty in inverting soil moisture from a natural surface using backscatter models is the inability to accurately characterize the roughness of natural surfaces which results in a dispersion between the modeled and observed backscattering coefficient (Álvarez-Mozos et al. 2006; Davidson et al. 2000; Mattia et al. 2003a; Merzouki et al. 2011; Satalino et al. 2002; Singh and Dubey, 2007; Wagner et al. 2007). Even with extensive *in situ* sampling of surface roughness it remains difficult to characterize surface roughness at the field scale or larger due to the natural variability of the soil (Srivastava et al. 2008). It is the inherent heterogeneity in field scale surface roughness that makes the accuracy of its measurement the limiting factor in the accuracy of satellite based soil moisture retrieval (Bryant et al. 2007).

Surface roughness is typically assumed to be a single scale stationary process (Altese et al. 1996; Álvarez-Mozos et al. 2006; Baghdadi et al. 2004), where roughness is measured once within a field campaign and is often considered to not have changed for subsequent SAR acquisitions, which does not properly account for the variability in natural soils (Wagner et al. 2007). In an attempt to analyze the impact of the spatial and temporal dynamics of surface roughness, particularly with respect to precipitation smoothing, in a real world setting Álvarez-Mozos et al. (2009) found that when roughness was assumed constant, actual changes in surface roughness resulted in a more specular behaving surface. The smoothing noted in agricultural fields largely results from splash erosion caused by rainfall (Jackson et al. 1997); where surface roughness decreases as the cumulative rainfall increases (Callens et al. 2006). The resulting change in backscatter is between 1-2 dB and 2-6 dB, in C and L bands respectively over five

months (Álvarez-Mozos et al. 2009). Thus, the assumption of surface roughness being a stationary process may be invalid under conditions where there is a high potential for erosion, such as recently tilled fields.

For natural surface conditions, as spatial scale increases multi-scale processes are introduced to surface roughness amalgamating the effects of topography, tillage and local single scale surface roughness (Davidson, 2003; Mattia et al. 2003a; Mattia and LeToan 1999). These are represented by the fractal dimension, which is a measure of the degree of self-similarity between small and large scale surface roughness (Mandelbrot, 1983). At small scales single scale processes occur, but as the sampling length of the profile increases multi-scale processes become dominant (Mattia et al. 2003a). The consideration of the surface as a fractal surface requires the use of another variable *D* to describe surface roughness, representing the fractal dimension.

The increased complexity with respect to modelling the surface roughness of the soil is furthered by the effects of volume scattering. It is well known, that radar may increasingly penetrate the surface as wavelength increases and the dielectric constant in the soil decreases (Fung, 1994; Ulaby, 1974; Ulaby et al. 1986). This effect makes it important to consider radar perceived surface roughness which may include roughness due to volume scattering of the penetrating radar waves (Rahman et al. 2008) and dielectric roughness caused by heterogeneity of water particles (Ulaby et al. 1986). In an attempt to grapple with the complexities of surface roughness and maintain model simplicity, Lievens et al. (2011) recommend the consideration of surface roughness parameters not as physically representative processes but as tuning parameters.

B.3.2.1 Surface Roughness Statistics

Because of its strong influence on radar scattering, the proper parameterization of surface roughness is important for modelling a radar signal. Rather than physically representing the surface, these parameters describe the statistical properties of a randomly rough surface. The characterization of surface roughness will be presented prior to discussion of popular models, despite reference to these models, as it provides helpful background. Both the Oh model (Oh et al. 1992) and the Semi-empirical model (Dubois et al., 1995a; b) use root mean square height as a model parameter, whereas the Integral Equation Model (IEM) (Fung et al. 1992) additionally requires the surface correlation length and autocorrelation function (ACF). The correlation length

and autocorrelation function are the parameters which are the most variable and consequently the most difficult to apply in a backscatter model.

Root mean square height is a measure of the vertical distribution of surface height and the correlation length is a measure of the horizontal distance where the surface profile is autocorrelated at a lag at least equal to 1/*e* (Ulaby et al. 1982).

$$\rho(l) \geq \frac{1}{e} \tag{B-1}$$

The autocorrelation function or correlation function is the third descriptor of surface roughness and is generally characterized as either Gaussian, exponential or 1.5 Power as given in equations B-2 to B-4. Selection of an appropriate ACF is important, as the backscattering coefficient changes as a result of the shape of the ACF which can be seen in Figure B-3 (Altese et al. 1996; Fung, 1994).

$$\rho(x) = e^{-x^2/l^2} \quad \text{Gaussian} \tag{B-2}$$

$$\rho(x) = e^{-x/l} \quad \text{Exponential} \tag{B-3}$$

$$\rho(x) = \left(1 + x^2/l^2\right)^{-1.5} \quad \text{1.5 Power} \tag{B-4}$$

Generally, the exponential ACF is applicable to smooth surfaces and is the most widely used, whereas the Gaussian ACF is recommended for rough soils (Fung, 1994). However, the shape of the ACF is dependent on the length of sampling. For very rough soils, the Gaussian ACF matches the roughness when sampling intervals are short, whereas the exponential ACF better represents the roughness at long sampling intervals (Davidson et al. 2000). As profile lengths increase, neither Gaussian nor exponential ACF are appropriate due to multi-scale features (Álvarez-Mozos et al. 2009; Davidson et al. 2000). A Brownian random fractional ACF was introduced by Zribi et al. (1998) to incorporate multi-scale processes into the IEM (Mattia and

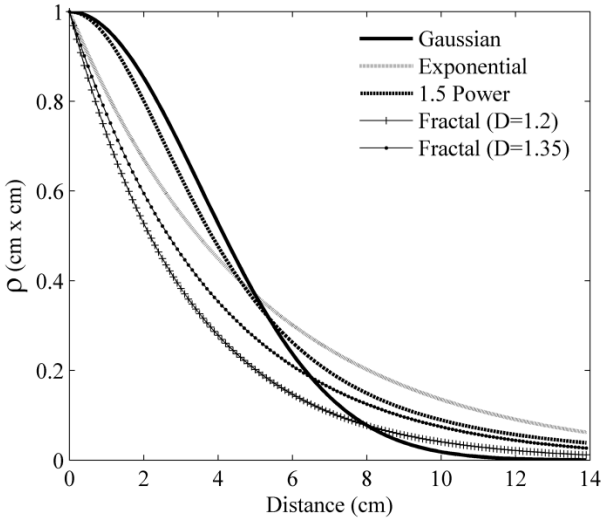

Figure B-3: The shape of three conventional and two fractal ACFs with respect to the ground correlation length.

LeToan, 1999; Zribi et al., 2000). The empirical form of the fractal ACF for all horizontal positions $x > 0$ is given by:

$$\rho(x) = e^{-(x/l)^r} \qquad \text{Fractal} \qquad \text{(B-5)}$$

where, and D is the fractal dimension for all two dimensional profiles x (Zribi et al. 2000). The fractal ACF has an interesting property, in that it can also approximate the Gaussian and exponential ACFs (Zribi et al. 2000). Since surface roughness at long intervals is better represented by a fractal ACF (Davidson et al. 2000), the incorporation of the fractal dimension in the IEM greatly improves its agreement with observed backscatter values (Baghdadi et al. 2004; Mattia et al. 2000; Zribi et al. 2000). Despite this, some disagreement remains as to whether the added complexity of a fractal dimension provides significant improvement with respect to soil moisture retrieval (Satalino et al. 2002).

B.3.2.2 In situ Surface Roughness Measurements

At the field scale it is possible to parameterize surface roughness using *in situ* sampling devices such as a meshboard profiler, pin profiler, laser profiler or through 3D photogrammetry. The two former techniques are criticized for causing surface disruption (Mattia et al. 2003a; Verhoest et al. 2008) and may not be accurate enough for soil moisture retrieval using the IEM (Su et al. 1997). While lasers and photogrammetry do not cause this disturbance, they are still greatly limited by sampling schemes which may not properly represent the surface as viewed by the SAR as well as suffering from light interference and reflection from vegetation. Photogrammetry allows for the retrieval of 3D roughness profiles as opposed to 2D profiles, but can also be limited in terms of sample size (Marzahn and Ludwig, 2009).

Since errors in the retrieval of soil moisture are often attributed to errors in measuring surface roughness (e.g. Álvarez-Mozos et al. 2006; Bryant et al. 2007; Davidson et al. 2000; Mattia et al. 2003a; Merzouki et al. 2011; Satalino et al. 2002; Singh and Dubey, 2007; Wagner et al. 2007), it is important that measurements of roughness be both as accurate and representative as possible. To that end several studies have been conducted in order to suggest requirements for sampling of roughness so that representative values can be found. Oh and Kay (1998) suggest a minimum sample length in order to estimate correlation length and a minimum length for RMS height, whereas others have proposed the ensemble mean of several profiles may give adequate representation without the burden of excessive sampling (Bryant et al. 2007;

Callens et al. 2006; Mattia et al. 2003a). In order to remove multi-scale effects, detrending of long profiles is often employed (Byrant et al. 2007; Callens et al. 2006; Lievens et al. 2009).

While errors in measuring RMS height propagate into larger retrieval errors than correlation length (Lievens et al. 2009), RMS height can be estimated from *in situ* sampling with a relatively high degree of accuracy. Estimation of correlation length is considerably more difficult owing to the parameter's inherent variability (Baghdadi et al. 2002b; Baghdadi et al. 2002c; Baghdadi et al. 2006; Davidson et al. 2003). To alleviate some uncertainty coinciding with the l parameter it can be corrected using 'effective roughness' parameters (Su et al. 1997). The effective roughness proposes to represent l by empirically fitting models using known values of RMS height (Baghdadi et al. 2006; Davidson et al. 2003) and without any *in situ* measurements by using the VH-VV depolarization ratio (Srivastava et al. 2008). Baghdadi et al (2002b; c) proposed a semi-empirical calibration technique of the IEM correlation length, referred to as l_{opt} or optimal correlation length, which incorporated the true correlation length and imperfections in the IEM. It was found that this calibration technique could extend the validity of the IEM from $k\sigma$=3 to $k\sigma\approx$5.3 (Baghdadi et al. 2004) and produces soil moisture inversion results which have lower and more consistent error values than non-calibrated retrievals (Álvarez-Mozos et al. 2008; Baghdadi et al. 2004; Baghdadi et al. 2011).

The approach of Baghdadi et al. (2004) requires extensive *in situ* measurements which is often impractical, but can be overcome by deriving surface roughness from SAR imagery using the Z-index or polarimetric decomposition (Iodice et al. 2011). Zribi and Deschambre (2002) noted that RMS height was poorly correlated with backscatter and proposed a new parameter *Zs*:

$$Zs = \frac{\sigma^2}{l} \tag{B-6}$$

where *Zs* is the dot product of the RMS height and the slope. They found that the difference in backscatter $\Delta\sigma^0$ between two acquisitions of different angles had a correlation coefficient of 0.995 with *Zs*. The strong correlation allows the Z-index to be used to derive soil roughness only from satellite SAR images (Rahman et al. 2008).

When the soil surface is dry the effect of soil moisture can be considered negligible; thus under periods of prolonged dry conditions (drought or dry season) the distributed soil moisture will be relatively homogenous at the permanent wilting point. From this simplification the correlation length of the surface can be determined through a relationship between RMS height

and *l* derived from the IEM for the moisture conditions, radar configuration and the sampled RMS height (Rahman et al. 2007). Rahman et al. (2008) were able to fit a polynomial function based on IEM simulations to approximate *Zs* as a function of the change in backscatter between two incidence angles ($\Delta\sigma^0$) and to approximate σ^0_{dry}.

B.3.3 Backscattering Models Over Bare Soils

B.3.3.1 Semi-Empirical Models

Many authors have presented empirical relationships to relate the radar backscattering coefficient to soil moisture (Le Hégarat-Mascle et al. 2002; Quesney et al. 2000; Oh et al. 1992; Wang et al. 1997; Zribi et al. 2005; Zribi and Deschambre, 2002). While the derived linear relationships can adequately determine soil moisture from radar backscatter, these empirical models are not generally transferrable (Dubois et al. 1995a; b; Le Hégarat-Mascle et al. 2002; Moran et al. 2000; Oh et al. 1992). Semi-empirical models combine the simplicity of widely sourced empirical models with a theoretical foundation providing a compromise between theoretical and physical representation of processes and simplicity. They are relatively widely applicable, as they are derived from multiple ground scatterometer measurements representing various radar configurations and surface conditions. The most commonly used semi-empirical models are the ones derived by Oh et al. (1992) and Dubois et al. (1995a; b).

The Oh model was developed following an investigation of the nature of radar backscatter using polarization ratios. It was found that both the cross ($q \equiv \sigma^0_{hv} / \sigma^0_{vv}$) and co-polarized ratios ($p \equiv \sigma^0_{hh} / \sigma^0_{vv}$) followed a function whose slope initially was increasingly steep and then formed an asymptote as a function of normalized surface roughness ($k\sigma$). The Oh model was updated to account for the effects of surface autocorrelation (Oh et al. 1994) and incidence angle (Oh et al, 1994; Oh et al. 2002). It was again modified to remove the effects of correlation length as well as inherently accounting for the conversion to volumetric soil moisture *Mv* (Oh, 2004).Correlation length was removed because the cross-polarized ratio was insensitive to the surface slope σ/l and because of the difficulties in measuring field correlation length. The resulting model has a validity range of $0.04 < Mv < 0.291$ m^3m^{-3}, $0.13 < k\sigma < 6.98$ at $10° \leq \theta \geq$

70°. The simplicity of the Oh model has the advantage over theoretical models of allowing for a relatively simple numerical inversion outlined by Oh (2004).

The Oh model has been applied over a variety of soil and moisture conditions with varying results. For high incidence angles and rough surfaces, the Oh model accurately estimated backscatter for the co-polarization ratio (Baghdadi and Zribi, 2006) and for all polarizations at C band using RADARSAT data (Álvarez-Mozos et al. 2007). In contrast, some studies have found that the Oh model tends to overestimate radar backscatter for the cross-polarization (Baghdadi and Zribi, 2006) or all polarizations (Merzouki et al. 2010; van Oevelen and Hoekman, 1999), whereas many others have found it tends to underestimate observed backscatter (Boisvert et al. 1997; Gherboudj et al. 2009; Gherboudj et al. 2011; Merzouki et al. 2011; Sahebi and Angles, 2010). It was noted by several authors that the bias in the Oh model for a specific study area could be accounted for by applying a correction factor (Baghdadi and Zribi, 2006; De Roo et al. 2001; D'Urso and Minacapilli, 2006; Merzouki et al. 2010), although more work is required to develop correction factors to operationally implement the Oh model for soil moisture retrieval (Merzouki et al. 2010).

The Oh model has also been calibrated to not require *a priori* knowledge of surface roughness (D'Urso and Minacapilli, 2006) and has been integrated into scattering models to account for the combined effects of vegetation and soil moisture on backscatter (De Roo et al. 2001; Gherboudj et al. 2009; Gherboudj et al. 2011).

Despite the models use for a variety of applications with airborne scatterometer data, until recently, the applicability of the Oh model has been limited by the availability of sufficient spaceborne radar technology. While ground based and airborne SAR systems have long been able to acquire multi-polarized data, only recently have satellite based SAR sensors (i.e. RADARSAT-2, ASAR, TerraSAR-X, PALSAR) been able to simultaneously acquire the multi-polarized data necessary to derive the polarization ratios.

The Dubois model was designed to be capable of modelling a variety of surface conditions and radar configurations because it was derived from a wide variety of scatterometer data (Dubois et al. 1995a, b). The model expresses the co-polarized backscattering coefficients as a linear function of the angle of incidence, RMS height, dielectric constant and wavelength and has a validity region of $k\sigma \leq 2.5$, $Mv \leq 35\%$ and angle of incidence is $\geq 30°$ (Dubois et al. 1995a). While the model was designed to be applicable to lightly vegetated surfaces, the presence of

vegetation will result in an overestimation of surface roughness and an underestimation of soil moisture. This tendency of the model resulted in the recommendation that a further constraint be placed on the validity where it is only applicable to surfaces where the normalized difference vegetation index (NDVI) < 0.4 or $\sigma^0_{hv} / \sigma^0_{vv}$ < -11dB at L band (Dubois et al. 1995a). The model can be analytically inverted to derive RMS height and the dielectric constant from the HH/VV ratio without *a priori* knowledge of the surface roughness.

Using radar backscattering coefficients in C band, Baghdadi and Zribi (2006) found the Dubois model had a bias of 1.7dB in HH polarization and no appreciable bias in VV polarization. The model tended to underestimate backscatter for smooth surfaces and overestimate backscatter for very rough surfaces, whereas it produced reasonable backscatter coefficients for surfaces with intermediate roughness and soil moisture. A similar comparison by Álvarez-Mozos et al. (2007) produced comparable results, where the Dubois model often underestimated the backscattering coefficient particularly in moist conditions. It was also found that the Dubois model contradicted a known relationship between backscatter and soil moisture, where backscatter is known to be insensitive to high soil moisture content (Álvarez-Mozos et al. 2007). In contrast, Wang et al. (1997) found the Dubois model to be in relatively good agreement with observations, but noted a large number of pixels failed to reach a soil moisture retrieval solution because of model errors and its validity range. This resulted in the removal of approximately 25-50% of the measured pixels, especially under dry conditions (Wang et al. 1997). In an attempt to improve the accuracy of the Dubois model, Sahebi and Angles (2010) calculated new coefficients for the model following the same procedure as in Dubois et al. (1995a, b), resulting in a site dependent model which was more accurate than competitive models.

B.3.3.2 Theoretical Backscattering Models

Theoretical backscattering models simulate the backscatter derived from a particular radar configuration and geometry based on the apparent surface conditions at the wavelength and resolution of the radar. While being more mathematically complicated, theoretical models can be used in a wide variety of conditions while making few *a priori* assumptions about the characteristics of the surface as compared to empirically derived models. Theoretical models are

derived from either an analytical or numerical solution to Maxwell's equations for the interaction of electric and magnetic fields in a conducting medium.

B.3.3.3 Numerical Backscattering Models

Numerical models solve Maxwell's equations through discretization of a conducting surface as a combination of basis functions in a mesh framework. The mesh is superimposed on a representation of the surface to be modeled and the numerical algorithms satisfy Maxwell's equations at each point on the mesh, which is then amalgamated to produce the backscattering value. While physically representative, numerical methods require costly computational resources, which, until recently were only available through supercomputing. Thus, while offering great potential for soil moisture retrieval, the advancement of numerical methods has been inhibited by the need for faster algorithms and more computational resources (Huang et al. 2010; Lawrence et al. 2011, Onier et al. 2011; Rabus et al. 2010).

The Method of Moments (MoM) is the most commonly applied numerical model as it requires fewer computational resources than other models because the electric and magnetic equations are solved across a surface mesh (Warnick and Chew, 2001). The use of a single dense surface mesh allows MoM to be efficient but also limits its capability to properly represent volume scattering and heterogeneous media (Huang et al. 2010; Lawrence et al. 2011). Historically, the MoM has been used as a tool to independently test the theoretical validity of analytical backscattering models (Huang et al. 2010; Fung, 1992; Fung 1994; Ulaby et al. 1986). The restriction of the MoM to this task was largely due to the fact that it could only be used to model small homogeneous surfaces. However, recent computational and algorithmic advances have allowed for the development of a 3D MoM algorithm proposed for a lookup table inversion scheme for missions such as SMAP (Huang et al. 2010).

The Finite Element Method (FEM) (Zienkiewicz and Taylor, 2000) and Finite Difference Time Domain method (FDTD) (Yee, 1966) discretize the media to be modeled using volume meshing or cells (Lawrence et al. 2011; Onier et al. 2011; Rabus et al. 2010). This allows for FEM and FDTD methods to be implemented across heterogeneous surface as well as having potential to model multi-layered media such as forests (Lawrence et al. 2011). Using an FDTD algorithm, Rabus et al. (2010) were able to assess the impacts of soil moisture gradients on backscatter, while Onier et al. (2011) and Lawrence et al. (2011) have used FEM to model the

impact of volume scattering related to clods and litter cover with various soil moisture conditions.

Numerical models offer a comprehensive and physically representative method of modelling backscatter and have high potential as forward models in an inversion scheme. However, despite recent advances in computing technology and algorithms, numerical models are still constrained to simplified examples often being limited to two dimensions (Lawrence et al. 2011; Onier et al. 2011; Rabus et al. 2011) or to an area no more than 8 to 10 times the wavelength (Huang et al. 2010). While a promising option, the development of faster algorithms, better computational facilities and further research is required before numerical models can be regularly used for soil moisture retrieval.

B.3.3.4 Analytical Backscattering Models

While several analytical backscattering models exist, the focus of this review will be the discussion and evolution of the Integral Equation Model (IEM) which was introduced in 1992 and has been subsequently revised to improve the representativeness and validity range of the model (Fung, 1992, Fung, 1994, Mattia and LeToan, 1999, Hsieh et al. 1997, Chen et al. 2000, Wu et al. 2001, Chen et al. 2003, Liu et al. 2003, Fung and Chen, 2004). Focus is placed on the IEM because other analytical models such as the Geometric Optics Model (GOM) (Ulaby et al. 1982) and Small Perturbation Method (SPM) (Ulaby et al. 1982) are approximated by the IEM at high and low frequencies respectively, leading to the conclusion that these are special cases of the IEM (Fung, 1992; Fung, 1994), and because it is by far the most widely used theoretical model found in the literature over the last decades.

When applied to the problem of soil moisture retrieval, it considers how an incident wave will be scattered from a bare soil surface as a result of SAR configuration and surface properties. The IEM can simulate co-polarized as well as cross-polarized radar for a given incidence angle and frequency. The surface parameters represented by the IEM are the RMS height, correlation length the ACF and the dielectric constant of the soil (ε^r). Surface parameters are often treated as unknowns or are determined by field sampling of sensitive targets. The problem of inversion of soil moisture from the IEM is compounded by the difficulty that has been found in applying the model to natural surface conditions (Altese et al. 1996; Baghdadi et al. 2002b; Baghdadi et al. 2002c; Baghdadi and Zribi, 2006; Zribi and Deschambre, 2002), which is not considered to be a

deficiency in the model itself but rather results from the poor characterization of the soil surface (Altese et al. 1996; Baghdadi et al. 2002a; Baghdadi and Zribi, 2006).

Most variants of the IEM commonly found in the literature with respect to soil moisture retrieval are the result of trying to better describe the surface roughness or to calibrate the IEM (Baghdadi et al. 2006; Mattia and LeToan, 1999; Zribi et al., 2000). Besides better representation of the surface roughness, several studies have focused on better representation of physical processes in the IEM by removing assumptions or including formulations which were not available during the original publication. Subsequent model updates include a revised Green's function, the inclusion of multiple scattering and the incorporation of a transition function for the Fresnel coefficients (Chen et al. 2000; Hsieh et al. 1997; Wu et al. 2001). These advances allow the IEM to better account for the spatial correlation of the surface (Hsieh et al. 1997), upwards and downwards multiple scattering (Chen et al. 2000; Fresnel reflection in Figure B-1) and the transition between the use of the incident and specular angle as surface roughness increases (Wu et al. 2001). The resulting model is termed the Advanced Integral Equation Model (AIEM) (Fung and Chen, 2004; Liu et al. 2003; Wu and Chen, 2004). A summary of the diverse versions of the IEM with various complexities and a focus on model users can be found in Fung and Chen (2010) and a summary of the above advances for the IEM and other models can be found in Table 2-3.

Comparison of the advantages and shortcomings of each version of the IEM shows that, at VV polarization, for soils that are relatively smooth when normalized by wavelength, the original IEM offers accurate representation of backscatter and is computationally efficient. The AIEM is more appropriate at high incidence angles with HH polarization, especially as roughness increases, because multiple scattering becomes important. While more mathematically complicated this model offers better representation of radar backscatter due to the effect of multiple scattering and the movement into the transition region between the incident and specular region of the Fresnel reflection coefficient.

As has been previously mentioned, inversion of soil moisture from the IEM is difficult because measured SAR backscatter in a natural environment does not necessarily behave in the same manner as the model predicts. In studying soil moisture in bare or near bare agricultural fields, the IEM has often been found to consistently overestimate the measured radar backscatter in some instances (Baghdadi et al, 2004; Baghdadi and Zribi, 2006; Mattia et al, 2003b;

Table B-3: Summary of backscatter model advances.

Integral Equation Model (IEM)		
Reference		Contribution
Fung, 1992	-	Initial derivation of IEM theoretical backscattering model
Fung, 1994	-	Separated IEM into simple and complex forms for single and multiple scattering
Hsieh *et al*. 1997	-	Removed simplification of Green's function
Mattia and LeToan, 1999 & Zribi, 2000	-	Introduced fractal ACF
Chen *et al*. 2000	-	Derived upward and downward scattering complimentary field coefficients
	-	Improved multiple scattering
Wu *et al*. 2001	-	Derived transition function between Fresnel and specular angles
Liu *et al*. 2003	-	Incorporated transition function and updated Green's function to derive AIEM
Fung and Chen, 2004	-	Incorporated transition function into original IEM
Wu and Chen, 2004	-	Incorporated transition function and updated Green's function to derive AIEM

Oh Model (OM)		
Reference		Contribution
Oh *et al*. 1992	-	Initial development of semi-empirical Oh model
Oh *et al*. 1994	-	Incidence angle function added
	-	Surface autocorrelation function added
Oh *et al*. 2002	-	Incidence angle function updated
Oh, 2004	-	Correlation length removed due to insensitivity of model to surface slope

Dubois Model (SEM)		
Reference		Contribution
Dubois *et al*. 1995a; b	-	Initial development of Dubois semi-empirical model
Sahebi and Angles, 2010	-	Updated coefficients to increase site specific model accuracy

Merzouki et al., 2010), while underestimating it in others (Kim and van Zyl, 2009). Often such errors are suggested to be the result of inappropriate characterization of the soil roughness (Altese et al. 1996; Baghdadi et al. 2002a; Merzouki et al. 2011). However, little attention has been paid in the literature to the role of volume scattering in radar return, at least in terms of its impact on the IEM. Volume scattering partially accounts for the reason why the IEM is expected to be more valid at higher dielectric values as penetration decreases (Fung, 1992). In contrast, for a dry soil surface, the dielectric constant is low and radar penetration is relatively high, in which

case the IEM may overestimate surface backscatter (Boisvert et al. 1997). Since the IEM and its derivatives do not inherently account for volume scattering, more study and potential model updates are required before definitive conclusions can be made about the applicability of the IEM when volume scattering may impact soil moisture retrieval.

B.3.4 Modelling Vegetation for Soil Moisture Retrieval

B.3.4.1 Impact of Vegetation on Radar Backscatter

A fundamental shortcoming of the IEM, the Oh and Dubois models in terms of soil moisture retrieval is that they are valid only for bare or sparsely vegetated soil surfaces. Many real soil surfaces are covered to some extent by vegetation, and almost all soil surfaces studied from the perspective of soil moisture retrieval are agricultural fields. Vegetation affects SAR backscatter in a fashion similar to soil surfaces. The scattering or attenuation of radar signals will vary based on the dielectric properties of the vegetation (i.e. water content of the leaves and trunk) as well as by the physical structure of vegetation (Ulaby et al. 1982). The vegetation canopy typically results in volume scattering, with the amount of penetration (optical depth τ) being determined by the radar frequency and the vegetation water content (VWC). Stalks, branches and trunks contribute to direct backscattering, trunk-ground double scattering and volume scattering, as seen in Figure B-1, as well as depolarization depending on their orientation relative to the incoming radar signal (Ulaby et al. 1982). When biomass is very low, vegetation is often neglected (Dubois et al. 1995a; Glenn and Carr, 2004; Lakhankar et al. 2006; McNairn and Brisco, 2004, Moran et al. 2000; and others), however as biomass increases the backscatter will become decreasingly sensitive to soil properties (Romshoo et al. 2002; Wigneron et al. 2004), until under forested environments the contribution of soil moisture to backscatter is indistinguishable. Fortunately, when biomass is high in agricultural fields, the backscattered signal is the sum of the vegetation and soil contributions, and has soil moisture information incorporated into it making retrieval of soil moisture possible (Joseph et al, 2010; Mattia et al. 2003b).

In order to separate the influence of vegetation and soil on radar signal, it is necessary to determine vegetation parameters for inclusion into models. Unfortunately, these parameters are difficult to set because they change with angle, wavelength and phenology (Wigneron et al, 2004). Currently, most research in soil moisture retrieval from vegetated surfaces occurs in an

agricultural context, leaving much work to be done to determine vegetation influences from natural vegetation and forests. As a result of this, the focus of the following sections is primarily derived from parameterization of crops.

In early stages, crops develop slowly, moderating their influence on SAR signal (Mattia et al. 2003b). Backscatter increases as the crop grows, until the radar saturation point is reached, after which very little signal from the soil surface will be distinguishable depending on crop type and water content. During this phase, vegetation scattering will increase in an approximately linear fashion in relation to VWC (Dobson and Ulaby, 1986) and plant density (Patel et al. 2006) where a wet canopy may increase the C band scattering by as much as 3 dB (Brown et al. 2003). For broad leaf crops, backscatter increases with VWC which eventually results in signal saturation at both C and L bands, whereas small leaf and grain crops have a dynamic backscatter relationship throughout their growth cycle allowing for differentiation of a soil signal (Paloscia et al. 2002; Soria-Ruiz et al. 2009). In contrast to grain crops, tree canopies tend to backscatter the majority of the radar signal from the crown (Dobson et al. 1992). The amount and impact of canopy penetration is difficult to determine in retrieval studies as models tend to underestimate the depth of penetration (Brown et al. 2003; Martinez et al. 2000).

B.3.4.2 Crop Parameterization

As discussed in the previous sub-section the backscattering response of vegetation changes with water content, phenology, biomass, and other factors. This makes it necessary to separately delineate between dominant land cover types prior to training a model inversion. Accurate classification of surface types is usually derived from optical remote sensing data (i.e. SPOT or Landsat satellites), although combinations of SAR data have also been successfully used (Kellndorfer et al, 1998).

The two most used vegetation parameters are VWC and LAI. VWC is related to the depth of the radar penetration, where the optical depth will decrease linearly with increasing VWC, therefore impacting the quality of soil moisture return. Thus, in order to derive VWC for an area, it must be empirically regressed requiring possibly extensive *in situ* data collection, often using the Normalized Difference Vegetation Index (NDVI) (Álvarez-Mozos et al. 2006; Jackson et al. 2004) or Normalized Difference Water Index (NDWI) (Jackson et al. 2004; Notarnicola et al. 2006). LAI is a measure of the leaf surface area per unit of ground surface, and has been used as

an indicator of vegetation health and density for soil moisture retrieval by some authors (Moran et al. 2000; Said et al. 2008). Similar to the derivation of VWC, the LAI is typically regressed based on a series of ground truth measurements. Using neural network inversion Said et al. (2008) found that LAI alone was comparable to using LAI in combination with VWC and crop height and better than the other vegetation indicators for soil moisture retrieval. Traditionally, LAI is empirically related to NDVI values (Manninen et al. 2005; Moran et al. 2000) requiring the addition of valid optical data at the time of SAR acquisition. However, LAI can also be empirically related to the HH/VV ratio as LAI has been found to have a stronger relationship to the HH/VV ratio than SPOT derived NDVI (Manninen et al. 2005). Other important crop parameters include crop height, total biomass, wet biomass and vegetation structure and geometry. Crop height and biomass can be derived from the depolarization ratio (Gherboudj et al. 2011), and used in empirical and semi-empirical models in a fashion similar to VWC and LAI. Vegetation structure and geometry parameters are used to create physically representative radiative transfer (RT) models of vegetation, where stalks, branches and leaves are modeled as a collection of dielectric discs and rods, with a representative geometry and density (Bracaglia et al. 1995; Moghaddam et al. 2000). These types of models are used when ground-trunk and multi-bounce scattering are important (Balenzano et al. 2011; Della Vecchia et al. 2008; Quesney et al. 2000).

B.3.4.3 Modelling of Vegetation Backscattering

Coupling a vegetation model with a soil surface model allows for the differentiation of the soil signal from the vegetation signal. Models are of variable complexity and accuracy and application is based on the availability of *in situ* data requirements for complex models.

The coupling of a RT model with a soil scattering model allows for the retrieval of soil moisture under vegetated canopies and has been widely applied to a variety of crop and vegetation types (Balenzano et al. 2011; Della Vecchia et al. 2008; De Roo et al. 2001; Dobson et al. 1992; Le Hégarat-Mascle et al. 2002; Lin et al. 2009; Quesney et al. 2000; Shi et al. 2004). There are several RT models such as Michigan Microwave Canopy Scattering Model (MIMICS) (Ulaby et al. 1990) and that of Bracaglia et al. (1995) which can model a variety of species within a vegetation type. While more complex crop models exist (i.e. Stiles and Sarabandi, 2000; Stiles et al. 2000), which model a single crop type in a highly physically representative manner,

these models are not often used for soil moisture retrieval, likely due to the added complexity. The advantage of RT models for soil moisture retrieval is their accuracy in modelling backscattered signal for a specific vegetation type. Conversely, the disadvantage is that they are often difficult to generalize, requiring the parameterization of many RT models for a given scene, requiring sometimes extensive parameterization of the vegetation (Dobson et al. 1992; Lin et al. 2009). Simplified RT models require less parameterization than MIMICS, but still require many vegetation parameters representing crop height, leaf and stalk shape, angle and dielectric properties (Belanzano et al. 2011; Della Vecchia et al. 2008; Quesney et al. 2000).

A significantly less complicated model which is frequently coupled with surface models is the water-cloud model (WCM) (Attema and Ulaby, 1978). A WCM represents the canopy, in a RT framework, as a uniform cloud of spherical dielectric droplets, all of which are held in place by dry solid matter which does not affect radar signal. The relative simplicity of the WCM makes it a good candidate for soil moisture retrieval (Bindlish and Barros, 2001). Conceptually the WCM makes the assumptions that multiple scattering between the canopy and the soil can be neglected and the only significant variables are the density of the cloud and the height of the cloud. The general WCM is represented by:

$$\sigma^0 = AV_1 \cos\theta\left[1-\exp\left(\frac{-2BV_2}{\cos\theta}\right)\right]+\exp\left(\frac{-2BV_2}{\cos\theta}\right)\left(C+Dm_s\right) \tag{B-7}$$

where the parameters V_1 and V_2 are canopy descriptors, and A-D are fitted parameters. The parameters A and B are typically vegetation parameters which determine the contribution from the vegetation canopy and C and D are considered soil parameters. In the general WCM formulation of Equation B-7 the first term represents the vegetation contribution to backscatter and the second term the soil contribution, although other forms of the WCM have been used (see Graham and Harris, 2003 for a review). The WCM has been widely used because its simplicity allows for its application in a variety of vegetation and climate regimes (Álvarez-Mozos et al. 2006; Bindlish and Barros, 2000; Bindlish and Barros, 2001; Crow et al. 2010; Gherboudj et al. 2009; Gherboudj et al. 2011; Lievens and Verhoest, 2011; Wang et al. 2011; Zribi et al. 2007), and it can be coupled with a surface model and multi-parameter radar data in order to retrieve soil moisture solely from remote sensing data (Wang et al. 2011). However, since the model assumes only single scattering is present, it may misrepresent backscatter under dense vegetation, which has a strong contribution from vegetation-surface multiple scattering, limiting

its applicability in some instances (Joseph et al. 2008; Pierdicca et al. 2010). Also, similar to complex RT models, the WCM may require separate parameterization for each vegetation class being studied (Álvarez-Mozos et al. 2006; Bindlish and Barros, 2001; Lievens and Verhoest, 2011).

B.3.5 Selection of Radar Parameters for Soil and Vegetation

Selection of the appropriate incidence angle, polarization and wavelength has the potential to either enhance or detract from the capability of retrieving soil moisture from a radar scene. Low incidence angles are preferred because they result in a shorter travel distance, less perceived vegetation biomass, a decrease in shadowing and minimization of multiple scattering, thereby lessening the effect of surface roughness (Aubert et al. 2011; Balenzano et al. 2011; Baghdadi et al. 2002c; Ulaby and Batlivala, 1976). This occurs because, as the incidence angle of the radar increases, the radar beam effectively has more vegetation through which to penetrate, and the high angle exaggerates the surface roughness (Dobson and Ulaby, 1986; Joseph et al. 2010; Paloscia et al. 2002). While increased perceived roughness will result in increased backscatter, the effect of the increased vegetation thickness may either result in greater attenuation of the signal or greater backscatter, depending on the dominating radar vegetation interactions.

In terms of polarization of radar signal, HH polarization is most often recommended as it allows for the highest amount of vegetation penetration by minimizing the impact of vertical stalks and trunks (Balenzano et al. 2011; Della Vecchia et al. 2008; Romshoo et al. 2002; Ulaby and Batlivala, 1976). Recently, the sensitivity of cross-polarization (HV or VH) and the depolarization ratio (HV/VV) to vegetation has resulted in some focus being placed on this configuration for removing the impact of vegetation. The degree of depolarization is determined by structure and density of the vegetation (Patel et al. 2006) and is proportional to the volume scattering (Shi et al. 2004). The high sensitivity to crop height and low sensitivity to VWC suggests the depolarization ratio may be predominantly controlled by dry biomass which can be used to determine crop parameters to be substituted into retrieval methods (Della Vecchia et al. 2008; Gherboudj et al. 2011; Lin et al. 2009).

Longer wavelengths are recommended as they will penetrate vegetation (Blumberg et al. 2000; Ulaby et al. 1986) and the perceived roughness will be less at long wavelengths (Blumberg et al. 2000; Lievens et al. 2011). Despite the lower penetration, some authors claim C band SAR

still shows appreciable sensitivity to soil moisture at high biomass (Romshoo et al. 2002), and its sensitivity to vegetation makes removal of the vegetation effect more reliable (Notarnicola et al. 2006). In L band, the higher penetration means the longer wavelength is total plant sensitive as opposed to canopy sensitive (Patel et al. 2006) however, the double bounce, trunk-ground scattering is dominant at L band (Moghaddam et al. 2000). In cereal crops, with HH polarization, the penetration of L band signals minimize the vegetation sensitivity increasing the strength of the relationship to soil moisture (Della Vecchia et al. 2008; Mattia et al. 2009). Unfortunately, the true value of L band and longer SAR wavelengths for soil moisture monitoring from space is difficult to assess since the majority of operational SAR sensors operate in the shorter C band. While there are a few examples of studies based on the L band PALSAR which display the aptitude of this wavelength for soil moisture retrieval (Kasischke et al. 2011; Lievens and Verhoest, 2011; Lucas et al. 2010), more work is required to better understand the capabilities and limitations of such data from spaceborne platforms.

Therefore, with single configuration radar the conventional wisdom introduced by Ulaby and Batlivala (1976) stated that for soil moisture retrieval the ideal radar parameters are a low frequency, low incidence angle and co-polarized configuration, and is largely maintained in the literature (Anguela et al. 2010; Aubert et al. 2011; Balenzano et al. 2011; Baghdadi et al. 2002a; Baghdadi et al. 2002c; Baup et al. 2007; Jackson et al. 1997).

B.3.6 Soil Moisture Inversion Methods

In the context of SAR soil moisture retrieval, model inversion is the process of reversing a selected backscatter model, where potentially unknown surface parameters can be determined from known backscatter. The problem of inversion can be under-defined where, at the pixel scale, a single backscatter value is decomposed into potentially multiple surface parameters (Loew and Mauser, 2006; Mattia et al. 2006; Shi et al. 1997). To mitigate the under-determined problem, many retrieval studies included a component of *in situ* data, where surface roughness was measured to reduce the number of unknowns to one (Baghdadi et al. 2002a; Jackson et al. 1999; Le Hégarat-Mascle et al. 2002; Quesney et al, 2000; Romshoo et al, 2002; Shi et al. 1997; and others). Besides being ill-posed, soil moisture inversion encounters a problem of equifinality, where different combinations of surface roughness and dielectric constant may result in the same backscattering coefficient (Hoeben et al. 1997). This means the primary

limitation of conventional soil moisture inversion is in regards to the necessity of characterizing the soil surface (Davidson et al. 2003; Mattia et al. 2003b; Singh and Dubey, 2007). It is sometimes suggested that this is largely an issue of scale, where a large homogenous unit should be considered (Glenn and Carr, 2004; Le Hégarat-Mascle et al. 2002; Sahebi and Angles, 2010; Thoma et al. 2006; Thoma et al. 2008). It follows that at large scale, the effects of surface roughness may be operationally ignored because pixels are averaged at a scale beyond which small scale roughness differences affect radar backscattering, although hydrologically this leads to problems with respect to soil moisture heterogeneity.

Because of the complicated nature of many backscattering models analytic inversion is impossible. Therefore, numerical, regression and other relational methods are often employed in order to establish a relationship between modeled and observed radar backscattering coefficient. Often a model is used to establish a theoretical relationship between a set of surface parameters, the soil dielectric properties and backscatter, which is validated against observed backscatter to determine the accuracy of the inversion. Unfortunately, the spatial variability of surface roughness and soil moisture often leads to disparity between simulations and measurements which may reach several decibels, making inversion results inaccurate (Álvarez-Mozos et al. 2006; Baghdadi et al. 2008a). The following section will discuss many different methods of inversion which have been commonly applied for soil moisture retrieval. The majority of the discussed methods are based on application with the IEM, although many are applicable with other scattering models.

B.3.6.1 Change Detection

The difficulty often encountered in using semi-empirical models and the IEM is the parameterization of surface roughness and vegetation (Baghdadi et al. 2006; Dubois et al. 1995a; b; Merzouki et al. 2010), making simple techniques to derive soil moisture where other surface variables can be inherently accounted for attractive. Change detection techniques are employed to remove the effect of surface roughness variables, which are assumed constant in time, while making no *a priori* assumption about the scale and homogenous distribution of surface roughness. The removal of these assumed constant effects in the radar images allows for an empirical relationship to be established with soil moisture (Kim and van Zyl, 2009; Kurucu et al. 2009; Moran et al., 2000; Pathe et al. 2009; Quesney et al. 2000; Zribi et al. 2007). It is often

assumed that vegetation growth does not change between acquisitions or that the time interval between acquisitions is inherently too short for change (Joseph et al. 2008), otherwise vegetation is explicitly accounted for in the model (Wickel et al. 2001).

An initial acquisition is made for a dry date, where the effect of the dielectric constant on radar backscatter can be negated. If the effect of the dielectric constant is eliminated then measured backscatter will be solely a function of surface roughness (Wagner et al. 1999). The backscatter coefficient is then subtracted from subsequent wet images to remove the influence of roughness. This simple technique for acquiring the dry backscatter coefficient is applicable in regions with distinct dry seasons, otherwise, soil moisture observations or model predictions are required to establish the relationship. The general equation is given by:

$$\sigma_{soil} = \alpha M_v + \beta \tag{B-8}$$

where β is a function of soil roughness which is eliminated from the equation by assuming it does not change between radar observations. Therefore,

$$\sigma_{soil} - \sigma_{dry} = \alpha\left(M_v - M_{v0}\right) \tag{B-9}$$

where M_{v0} is the moisture content of the dry image which is assumed to be approximately zero resulting in

$$\sigma_{soil} - \sigma_{dry} \approx \alpha M_v \tag{B-10}$$

The resulting retrieval method is much simpler than the use of models (Zribi et al. 2007) and can often allow for a simple linear relationship between backscatter and soil moisture (Moran et al. 2000). Moran et al. (2000) noted that in semi-arid regions with little change in vegetation cover, the derived empirical relationship produced accurate results when soil moisture was above 20% but when soil was relatively dry the change detection method produced results which had a difference of less than 1 dB.

Other research has applied the change detection technique to non-homogenous areas by incorporating multiple empirical models. For example, Quesney et al. (2000) used unsupervised classification to determine crop classes, allowing for the isolation of wheat fields. Le Hégarat-Mascle et al. (2002) and Zribi et al. (2007) used a similar procedure where Le Hégarat-Mascle et al. extended the procedure to include multiple crop types and accounted for the crop growth through use of a RT model (Karam et al. 1992), whereas Zribi et al. (2007) used a WCM to account for vegetation at a regional scale, resulting in a more parsimonious retrieval model.

Kurucu et al. (2009) attempted to account for vegetation using a change detection methodology by incorporating a near infrared (NIR) optical image through an image fusion technique of RADARSAT-1 and SPOT data, which provided a better soil moisture return than either data type alone. In a similar attempt, Kim and van Zyl (2009) combined radiometer data with SAR to improve the retrieval response. The application of polarization ratios in a time series context can also be an effective way of decoupling the influence of vegetation (Balenzano, et al. 2011). Currently, there are two primary limiting factors of the change detection technique. The main limitation is the availability of data with a short time lag and similar viewing geometry. The other, is with respect to the constant surface roughness condition when applied at the field scale as in Lievens and Verhoest (2012). Change detection techniques may only be valid over short time periods since tillage, erosion and other processes may alter surface roughness, which is especially influential at smaller scales.

B.3.6.2 Optimization of a Cost Function

For the optimization problem, a cost function is established which is a measure of error in the system. Once acceptable error criteria have been established, an iterative procedure is employed to search for the parameters that provide the minimum error. Pauwels et al. (2002) minimized the difference between the IEM and Oh model finding the solution that optimally fit both models. Joseph et al. (2010) employed an optimization algorithm to find the RMS height and correlation length characteristics, which minimized the difference between model predictions generated with observed soil moisture and a series of backscatter measurements. Wang et al. (2011) and van der Velde et al. (2012) used a minimization algorithm to optimize soil moisture from the IEM by using multi-angular data to determine surface roughness. The most common approach found in the literature uses the Bayesian MAP framework to determine the parameters which minimize the cost function using the posterior probability distribution (Pierdicca et al. 2010; Mattia et al. 2009; Mattia et al. 2006).

Considering the known uncertainty in soil moisture retrieval resulting from the inherent variability of surface roughness and soil moisture over the field to watershed scale, optimization techniques represent an opportunity to better parameterize retrieval models in an operational algorithm especially when field sampling at a representative scale is not feasible.

B.3.6.3 Linear and Non-Linear Interpolation

Linear inversion is based on a simple least squares fitting of the desired parameter and model simulations of the backscattering coefficient, and is often applied at larger spatial scales (i.e. Anguela et al. 2010; Baghdadi et al. 2006; Baghdadi et al. 2007; Baup et al. 2007; Bourgeau-Chavez et al. 2007). Direct linear inversion models are site specific and require classification of a radar scene prior to inversion. For example, Bourgeau-Chavez et al. (2007) grouped areas with different burn severity before retrieving surface soil moisture in the Boreal forest, and in agricultural settings many authors have classified fields based on the crop type, tillage practices or soil texture (Álvarez-Mozos et al. 2006; Baghdadi et al. 2007; Balenzano et al. 2011; Quesney et al. 2000; and others). While the accepted relationship between backscatter and soil moisture is non-linear, Quesney et al. (2000) found that if the scale is sufficiently large, the relationship between radar backscattering signal and surface soil moisture could be approximated by a linear function. If the relationship is assumed non-linear, the form of the non-linear model used in inversion varies from exponential (Baghdadi et al. 2006) to polynomial (De Roo et al. 2001; Shi et al. 1997; Su et al. 1997; Altese et al. 1996) depending on the nature of the relationship between backscatter and surface conditions. Both linear and non-linear least squares inversions, are only applicable when one soil parameter (i.e. soil moisture) is considered unknown. Since the surface roughness parameters must be known in order to derive soil moisture, multiple inversion models are required to determine distributed soil moisture, where surface roughness parameters are required for each area under examination.

B.3.6.4 Look Up Table Inversion

Inversion by look up table (LUT) is similar in process to inversion by least squares interpolation and has been successfully applied in a number of studies (i.e. Bryant et al. 2007; Merzouki et al. 2011; Rabus et al. 2010; Rahman et al. 2007; Rahman et al. 2008). A forward model is run for a diverse range of variables for a given radar configuration. The soil parameters which resulted in a specific backscattering coefficient are stored in tabular format to be retrieved when a similar backscattering coefficient occurs. This method of inversion is relatively simple, computationally efficient and can be used to find an inversion solution for more than one parameter (Rahman et al. 2007). Interpolation can also be used when an observed value is between table values (Huang et al. 2010; Rabus et al. 2010).

B.3.6.5 Statistical Probability Inversion

Statistical inversion is used to create a probability distribution function (PDF) from which the statistical likelihood of particular soil moisture value can be determined from the surface properties of the soil and the backscattered signal. A Bayesian approach is employed where a forward model is used to establish a PDF to determine the probability of finding a backscattering coefficient given a set or subset of soil parameters (Notarnicola et al. 2008), which have been assumed to follow a Gaussian distribution (Notarnicola et al. 2008; Zhan et al. 2006; Notarnicola et al. 2006).

The standard Bayesian approach seeks an optimal estimator as the mean value of the PDF (Notarnicola et al. 2006; Pierdicca et al. 2008). This approach is extended to derive a joint Gaussian PDF accounting for noise as well as surface roughness parameters (Notarnicola et al. 2008). Pierdicca et al. (2008, 2010) compared the minimum variance (MV) Bayesian technique to the maximum posterior probability estimator (MAP) and found that if no correlation between RMS height and *l* were assumed that MAP would retrieve soil moisture with a slightly higher accuracy than MV, and would retrieve roughness parameters slightly worse than MV. If RMS height and *l* were assumed correlated MV retrieval produced higher accuracy than MAP for both soil moisture and surface roughness (Pierdicca et al. 2008). For both methods, the quality of the *a priori* information used to establish the PDF is important, as poor data can lead to a wide distribution and poor retrieval results (Pierdicca et al. 2008).

In a time series approach Pierdicca et al. (2010) used the MAP method to incorporate a time series of SAR acquisitions, where the conditional probability of the current surface parameter (i.e. soil moisture) was determined given the current backscattering signal and the parameter's previous state, if the two observations were assumed close enough to be related. In order to invert soil moisture from a model without *a priori* knowledge of surface roughness, Verhoest et al. (2007b) related surface roughness to tillage and agricultural calendars through possibility distributions. A possibility distribution represents a fuzzy subset of a known variable and provides a weaker association than a probability distribution because it accounts for inherent uncertainty in large scale measurements (Verhoest et al. 2007b). Since the possibility distributions were wide, leading to uncertain soil moisture retrieval, a following study proposed the use of a joint possibility distribution to account for interactions between RMS height and *l*

(Vernieuwe et al. 2011). Unfortunately, different agricultural practices (i.e. speed, tiller depth, etc.) lead to different roughness characteristics, making identification of accurate possibility distributions difficult, especially at the core of the distribution, which hindered retrieval accuracy (Vernieuwe et al. 2011).

B.3.6.6 Artificial Neural Networks and Fuzzy Logic

Artificial neural networks (ANN) link a defined set of training inputs through a weight matrix with a hidden layer composed of artificial neurons. A feedforward ANN is considered to be a universal approximator of any non-linear function and has been used with success for inverting soil moisture models (i.e. Baghdadi et al. 2002a; Fung, 1994; Lakhankar et al. 2006; Notarnicola et al. 2008; Pierdicca et al. 2008; Said et al. 2008; Satalino et al. 2002). Once the architecture of the network is determined, a dataset is created using a surface scattering model in forward mode to train the ANN, which can subsequently simulate a response for a given input value. In comparison studies, ANNs have been found to be able to retrieve soil moisture with greater accuracy than conventional or multiple regression (Said et al. 2008), fuzzy logic (Lakhankar et al. 2006) and Bayesian inversion approaches (Notarnicola et al. 2008; Pierdicca et al. 2008). Also, in comparison to a Bayesian approach, Notarnicola et al. (2008) found that increasing the number of parameters used in the inversion decreased the accuracy of the Bayesian method, while increasing the accuracy of the ANN retrieval. Ancillary data, such as NDVI, LAI, optical depth and land-cover, are used to supplement backscatter and surface roughness to improve the inversion results (Lakhankar et al. 2006; Notarnicola et al. 2008; Said et al. 2008). When a separate ANN was trained for each season, it was found that soil moisture retrieval performance increased (Said et al. 2008), suggesting that an ANN trained for specific rather than general conditions may be a more capable tool for retrieving soil moisture. Baghdadi et al. (2002a) used multiple backscattering coefficients from different configurations (i.e. VV23°, HH39°, etc.) as input to the ANN and was able to simultaneously invert soil moisture and RMS height. This demonstrates one of the inherent advantages of the ANN over least squares regression methods, in that a single network, if provided adequate input, is better at simultaneously retrieving multiple surface parameters, even when the surface is slightly vegetated (Notarnicola et al. 2008; Said et al. 2008).

The use of fuzzy logic in soil moisture retrieval has been proposed by a few authors (Lakhankar et al. 2006; Verhoest et al. 2007a) as a method for soil moisture retrieval comparable to PDFs and also allowed for the computation of confidence intervals. A fuzzy set is designed to allow partial membership into a Boolean set based on the amount of correlation to observations. Lakhankar et al. (2006) found that in a classification algorithm fuzzy logic models were able to retrieve soil moisture slightly less accurately, but with more consistency than artificial neural networks.

B.3.7 Soil Moisture Retrieval Uncertainty

In light of the hydrological applications of soil moisture retrieval, it is important that some consideration be given to the quantification of uncertainty in retrieval results, an aspect which has been little explored. Sources of uncertainty arise from speckle, surface roughness characterization, forward models and inversion techniques. Even the choice of dielectric model can introduce uncertainty as high as 4% (Fernández-Gálvez 2008). The uncertainty caused by speckle is mitigated by averaging and filtering, but cannot actually be known since it is impossible to retrieve backscatter without speckle (De Keyser et al. 2012). To account for the contribution of speckle to uncertainty, Barber et al. (2012) use a polarimetric mutli-look technique to infer speckle statistical properties, which can be accounted for as a source of error in Bayesian retrieval. Lakhankar et al. (2012) found that uncertainty, represented as the variation in backscatter, was positively related to land-cover heterogeneity, which poses a theoretical obstacle to large scale filtering. When retrieving soil moisture from grasslands and wetlands using ASAR wide swath data, van der Velde et al. (2012) found that the parameterization of surface roughness was a larger source of uncertainty than soil moisture variability, vegetation heterogeneity or models errors. If the assumption can be made that surface roughness is the primary source of uncertainty, Monte Carlo methods are a potential mechanism to quantify uncertainty (De Keyser et al. 2012; Salama et al. 2012). Inverting soil moisture using Bayesian methods (Notarnicola et al. 2006; Pierdicca et al. 2010), possibility distributions (Verhoest et al. 2007b) or fuzzy logic (Lakhankar et al. 2006; Verhoest et al. 2007a) all have the advantage of inherently deriving soil moisture retrievals with uncertainty. Unfortunately, the support of the retrieval distributions may be unacceptably large and require refinement before they can be considered for operational use (Vernieuwe et al. 2011). From an operational perspective,

accounting for uncertainty is a critically important topic, which requires a focused research effort, especially as a consensus begins to emerge towards accepted soil moisture retrieval techniques.

B.3.8 Improvement of Retrieval through an Initial Guess and Ancillary Data

The spatial and temporal variability of soil moisture combined with the multiple influences and speckle inherent in SAR elicits a situation where accurate retrieval of soil moisture from a backscattered radar signal is difficult. In these circumstances, the incorporation of *a priori* information in the retrieval algorithm to either constrain the range of potential soil moisture outcomes or supplement retrieved data to improve the retrieval result may be beneficial. One method of constraining the retrieval algorithm to a likely outcome is the use of possibility and probability distributions to determine the most probable value of soil moisture from observed backscatter and the probable state of the surface roughness (Notarnicola et al. 2008; Pierdicca et al. 2008; Pierdicca et al. 2010; Verhoest et al. 2007b; Vernieuwe et al. 2011). Baghdadi et al. (2002a) improved retrieval from an ANN by introducing a constraint through weather forecasts and field knowledge. The analogous data was used to classify the soil condition as either moist or very moist prior to input into the ANN.

Passive remote sensing offers a source of commensurate information which can aid in the retrieval of soil moisture. Radar emissions are more strongly correlated to soil dielectric properties than active SAR, but can only be collected at coarser resolution (~30-40 km) (Anderson and Croft, 2009). Low resolution radiometer data has been combined with high resolution SAR data through a Bayesian framework (Zhan et al. 2006) and through time series analysis (Narayan et al. 2006). In the case of the latter, SAR was used to determine the relative amount of change in soil moisture within the footprint of a radiometer. This method is planned for SMAP where active and passive data will be simultaneously acquired and radiometer data will be disaggregated with SAR (Entekhabi et al. 2010). The result is the retrieval of soil moisture which is more accurate than with SAR alone, at the cost of a compromise in spatial resolution (Narayan et al. 2006; Zhan et al. 2006).

Another potentially powerful source of *a priori* information comes from hydrological models. The hydrologic model is used to create an ‘initial guess’ of soil moisture conditions. Using a Bayesian approach, Mattia et al. (2006) combined a simplified Thornwaite lumped water

balance model on a monthly scale with an IEM inversion algorithm. The quality of the initial guess proved to be very important in reducing the error in the retrieval. It was found that a poor quality 'initial guess' decreased RMSE from 9.6 % to 5.9 %, whereas a high quality 'initial guess' reduced the error to 2.8 % (Mattia et al. 2006). Advanced hydrological models such as the TOPMODEL land-atmosphere transfer scheme (TOPLATS) (Famiglietti and Wood, 1994) and the Process Oriented Multiscale Evapotranspiration Model (PROMET) (Mauser and Schadlich, 1998) have also been combined with SAR to improve the accuracy of soil moisture estimation (Mattia et al. 2009). The model data were combined with multi-temporal 25 km^2 resolution L band SAR images to derive soil moisture at 0.01 km^2 resolution (Mattia et al. 2009). In all cases, incorporation of *a priori* information improved the accuracy of the soil moisture retrieval.

B.3.9 Advantages of Multi-configuration SAR

With the increasing availability of SAR data, more emphasis is being placed on the retrieval of soil moisture from multi-configuration data. It has long been clear that a single radar configuration (i.e. C Band, 23°, HH) was not sufficient to accurately retrieve soil moisture from distributed surfaces (Altese et al. 1996). Modern satellites such as ENVISAT ASAR, RADARSAT 2 and ALOS PALSAR are capable of imaging the surface with multiple polarizations from multiple angles with a relatively short revisit time. These advances have allowed for the retrieval of soil moisture from solely remote sensing sources (i.e. Wang et al. 2011).

Multiple polarization radar allows for the simultaneous acquisition of two (ASAR, RADARSAT 1/2, PALSAR) or four (RADARSAT 2) polarizations (i.e. HH, VV, HV, VH). The availability of multi-polarized data has allowed for the application of the Oh model (Oh, 2004), which requires the co-polarization and depolarization ratios, as well as polarimetric decomposition. Perhaps the greatest operational advantage of multiple polarizations is the calculation of the co-polarization ratio (HH/VV) and the depolarization ratio (HV/VV), which can be used to quantify vegetation parameters (McNairn and Brisco, 2004). The depolarization ratio is useful for determining crop height and plant density (Gherboudj et al. 2011), while cross-polarization is more closely related to NDVI than other forms of backscatter (Bindlish and Barros, 2001). Despite this advantage over vegetated soil, when soils were bare Baghdadi et al. (2006) found that using ASAR dual-polarization data only improved soil moisture retrieval

accuracy by ~1%, whereas Notarnicola et al. (2008) noted that multi-polarization configurations only improved retrieval results under dry soil moisture conditions. Despite this, polarimetric SAR offers great promise for soil moisture retrieval. Polarimetric methods model the amplitude and phase difference in the backscatter response of different polarizations. By decomposing the polarimetric response, the individual contributions of vegetation, surface roughness and soil moisture can be separated (Iodice et al, 2011; Hajnsek et al. 2009; Jagdhuber et al. 2009). Unfortunately, polarimetric decomposition is currently limited to simple models, and has only recently begun to be explored in terms of soil moisture research.

Of greater importance in terms of soil moisture retrieval is the ability of satellites to collect data at multiple incidence angles (Baghdadi et al. 2006; McNairn and Brisco, 2004; Notarnicola et al. 2008; Sahebi et al. 2002; Srivastava et al. 2003; Zribi and Deschambre, 2002; Zribi et al. 2005). The greatest advantage of multiple incidence angle data with conventional inversion techniques is the ability to quantify surface roughness without the need for *in situ* measurements. Between two different incidence angles, the impact of soil moisture on the difference in backscatter is negligible compared to that of surface roughness (Srivastava et al. 2003; Zribi and Deschambre, 2002; Zribi et al. 2005). This allows for the determination of surface roughness if both a high and low incidence angle SAR image is available over a time period in which soil moisture has not changed. Utilizing high and low incidence angle data allowed Wang et al. (2011) to fully parameterize surface roughness (with an assumed ACF) and soil moisture using the *Zs* index (Zribi and Deschambre, 2002) and the calibrated correlation length (Baghdadi et al. 2006).

Currently, simultaneous acquisition of multi-incidence angle data is only possible with Terra SAR-X and TanDEM-X. Otherwise, collecting backscatter from an area is only possible at one incidence angle per orbit. This results in a necessary delay between the images at low and high incidence angles of typically 3 to 5 days during which time the soil moisture hydrology may change significantly. Most studies using satellite based SAR assume that the change in soil moisture between the two acquisitions is negligible and care must be taken to ensure that precipitation has not occurred between scenes. In order to improve the retrieval of soil moisture, Baghdadi et al. (2008) suggested that future satellites and the launch of more SAR satellites will decrease the temporal distribution between low and high incidence angle scenes. Planned satellite constellations such as Sentinel-1, RADARSAT Constellation and SAOCOM incorporate

multiple SAR satellites decreasing the revisit time (Attema et al. 2009; Flett et al. 2009; Torres et al. 2012). The decrease in revisit time of satellites with the same configurations will also enhance the possibilities of change detection techniques.

B.4 Hydrological Applications: Advances and Key Issues

B.4.1 Watershed Characteristics and Soil Moisture Variability

Retrieval of soil moisture from SAR requires several necessary assumptions about the spatial characteristics of soil moisture at the SAR scale. It should be noted at this point that each of the backscatter models discussed so far (IEM, Oh, Dubois) all explicitly account for the soil structure variability, while soil moisture is assumed either homogenous or to represent the mean soil moisture state. This assumption is often accounted for in the retrieval process rather than the model, where soil moisture is inverted as a mean at the field scale or larger (Baghdadi and Zribi, 2006; Lievens and Verhoest, 2011; Paloscia, 2002; Verhoest et al. 2007b; and others). The central limit theorem gives the basis for these assumptions, where small scale non-linear variability is expected to average out and behave in a linear fashion as scale increases (Quesnel et al. 2000; Sivapalan and Wood, 1986). However, the presence of non-random influences on soil moisture structure causes only some non-linearity to be averaged out with scale (Blöschl and Sivapalan, 1995). Therefore, an appropriate discussion of the hydrological value of SAR soil moisture begins with a discussion of the spatial structure and variability of soil moisture at the sub-catchment scale.

B.4.1.1 Soil moisture spatial variation

Soil moisture spatial variation is known to increase with observation scale (Das and Mohanty, 2008; Famiglietti et al. 2008), where the variability is related to the mean soil moisture content for a given scale as a result of dominant surface processes at work. Some authors have found that soil moisture variability decreases with increasing soil moisture content (Brocca et al. 2010; Choi and Jacobs, 2007; Famiglietti et al. 1999; Hupet and Vanclooster, 2002; Western et al. 2004; and others), while others have found the opposite trend of increasing variability with mean soil moisture (Famiglietti et al. 1998; Vivoni et al. 2008; Western et al. 1998; Western et al. 2004; and others). Despite the literature showing some contention based on local conditions, a

consensus is beginning to emerge on the processes and relationships between surface soil moisture and its spatial variability. At sub-field scales (~50m-200m) spatial structure and dominant processes are difficult to establish. Famiglietti et al. (2008) found that at scales of less than 100 m soil moisture mean and standard deviation were unrelated, but noted under-sampling as a problem. Famiglietti et al. (1998, 1999) found that the spatial distribution of soil moisture was influenced by landscape features on hillslopes, whereas Hupet and Vanclooster (2002) determined that soil spatial structure and topography were unimportant factors and LAI influenced soil moisture variability through evapotranspiration. At field scales, the spatial variability of soil moisture changes with its mean as a result of the transition between wet and dry conditions and the regional climate (Brocca et al. 2007). In a review of several soil moisture datasets Brocca et al. (2007) point out that across various scales, in semi-arid environments soil moisture variability increases as the soil becomes wetter, whereas humid environments tend to have greater variability during drydown. The controlling influence of this relationship is the pattern of precipitation. When the local climatology results in heterogeneous precipitation, the non-uniform wetting results in increased soil moisture variability (Brocca et al. 2007; Cosh et al. 2004; Famiglietti et al. 2008; Vivoni et al. 2008), whereas when precipitation is homogeneous, and at smaller scales, precipitation results in a decrease in the variability of soil moisture (Brocca et al. 2007; Cosh et al. 2004; Das and Mohanty, 2008; Jacobs et al. 2004; Ryu and Famiglietti, 2006; Wilson et al. 2003). For the purposes of further discussion, it will now be assumed that at the typical scale of SAR soil moisture retrievals (approx. < 1 km^2) precipitation results in decreasing variability. Shortly following wetting, soil moisture variability is dominated by simple scaling processes resulting from differences in porosity, hydraulic conductivity and topography (Das and Mohanty, 2008; Famiglietti et al. 1999; Famiglietti et al. 2008; Heathman et al. 2009; Peters-Lidard et al. 2001). During drydown, soil moisture variability transitions to have a greater influence from multi-scale processes where vegetation type, soil moisture redistribution, drainage and texture dominate, thereby increasing the variability of soil moisture at moderate wetness conditions (Bosch et al. 2006; Das and Mohanty, 2008; Entin et al. 2000; Famiglietti et al. 2008; Peters-Lidard et al. 2001). Continued drying results in a decrease in soil moisture variability, where the preserved variation is a result of soil texture, particularly clay content, and the concavity of the landscape (Das and Mohanty, 2008; Famiglietti et al. 1998). The dominant processes determining the variation of soil moisture are the result of a change

from control of infiltration in wet conditions to control by drainage and evapotranspiration in dry conditions (Peters-Lidard et al. 2001). Therefore, the overall relationship between mean soil moisture and the coefficient of variation is concave down with the smallest amount of variability during wet and dry conditions (Ryu and Famiglietti, 2005). This soil moisture variation relationship has implications for when SAR soil moisture would have the most potential benefit for hydrological applications. Since watershed scale precipitation (if relatively uniform) and long dry periods tend to homogenize soil moisture, the benefit to be realized from the spatially distributed SAR retrieved soil moisture, as compared to less costly radiometer derived soil moisture, may not be significant. However, the spatial variability introduced at intermediate soil moisture states is a potential source of uncertainty which can be mitigated by spatially distributed soil moisture as derived from SAR.

B.4.1.2 Soil moisture temporal variation

The concept of temporal stability was introduced by Vachaud et al. (1985) and states that soil moisture at different locations, while variable in time, exhibit stability of rank when compared to the mean of other surrounding locations. That is, when several sites are ranked based on soil moisture, while the mean soil moisture may change, the ranking of any individual sites with respect to the others remains relatively constant in time. This concept has implications in terms of soil moisture retrieval from the perspective of sampling validation data and more importantly from the perspective of the retrieval model and hydrological applications. By sampling at a mid-rank representative field site Brocca et al. (2010) were able to collect 10 samples from a field which represented the catchment (60 km^2) soil moisture with an R^2 of 0.96 and RMSE of 2.38%.

Many authors have noted soil moisture measurements displayed temporal stability at both the field and catchment scale (Bosch et al. 2006; Cosh et al. 2004; Entin et al. 2000; Vivoni et al. 2008). Despite the propensity for temporal persistence shown in various studies, the strength of the temporal stability remains difficult to predict by external indicators. Topography influences temporal stability, where topographically flat areas have been found to have poor stability of temporal variations (Mohanty and Skaggs, 2001), whereas on mountainous terrain time stable locations were located at mid elevations while high and low elevations were relatively less stable (Vivoni et al. 2008). Additional factors influencing temporal stability include soil characteristics (Jacobs et al. 2004), which change only slowly, and vegetation (Vivoni et al. 2008) which plays

only a minor role in temporal persistence (Cosh et al. 2004). Unfortunately, temporal persistence may be interrupted at different spatial and temporal scales due to disturbances such as rainfall or vegetation growth at large scales and overland flow or land management at small scales (Cosh, et al. 2004; Mohanty et al. 2000a; Mohanty et al. 2000b). This naturally leads to two primary questions, the first being how long does it take for the temporal pattern to re-emerge following a disturbance and second being about the reliability of temporal stability for hydrological applications. The answer to the first question is naturally dependent on watershed characteristics, while the second is related to the mean soil moisture state. For example, Parada and Liang (2008) were able to assimilate coarse soil moisture data disaggregated with temporal ratios and found recovery of information lost by coarse soil moisture data alone. However, this advantage was lost following precipitation which disturbed the rank stability relationship (Parada and Liang, 2008). Notwithstanding that limitation, the use of temporal stability has some hydrological merit but requires parameterization without *in situ* measurement to be useful at the watershed scale or larger. Since backscatter is related to soil moisture, temporal stability patterns in soil moisture are also detectable in SAR backscatter signals and can therefore be derived in the absence of ground measurements using change detection methods for sensors such as ASAR (Wagner et al. 2008).

B.4.1.3 Soil moisture distribution

Remotely sensed soil moisture retrieves a spatially averaged value of soil moisture at the spatial resolution of the sensor, or the speckle filter applied. The spatial averaging neglects the sub-pixel variability inherent in surface soil moisture. Thus, it is important from the perspective of validation of soil moisture retrievals, particularly at the field scale, that the selected PDF of the soil moisture observations is consistent with that being observed by the sensor (Ryu and Famiglietti, 2005). Often soil moisture distribution is assumed normal, and that multiple point samples will represent the mean field scale soil moisture as determined from the remote sensing platform (Baghdadi and Zribi, 2006; Lievens and Verhoest, 2011, Paloscia, 2002; Verhoest et al. 2007b; and others). While normality is generally assumed, the behaviour of soil moisture distributions during wetting and drying are difficult to predetermine. Under moderate to wet conditions Wilson et al. (2003) found that field sites in New Zealand were normally distributed. During the Southern Great Plains ’97 field experiment it was found that on approximately 36%

of sampling dates the soil moisture distribution was non-normal (Famiglietti et al. 1999) and similarly in central Italy when considering volumetric soil moisture nearly half of measured sites were non-normal ($p = 0.05$) (Brocca et al. 2007). At the field scale, soil moisture distribution was unimodal and skewed based on soil moisture conditions. Since the soil moisture distribution is bounded by the wilting point and saturation point, the distribution of soil moisture has been found to become increasingly skewed and non-normal as the mean soil moisture value approaches either boundary (Das et al. 2008; Famiglietti et al. 1999). Only when soil moisture was in an intermediary state was it normally distributed (Das et al. 2008; Famiglietti et al. 1999). At catchment to watershed scales soil moisture is less likely to be normally distributed as soil moisture distribution is no longer the result of random processes but is rather structured by vegetation, topography, precipitation and soil characteristics. The influence of structure on the soil moisture distribution causes a shift towards heavy-tailed distributions and multimodality. As a result of this, Ryu and Famiglietti (2005) suggest at larger scales that soil moisture distribution is best described by multiple normal distributions which are superimposed upon each other as soil moisture exceeds 18% and by the beta distribution below that threshold. Therefore, the use of spatially averaged soil moisture to validate remote sensing observations may produce a small bias at the field scale when conditions are wet or dry and uncertainty exists for coarse resolution or aggregated SAR data (i.e. Kong and Dorling, 2008; Le Hégarat-Mascle et al. 2002; Thoma et al. 2006). By extension, the above discussion also suggests the field scale as optimal for soil moisture retrieval that is coarse enough to reduce radiometric uncertainty while being fine enough to not introduce uncertainty as a result of hydrological processes.

B.4.2 SAR Retrieval Error and Soil Moisture Scaling Issues in Hydrology

The single major advantage of retrieving soil moisture from SAR sensors is the high spatial resolution, which comes at a significant cost of (currently) low temporal resolution and relatively poor radiometric resolution. This is the reason for the recent trend where platforms such as AMSR-E, SMOS and to some extent SMAP have received much greater attention, as side by side comparisons show radiometer retrievals to have significantly lower retrieval error than SAR (Bindlish et al. 2009; Bolton et al. 2003). With respect to retrievals, and in the absence of uncertainty in surface roughness and vegetation, errors largely result from the influence of speckle. Speckle emerges from random constructive and destructive wave interference as the

signal travels between the target and antenna (Mattia et al. 2003a; Thoma et al. 2008) resulting in fluctuations between adjacent pixels of up to 18 dB (Ulaby et al. 1982). In order to compensate for the effects of speckle, radar images are regularly filtered in order to determine the 'true' backscattering coefficient using one of many potential speckle reduction filters (See Lee et al. 1994 for a review). The effect of filters is to capture the central tendencies of a group of pixels, thereby negating speckle. Also common, is the retrieval of soil moisture at the pixel scale and averaging soil moisture values across larger scales (Baghdadi and Zribi, 2006; Lievens and Verhoest, 2011; Paloscia, 2002; Verhoest et al. 2007b; and others). In essence, both methods result in a tradeoff, where spatial resolution and/or soil properties heterogeneity are sacrificed in order to compensate for poor radiometric resolution (Barber et al. 2012; Thoma et al. 2006, 2008). As the spatial scale increases to catchment scale, the variability of surface characteristics is reduced and backscatter is more sensitive to changes in soil moisture (Alvarez-Mozos et al. 2005). This becomes a difficult balance, as decreasing spatial scale decreases soil moisture variability making soil moisture retrieval more representative of the hydrological state. This must be balanced with increasing spatial scale through aggregation, which increases retrieved soil moisture accuracy (Alvarez-Mozos et al. 2005), at the cost of simultaneously increasing sub-grid soil moisture variability and uncertainty. In essence, the balance described is between hydrological and radiometric uncertainty. Famiglietti et al. (2008) characterize soil moisture variability as a function of spatial scale as

$$Var(S) = C \cdot S^{D} \tag{B-11}$$

where S is the log of the spatial scale, C is a parameter and D is the fractal dimension determined to be 0.86. Since the variability increases exponentially with scale, the goal of soil moisture retrieval from SAR must be to balance the need for accurate soil moisture retrieval without compromising the benefit of higher resolution.

Thoma et al. (2006, 2008) studied the impact of using a median filter to reduce the effect of speckle on soil moisture retrievals using change detection, theoretical and semi-empirical models. They found that a 7 X 7 median filter best represented the mean backscatter value, whereas a 15 X 15 median filter removed most of the variability induced by speckle (Thoma et al. 2006). Despite these findings, they opted to use a 5 X 5 filter as a balance to maintain the spatial heterogeneity of soil moisture. However, filtering alone was not adequate to reduce the errors in the soil moisture retrieval and spatial averaging was still required (Thoma et al. 2008).

The net result was that to have an acceptable level of accuracy in soil moisture retrieval (+/- 0.05 m^3m^{-3}) a combination of median filtering and spatial averaging reduced the ground resolution to 25 to 160 times that of the SAR pixel resolution (Thoma et al. 2008). The specific size of the effective resolution changes with surface characteristics of the field, where greater variability leads to coarser effective resolution. In order to maintain the fine scale soil moisture features it is necessary to use hyper-resolution sensors such as TerraSAR-X (1 m) (Anguela et al. 2010; Aubert et al. 2011) or to minimize the effects of speckle without compromising spatial resolution. Along this line, Doubkova et al. (2012) suggest that the upcoming launch of Sentinel-1 with a radiometric resolution greater than ASAR (Torres et al. 2012) will have a positive impact on the potential for soil moisture retrieval.

The above has implications for field sampling campaigns carried out to verify soil moisture retrievals. At the field scale, in order to estimate soil moisture to within 2% of the true mean, between 4 to 15 samples are required for dry conditions and 15-40 samples in intermediate conditions decreasing again to 4 to 15 samples in extremely wet conditions (Brocca et al. 2010; Famiglietti et al. 2008; Jacobs et al. 2004). The variability in the number of samples required to characterize soil moisture changes as the variability of soil moisture increases at intermediate states of wetness (Brocca et al. 2010; Brocca et al. 2007; Famiglietti et al. 2008; Jacobs et al. 2004). Sampling depth is also an important consideration as the maximum variability of soil moisture occurs at the 5 cm depth (Choi and Jacobs, 2007), which is inconveniently the penetration depth of most SAR sensors in C and L band (Ulaby et al. 1986). The large range of variability experienced at the field/catchment scale brings into question the appropriateness of using a static goal for the acceptable error in soil moisture retrieval. Often a goal of 4-5% is considered an acceptable level of accuracy, as this is presumed to be representative of the variability at the watershed scale (Kerr et al. 2010). However, soil moisture variability is known to vary as a function of the mean value with the greatest variability, and presumably the largest potential for retrieval error, at intermediate values with less variability during dry and wet periods. Assuming that sub-pixel soil moisture variability is normally distributed, a variable error target can be established where the threshold of acceptable error ε_{target} is given by

$$\varepsilon_{target} = \frac{f\left(x;\mu,\sigma^2\right)}{\exp\left(2\varepsilon_{sat}\left(\sigma\right)^{-1}\right)} + \varepsilon_{sat} \tag{B-12}$$

where the function *f* represents the standard normal distribution about the mean μ of watershed scale soil moisture for pixel value *x* normalized by the standard deviation σ and shifted by the radiometric sensitivity of the sensor to soil moisture ε_{sat}. The parameter ε_{sat} can be derived as the product of the slope of a regression analysis between soil moisture and backscatter and the radiometric accuracy of the sensor. Therefore in wet and dry conditions which experience less heterogeneity the target error is closely related to the radiometric accuracy of the sensor, and the allowable error increases at intermediate and highly heterogeneous soil moisture states. Figure B-4 shows the target error function for sub-watershed soil moisture retrieval applied to select measurements from the 1120 km^2 Seine River watershed (Aubert et al. 2003). Using the proposed function, greater watershed variability results in less stringent accuracy requirements, with wet and dry extreme areas have the most strict retrieval targets as the sub-pixel variability is lowest.

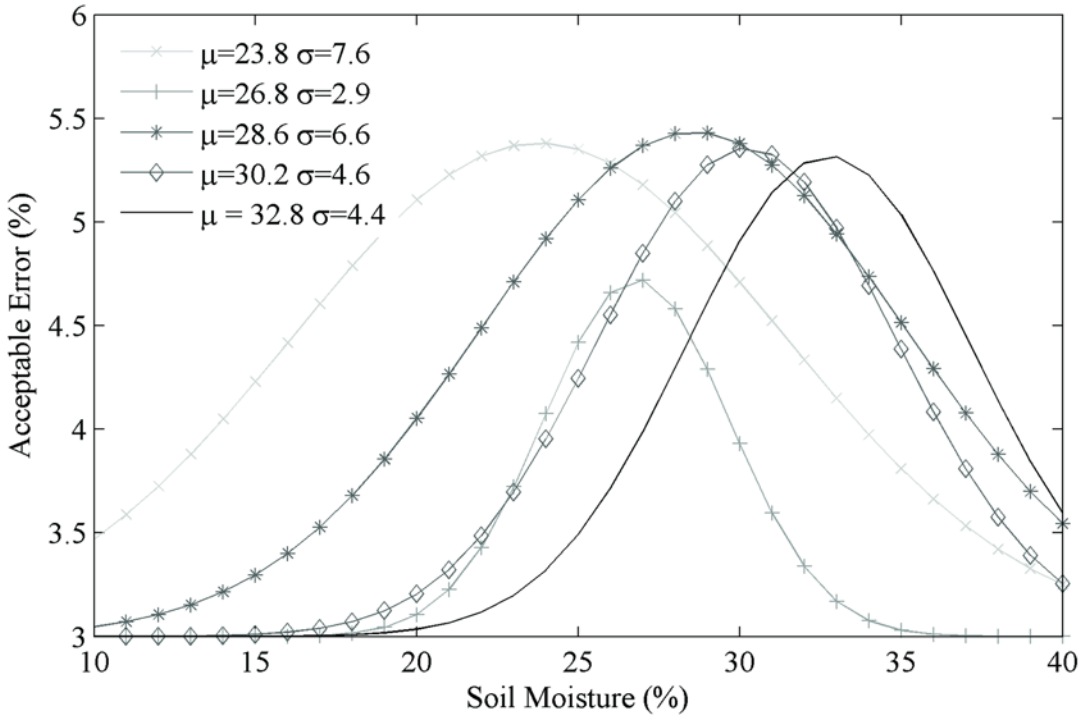

Figure B-4: Soil moisture retrieval target error accounting for variability. The target error assumes a normal soil moisture distribution, a sensor radiometric accuracy of 1dB and the slope of the soil moisture regression is 0.3%/dB using watershed characteristics from the Seine River Watershed, France (Aubert et al., 2003)

B.4.3 Advances in Applications of SAR Soil Moisture in Hydrology

B.4.3.1 Requirement for High Resolution Soil Moisture

Following the launch of SMOS (Kerr et al. 2010) and AMSR-E (Njoku et al. 2003), which both provide operational soil moisture retrieval algorithms, attention has arguably shifted away from active radar remote sensing for providing soil moisture products. This shift is natural, considering the higher temporal resolution (3 days for SMOS) and higher accuracy of soil moisture retrievals. However, the radiometer spatial resolution is inadequate for many hydrological purposes. Using a soil moisture product at a coarse resolution results in significant error as a result of not properly accounting for spatial variability (Crow and Wood, 2002). As was discussed previously soil moisture variability is a product of stochastic forces operating over an organized landscape, where the true variability of soil moisture can only be represented when both components are considered (Bronstert and Bardossy, 1999). Complicating the interaction between organized and stochastic influences on soil moisture patterns, is the strong non-linear relationship between soil moisture and hydrological processes (Brocca et al. 2009; James and Roulet, 2009; Merlin et al. 2006; Penna et al. 2011) where, the stronger the organized heterogeneity and non-linearity, the greater the requirement to account for fine scale heterogeneity to accurately represent hydrological processes (Bronstert and Bardossy, 1999; Merlin et al. 2006). Unfortunately, the variability of soil moisture within a watershed or catchment is often not considered and the assumption that the mean state is representative is applied in its stead (Ivanov et al. 2010), partially due to lack of available data and partially due to the lack of adequate high resolution models (Wood et al. 2011).

Integration of high resolution soil moisture into hydrological applications provides many potential benefits including correcting for weak model state interactions/variability (Alavi et al. 2010; Bronstert and Bardossy, 1999; Merlin et al. 2006; Mohanty et al. 2000b; Mohr et al. 2000), quantifying sub-grid variability (Das et al. 2008; Crow and Wood, 2002), parameterization and calibration (Koren et al. 2008; Merz et al. 2011), reducing uncertainty (Loew et al. 2009; Zehe and Blöschl, 2004) and state initialization (Koster et al. 2010a; Koster et al. 2010b; Mahanama et al. 2008; Mohr et al. 2000). Despite these many potential benefits, the greatest impediment to the use of SAR as a source of high resolution soil moisture data for hydrological applications is the scarcity of data (Weisse et al. 2001). For this reason, many of the applications discussed in the following section show the benefits of high resolutions soil moisture, but are the result of

modelling exercises, ground/airborne remote sensing and in some instances coarse resolution spaceborne remote sensing.

B.4.3.2 Assimilation of Soil Moisture

Hydrological models are dependent on the availability and accuracy of a variety of input data for simulating surface water and energy fluxes. Uncertainty in meteorological data, soil moisture initial state, and land surface parameters causes divergence between simulation and reality sometimes leading to significant errors. The goal of data assimilation is to reduce the uncertainty which may exist in the model by integrating observations, as imperfect as they may be, of the state which the model represents, in this case soil moisture. The skill of the model assimilation product results from the synergetic relationship between the skill of the model and the quality of the soil moisture data/retrieval (Reichle et al. 2008). Within the context of this review, the potential for the assimilation of SAR retrieved soil moisture must be considered taking into account the advantage of high spatial resolution at the cost of lower accuracy and temporal resolution compared to available radiometer data. Many studies have shown that in general, assimilation of observed data, even poor or coarse data, provides at least a modest improvement over open loop models (Alavi et al. 2010; Loew et al. 2009; Pan and Wood, 2010; Parada and Liang, 2008; Reichle et al. 2008). As is expected, the natural result of higher skill soil moisture retrievals is greater improvement in the modeled product when compared to open loop simulations (Pan and Wood, 2010; Reichle et al. 2008). What may not be expected is that the performance of soil moisture assimilation is more dependent on spatial availability than retrieval accuracy (Pan and Wood, 2010). Pan and Wood (2010) found that at the watershed scale, in decreasing order, assimilation skill was most strong related to spatial availability, revisit time and retrieval accuracy, suggesting a significant benefit to high resolution soil moisture products. Unfortunately, few examples exist of assimilation of actual high resolution soil moisture data comparable to SAR (i.e. Das et al. 2008; Merlin et al. 2006; Parada and Liang, 2008; Pauwels et al. 2002).

The accurate representation of surface energy fluxes is critically important in climate and weather prediction. At the continental scale, there are many examples of the assimilation of soil moisture data to improve global climate model outputs using data from SMOS, AMSR-E, etc. (Draper et al. 2011; Draper et al. 2012; Kumar et al. 2008; Reichle et al. 2008; Reichle and

Koster, 2005; Walker and Houser, 2001; and others). Assimilation of coarse soil moisture data lowers the error in soil moisture storage (Walker and Houser, 2001) increasing the correlation between modeled storage and observations by 10% (Draper et al. 2012). Despite the improvements found by assimilation of coarse soil moisture data, neglecting the fine scale heterogeneity of soil moisture limits the increase in assimilation skill that could otherwise be realized (Crow and Wood, 2002a; Crow and Wood, 2002b). This is because the error in latent energy flux at the regional scale increases proportionately to the variance of field-scale soil moisture within the region (Crow and Wood, 2002b). The impact of this is found in sub-regional differences in soil moisture control of surface energy flux. For example, weak thermal conductivity in a model can hamper evapotranspiration causing the overestimation of soil moisture, leading to overestimation of convective precipitation (Mohr et al. 2000). Similarly, the variability of rainfall has an impact on boundary layer conditions at different locations resulting in heterogeneous surface fluxes and misrepresenting the genesis of convective storms (Mohr et al. 2003). In both cases, poor or under-representation of soil moisture and its spatial organization has a detrimental impact on the capacity of a numerical weather prediction model to predict near time weather events. After disaggregation of coarse soil moisture data, Merlin et al. (2006) were able to show that assimilation of 1 km soil moisture values represented *in situ* measured soil moisture conditions better than either microwave or model predicted soil moisture alone. The result to the model was the removal of some of the uncertainty inherent in coarse resolution weather data, particularly through the improvement of latent heat flux estimates (Merlin et al. 2006). At even finer resolution Alavi et al. (2010) found that assimilating soil moisture into an LSM improved latent heat flux by 14% with greater gains made during the start of the growing season. The improvements over the model alone only occurred with the assimilation of spatially distributed soil moisture and not with assimilation of mean soil moisture (Alavi et al. 2010). Much of the error surrounding latent heat flux is a result of the poor characterization of rainfall distribution in poorly gauged basins (Loew et al. 2009). Even when using spatially distributed satellite observations of precipitation there is still a high degree of uncertainty in the precipitation. The assimilation of soil moisture modifies soil water storage, compensating for precipitation error and reducing overall model uncertainty (Bolton et al. 2010; Loew et al. 2009). Streamflow and runoff modelling also has the potential to benefit greatly from SAR retrieved soil moisture. Often hydrological models approximate the soil moisture state by a mean value,

which is representative for an entire catchment (Ivanov et al. 2010). At the field scale, the assumption of mean soil moisture as representative can underestimate peak runoff during a storm by greater than 50% (Minet et al. 2011). The discrepancy is significant at the field scale and represents an even greater uncertainty at the watershed scale when mean soil moisture conditions are assumed. Soil moisture patterns that accurately account for the structure of soil moisture provide more realistic runoff values (Minet et al. 2011). The importance of the spatial structure of soil moisture to runoff response is important at micro and catchment scales because different processes dominate the runoff response (Zehe and Blöschl, 2004). The result is that different spatial patterns of soil moisture within a watershed can produce different runoff responses (Zehe and Blöschl, 2004), even when the mean and statistical distribution of soil moisture is the same (Zehe et al. 2005). The greatest potential gain in terms of assimilation of soil moisture for runoff modelling occurs when the soil moisture state is at an intermediate level. When conditions are persistently dry or in arid regions the soil moisture distribution becomes more homogenous and skewed (Das et al. 2008; Famiglietti et al. 1999) and the low variability of soil moisture results in antecedent conditions lacking importance in the prediction of runoff (Zhang et al. 2011). Similarly, in wet conditions soil moisture becomes homogenous near saturation and the importance of soil moisture spatial distribution decreases (Zehe et al. 2005). Runoff is least predictable in the intermediate range of soil moisture as large spatial variations (Familietti et al. 1999) and runoff threshold behaviour (Zehe et al. 2005) cause instability in the runoff response (Zehe and Blöschl, 2004). Using a lumped model, Aubert et al. (2003) were able to show that assimilation of soil moisture was effective for improving stream flow prediction and noted that soil moisture assimilation was particularly effective for flood forecasting. Using a distributed model, the assimilation of soil moisture was found to significantly improve forecasts for large floods decreasing the error from 25% to 12% for a 3 hour forecast and 25% to 19% for a 48 hour forecast (Komma et al. 2008). The benefit of assimilation is particularly noted when assimilation begins with low flow conditions (Komma et al. 2008). Pauwels et al. (2002) assimilated soil moisture retrieved from ERS using the IEM and Oh Model into TOPLATS to model runoff behavior in Belgium. Despite the relatively high RMSE (5.7%) of the retrieved soil moisture, assimilation significantly improved simulated discharge (20-60%) in three of five catchments while only slightly decreasing (5-10%) discharge accuracy in the remaining two catchments (Pauwels et al. 2002). These demonstrations show the potential of assimilation of SAR retrieved

soil moisture for hydrological forecasts, but the fundamental shortcoming remains the availability of data as the temporal resolution and operational capacity of current SAR sensors is limiting.

B.4.3.3 Model Initialization

Notwithstanding the limitation of current sensors for assimilation due to poor repeat coverage, non-temporal soil moisture data may be useful in the context of initializing models, or providing sparse updates to re-establish correct spatial heterogeneity. Realistic initialization of soil moisture has been shown to improve climate simulations (Koster et al. 2010), positively impacting model skill at lead times of over six months (Mahanama et al. 2012). The greatest benefit to the climate model from soil moisture initialization comes in terms of improvements in temperature simulation as a result of better constrained energy fluxes, while the improvement to precipitation modelling is generally minimal (Bisselink et al. 2011; Koster et al. 2010). However, in the absence of hyper-resolution models, many climate models operate at a resolution of 50 km or greater resulting in passive sensors such as AMSR-E (Bisselink et al. 2011) being more appropriate for climate simulation.

Perhaps a more appropriate current application of SAR soil moisture is the initialization of flood forecasts. There is a strong link between flood formation and antecedent soil moisture (Aronica and Candela, 2004), where the spatial distribution of soil moisture is important for properly characterizing the flood hydrograph (Minet et al. 2011). Realistic initialization is particularly important when runoff genesis exhibits threshold behavior (Penna et al. 2011; Zehe et al. 2005), where standard model initializations may not adequately represent the difference between the state and the threshold at which runoff is induced and the state of antecedent wetness is therefore also strongly linked to catchment response time. Due to the influence of the runoff threshold and the representation of soil moisture heterogeneity it is important to properly account for the contribution from variable source areas (Penna et al. 2011).

B.4.3.4 Parameterization and Calibration of Hydrological Models

Prior to now the discussion has focused primarily on the potential of high resolution soil moisture data to improve hydrologic models through incorporation of the observed soil moisture during simulation and forecasting. Within this framework, soil moisture is used to artificially

force the model state to a better representation of the true state in order to compensate for model errors and input data limitations (Alavi et al. 2010; Loew et al. 2009; Merlin et al. 2006; Reichle et al. 2008; and others). However, as Reichle et al. (2008) point out, the skill of techniques such as data assimilation are also strongly dependent on the skill of the model, which may also potentially be improved by use of soil moisture data. This is done through incorporation of soil moisture into the calibration scheme of the model, in order to increase the confidence in model parameters and decrease the potential of equifinality (Koren et al. 2008). Typically, the calibration/parameterization of operational hydrological models is carried out by calibrating the model parameters based on a historically observed hydrograph. Koren et al. (2008) found that when the Sacremento (SAC) model was calibrated in several basins with only the hydrograph there were significant deficiencies in the model, whereas the inclusion of basin averaged soil moisture improved runoff simulation by 45%. The strongest benefit of this calibration procedure was found in dry watersheds, when the interconnection between runoff and soil moisture was weak (Koren et al. 2008; Wooldridge et al. 2003). The use of soil moisture as an additional calibration parameter is especially useful in ephemeral catchments where hydrograph calibration may be limited by sparse data, but the model can be calibrated to better represent evapotranspiration or other processes (Wooldridge et al. 2003). The temporal limitations in calibrating models to ephemeral streams is not as prevalent for soil moisture calibration as temporally discontinuous soil moisture measurements that represent a wide range of conditions, which better trains a model than long homogeneous records (Wooldridge et al. 2003). Calibration with spatially structured soil moisture is also important to account for variable source areas within the catchment. Zehe et al. (2005) point out that the spatial pattern of macroporosity is very important, as it is a controlling factor of threshold behavior; but unfortunately is typically parameterized using sparse field measurements supplemented by expert knowledge. More difficult to predict than the spatial heterogeneity of model parameters is the temporal heterogeneity of parameters. The use of historical flows for model calibration assumes stationarity in the catchment, whereas global climate change is known to be altering the hydrological cycle requiring compensation for this effect in hydrological model parameters (Merz et al. 2011). For example, the use of historical flows to calibrate a lumped model in Austria has been found to expose significant bias (15-35%) in recent simulations as the parameter controlling runoff doubled on average every thirty years (Merz et al. 2011). Since this

finding would suggest the calibration window is relatively short, alternate sources of information may be effective in supplementing the hydrograph, particularly for parameterization of evapotranspiration which Merz et al. (2011) discuss as being one of the factors which are most important for altering hydrological behavior and for which the inclusion of soil moisture in calibration is known to improve (Koren et al. 2008; Wooldridge et al. 2003).

B.4.4 Contribution of Multiple Data Sources

The use of a single radar data source to derive surface soil moisture, even at multiple incidence angles and polarizations, bears many inherent limitations. High revisit time, speckle induced error, surface roughness and vegetation effects all constrain the effectiveness of retrieving soil moisture from a single SAR satellite. In order to alleviate these constraints, a host of methods have been demonstrated in which one or more of the limitations of SAR are mitigated through the amalgamation of SAR soil moisture products with complimentary data products. The Soil Moisture Active Passive (SMAP) mission is one such example where the relatively noisy high resolution soil moisture will be retrieved with active radar to capture spatial dynamics and combined with more accurate passive radar (Entekhabi et al. 2010). In the SMAP example, SAR soil moisture is filtered to reduce speckle and used to determine the distribution of soil moisture within the footprint of the radiometer retrieval (Narayan et al. 2006; Piles et al. 2009). The SMAP approach establishes a linear relationship between simultaneous active and passive soil moisture retrievals over the same site, which is used to disaggregate the radiometer soil moisture signal (Entekhabi et al. 2010; Narayan et al. 2006; Piles et al. 2009). Zhan et al. (2006) proposed the use of a Bayesian merging method, similar to the Kalman Filter, to combine radar and radiometer retrievals. The Bayesian method uses the error covariance of SAR and radiometer soil moisture to calculate optimized soil moisture. This procedure worked best when speckle was removed, thus compromising spatial resolution (Zhan et al. 2006). In both cases the final result was a soil moisture product with lower RMS error than inversion of either SAR or radiometer individually, with a slight compromise in spatial resolution compared to SAR (Entekhabi et al. 2010; Zhan et al. 2006).

In contrast to methods previously discussed which combine SAR and radiometer data at the soil moisture retrieval stage, Parada and Liang (2008) combine active and passive soil moisture retrievals at the application stage. The Multi-scale Kalman Filter (MKF) allows the

assimilation of data sources at different spatial resolution into a hydrological model (Parada and Liang, 2004). With the assimilation of both active and passive data, spatial features which were lost at the coarse 25 km resolution were recovered, improving model performance (Parada and Liang, 2008). Similarly, the assimilation of coarse soil moisture data with temporal ratios based on the rank stability of SAR return also lead to significant improvement and information recovery of about 40% (Parada and Liang, 2008).

A major shortcoming of current SAR sensors is the issue of low temporal resolution, which limits the applicability of many simple retrieval methods such as change detection. The launch of upcoming SAR constellations will mitigate this issue somewhat, but in the meantime data from currently available sensors with similar configurations may also be able to be used together through distribution fitting. Demonstrated with 25 km data from the active ASCAT and passive AMSR-E, the distribution fitting techniques uses interpolation to match the cumulative distribution function (CDF) of the remotely sensed data sources to that of a common reference data set, usually a model (Liu et al. 2011). Distribution fitting requires complimentary data sets which have a correlation of greater than 0.65 and matches the range of the satellite data to that of the model while maintaining the relative dynamics of the satellite soil moisture (Liu et al. 2011). Despite the shift of absolute volumetric soil moisture as a result of the CDF fitting, assimilation of fitted soil moisture data can still increase model skill by R greater than 0.1 compared to open loop simulations (Draper et al. 2012) because the distribution of soil moisture is more important for the determination of surface processes than absolute values (Brocca et al. 2010).

B.5 Hydrological Modelling Potential of SAR Soil Moisture

B.5.1 Overview

Soil moisture information derived from synthetic aperture radar has a high potential to benefit hydrological and meteorological applications. These benefits will be realized by implicitly accounting for heterogeneity at what is currently sub-pixel resolution for many LSM applications as future technological advances allow for hyper-resolution modelling (Wood et al. 2011). The assimilation of soil moisture can compensate for noisy model forcing data (Loew et al. 2009), the misrepresentation of model processes such as latent energy flux (Mohr et al. 2003) and for poor model parameterization. Also, the uncertainty introduced by global climate changes brings into

question the temporal stability of model parameters (Peel and Blöschl, 2011) which can be better calibrated by the incorporation of soil moisture data (Koren et al. 2008; Wooldridge et al. 2003). While not well established in the literature, the potential of the global coverage of SAR satellites also has prospective benefits for the parameterization of ungauged basins.

The benefits of soil moisture data have already been proven for many of the above applications using *in situ* measurements with poor spatial representativeness (Aubert et al. 2003; Loew et al. 2009; Koren et al. 2008,) or with coarse spatial resolution (Bolten et al. 2010; Crow and van den Berg; 2010; Draper et al. 2012; Reichle et al. 2008; Walker and Houser, 2001), although some examples do exist as to the benefits of high resolution SAR soil moisture (Pauwels et al. 2002). Perhaps the two most compelling reasons for the lack of examples of SAR retrieved soil moisture use in operational hydrology are the lack of data availability and the relatively poor performance of most soil moisture retrievals. However, performance metrics in soil moisture retrieval studies are often based on retrieval error with respect to soil moisture measurements aggregated beyond the field scale, which is a natural comparison, but does not account for the spatial heterogeneity of soil moisture or the benefits at the application level. In many cases the volumetric accuracy of soil moisture are less important than the spatial heterogeneity (Bronstert and Bardossy, 1999; Pan and Wood, 2010) as the spatial patterns allow for the better identification of important patterns such as runoff source areas. In this way, the benefit of SAR soil moisture lies less in the retrieval of soil moisture itself, but in the identification of spatial patterns across a catchment (Parada and Liang, 2008; Wagner et al. 2008). Even so, noisy high resolution soil moisture data still provides noted benefits at application time (Reichle et al. 2008). With these considerations in mind, the following discussion will seek to identify the key issues and necessary compromises in soil moisture retrieval focusing primarily on the implications of soil moisture retrieval techniques discussed in the first section of this review to the applications discussed in the second section.

B.5.2 Key Issues

The separation of the influences of vegetation and surface roughness from that of soil moisture in backscatter signals remains a primary consideration. In terms of surface roughness, this has traditionally been accomplished through the use of under-representative *in situ* surface roughness measurements, which limits the established relationship to selected sampling areas. Fortunately,

recent studies have demonstrated that roughness can be effectively parameterized solely from SAR. Considering surface roughness as a model tuning parameter that is sensor and scene dependent (Lievens et al. 2011; Su et al. 1997) allows for inversion making few *a priori* assumptions about surface conditions. Since parameters may vary between SAR acquisitions, Lievens et al. (2011) have suggested a technique using normalized backscatter to derive effective roughness. The added benefit of such procedures that rely on few ground based data sets is the representation of the spatial distribution of surface roughness and soil moisture, which is of greater benefit to hydrological applications. Unfortunately, this process is based on the assumption that the soil moisture and its distribution do not vary between multi-angular acquisitions. In order to decouple the backscattered vegetation signal from that of soil moisture many recent publications have coupled the WCM with the IEM (or AIEM) (Álvarez-Mozos et al. 2006; Joseph et al. 2008; Lievens and Verhoest, 2011; Lievens et al. 2011; Wang et al. 2011; and others). The WCM approach offers the advantages of simplicity and relatively small computation requirements compared to more physically representative models. It is these advantages that currently make the WCM the best candidate in a potential hydrological application, especially when parameterizing vegetation from multi-polarized data (Gherboudj et al. 2011; Kasischke et al. 2011; Lin et al. 2009; Shi et al. 2004). The aforementioned modelling approach reduces the number of unknown parameters prior to inversion, alleviating some of the difficulty associated with the under-determined problem.

The simplest method of removing the surface roughness (and some vegetation) influences is through the use of change detection techniques. The limiting factor for the use of change detection is the underlying assumption that neither roughness nor vegetation has significantly changed between acquisitions. While current sensors can use multi-angular data to make acquisitions of the same area in a relatively short times span through adjustable viewing geometry, both surface roughness and vegetation influences are known to increase with incidence angle changing the backscatter behaviour (Baghdadi et al. 2002a; Mattia et al. 2003b; Zribi & Deschambre, 2002). Therefore, this method can only be truly implemented in an operational sense when the same viewing geometry is used with short revisit times. The key shortcoming of the change detection approach is that the change in backscatter is often related to uniformly dry conditions [Eq B-10] (Moran et al. 2000; Pathe et al. 2009; Zribi et al. 2011) or to known surface soil moisture (Narayan et al. 2006; Wickel et al. 2001). These limitations can be

ignored by using the change detection method not for soil moisture retrieval but the establishment of temporal ratios using the assumption of rank stability (Parada and Liang, 2008; Wagner et al. 2008).

In order to properly remove the vegetation and surface roughness effects using model inversion and change detection requires the acquisition of satellite data at a temporal resolution higher than is currently available with operational SAR satellites (35 days ASAR and 24 days RADARSAT-2). While the geographic location can be imaged more frequently because of variable imaging geometry, the change in incidence angle will impact the perceived roughness and optical depth (Aubert et al. 2011; Balenzano et al. 2011; Blumberg et al. 2000; Ulaby et al. 1986; Ulaby and Batlivala, 1976). With this in mind, multiple agencies are planning the launch of SAR constellations to improve the scene revisit time (Attema et al. 2009; Du et al. 2010; Flett et al. 2009; Snoeij et al. 2010; Torres et al. 2012). Soil moisture retrieval may benefit the greatest from L band sensors such as ALOS-2 PALSAR-2 and the SAOCOM constellation. However, from a soil moisture retrieval perspective the C band multi-configuration constellations Sentinel 1 (Torres et al. 2012) and RADARSAT Constellation (Flett et al. 2009) will significantly reduce the error in the assumption of no change in surface roughness or vegetation between acquisitions for change detection methods and may even permit near-simultaneous multi-angular acquisitions for deriving surface roughness and vegetation parameters. The validity of the zero change assumptions are important, as both surface roughness and vegetation change over time (Callens et al. 2006), albeit by different rates and magnitudes. This is particularly important when using multi-angular observations to determine surface roughness, as the difference in backscatter between the two angles is considered to be purely the result of surface roughness as soil moisture at different incidence angles will produce the same backscatter contribution (Zribi and Deschambre, 2002). By coordinating two satellites Sentinel-1 can achieve an exact revisit time of 6 days (Torres et al. 2012) and the use of three satellites will allow the RADARSAT Constellation an exact revisit time of 4 days (Flett et al. 2009). Thus, while not optimum in terms of wavelength, the proposed SAR constellations can be expected to reduce the uncertainty in current soil moisture retrieval model parameterizations, with greater benefits possible from the L Band SAOCOM Constellation.

In terms of hydrological applications this sparse revisit time limits the availability of soil moisture data for techniques such as assimilation or the initialization of an operational flood

forecast model, where data may not be available when it is needed in emergency situations. Despite the findings that spatial data availability is more important than temporal availability (Pan and Wood, 2010), current sensors only allow retrieval of soil moisture at weekly to bi-weekly resolution using multi-angular data which still may not be adequate for hydrological applications. Distribution matching (Liu et al. 2011) is a possible solution to this problem allowing complimentary data from ASAR and RADARSAT to be combined thereby increasing the temporal resolution and possibly producing soil moisture retrievals from the forthcoming constellations that allow for the assimilation of near daily soil moisture.

B.5.3 Compromises in Soil Moisture Retrieval

With regards to retrieval of soil moisture from SAR many conciliations are necessary, some of which are imposed and others at the discretion of the researcher. Issues such as the bandwidth of satellites, particularly the use of C instead of L band, radar design and the fact that many SAR satellites are primarily tasked to duties such as ship and sea ice monitoring (Flett et al. 2009; Torres et al. 2012) limits the usefulness of SAR for hydrological applications. These considerations combined with the low operational cycle (~30% per orbit) of most SAR limits the availability of data for hydrological purposes, increasing the importance of data sets such as WideSwath and ScanSAR for ASAR and RADARSAT respectively. The ASAR also had a Global Monitoring (GM) mode, which operated as a background state continually collecting data, with a resolution of 1000 m (Doubkova et al. 2012; Pathe et al. 2009). After filtering and spatial averaging ASAR GM data had a useful resolution of 3-10 km (Pathe et al. 2009), similar to SMAP, but adequate for many hydrological and particularly meteorological applications.

Those compromises which researchers have control over are primarily related to the retrieval of soil moisture itself, and the models chosen for the retrieval. This largely results as a balance of model physical representativeness, computational effort and ancillary data for parameterization. The change detection approach is arguably the most parsimonious retrieval method with comparable results to more physically representative means, but is severely limited by the availability of data. Numerical methods make few assumptions and are highly representative, but the computational cost limits real world applicability (Lawrence et al. 2011; Onier et al. 2011; Rabus et al. 2010), whereas the IEM is both physically representative without requiring an arduous computation effort. Similarly, the removal of vegetation is important and

can be carried out through extensive parameterization of a physically representative model (Dobson et al. 1992) or as is often the case, is simplified to representation through the WCM (Álvarez-Mozos et al. 2006; Bindlish and Barros, 2001; Joseph et al. 2010; Zribi et al. 2007; Zribi et al. 2011; and others).

The largest compromise in terms of soil moisture retrieval from SAR, and arguably the most important from a hydrological perspective, is the result of balancing the accuracy of the retrieval algorithm, with the representation of soil moisture heterogeneity. The filtering of speckle and averaging of retrieved soil moisture values results in a significant reduction in spatial resolution (Thoma et al. 2008), thereby increasing the variability of the soil moisture represented by the retrieval (Brocca et al. 2010; Famiglietti et al. 2008). The cost in spatial resolution may be acceptable from the standpoint of soil moisture retrieval, but quickly reduces the hydrological applicability of the derived data. This is especially true considering the importance of spatial heterogeneity as compared to accuracy (Pan and Wood, 2010). It appears that combined spatial filtering and averaging beyond the sub-catchment scale (< 10 km) would cease to be beneficial for hydrological applications as further filtering would produce a retrieval result with an order of magnitude similar to that of more accurate passive radar. Hopefully, future research will help to address this balance and provide guidelines to determine at what level of error are the benefits of spatial heterogeneity compromised as filtering decreases the spatial resolution.

The compromise between accuracy and scale is also manifested when merging SAR soil moisture with radiometer soil moisture (Entekhabi et al. 2010; Narayan et al. 2006; Zhan et al. 2006). This discussion of scale and accuracy then leads to the question which should be fundamental to soil moisture retrieval, “what kind of information is actually required from SAR to add value to the hydrological application in question?” Many suggest that relative spatial dynamics are potentially more important (Pan and Wood, 2010; Parada and Liang, 2008; Wagner et al. 2008) allowing the magnitude of the soil moisture state to be determined from radiometers or model state (Draper et al. 2012; Liu et al. 2011; Parada and Liang, 2008), which is itself a compromise to increase the temporal resolution (Draper et al. 2012; Liu et al. 2012). In terms of SMAP using SAR to disaggregate the radiometer only represents a relatively small compromise where 3 km SAR is merged into a 10 km soil moisture data product (Entekhabi et al. 2010). The slightly greater than three times increase in scale with SMAP is far less of a spatial compromise than the 25-160 times increase suggested to directly use SAR by Thoma et al. (2008).

B.6 Conclusion

In order to better understand the impacts of the hydrological cycle on climate, water resources, the spread of pollution and a host of other environmental phenomena, further information are required on the distribution and amount of soil moisture. Because the traditional methods of *in situ* measurement are too costly, synthetic aperture radar (SAR) from spaceborne platforms allows for the monitoring of this key state variable in a distributed manner at relatively high resolutions. Consideration of soil moisture dynamics at the sub-watershed scale is hydrologically important as it regulates runoff response (Minet et al. 2011; Zehe and Blöschl, 2004) and influences surface energy flux (Mohr et al. 2000; Mohr et al. 2003). Through assimilation, soil moisture data reduce the uncertainty associated with meteorological inputs (Bolten et al. 2010; Loew et al. 2009; Merlin et al. 2006) and compensate for under-representative model physics (Mohr et al. 2000), thereby increasing the skill of the final simulation or forecast product (Draper et al. 2012; Kumar et al. 2008; Pauwels et al. 2002; Reichle et al. 2008; Walker and Houser, 2001). Soil moisture information can also aid in the parameterization and calibration of hydrological models, particularly in watersheds where flow is ephemeral or highly variable (Wooldridge et al. 2003) or are impacted by the effects of climate change (Merz et al. 2011; Peel and Blöschl, 2011).

While the benefits of soil moisture data is evident, retrieval of surface soil moisture remains difficult, confounded by the competing influences of surface roughness, vegetation and speckle each of which introduce a unique set of influences. Change detection approaches can be effective and simple methods for estimating soil moisture by contrasting the changes in a dry image to that of subsequent acquisitions and assuming no other changes besides soil moisture have occurred. As an alternative, inversion of semi-empirical and theoretical models provides the ability to determine soil moisture considering a wide variety of conditions at the expense of simplicity. The Integral Equation Model (IEM) provides the most accurate representation of radar backscatter from a bare surface at a moderate computational cost, but requires multi-angular data in order to adequately remove the influence of surface roughness across the entire scene (Wang et al. 2011). Unfortunately, the applicability of both of these retrieval methods for hydrological applications suffer from the poor temporal resolution of current sensors as well as

the need to reduce spatial resolution to increase retrieval accuracy (Baghdadi and Zribi, 2006; Lievens and Verhoest, 2011; Paloscia, 2002; Thoma et al. 2008; Verhoest et al. 2007b; and others).

Unquestionably, lower error retrievals result in higher assimilation skill (Reichle et al. 2008), yet the absolute skill of the retrieval is less important to hydrological applications than spatial or temporal availability of soil moisture information (Pan and Wood, 2010). The importance of the soil moisture spatial distribution indicates that, from the perspective of the end use of the data, the need to compromise in spatial resolution to attain an acceptable level of accuracy may not be as extensive as suggested (Thoma et al. 2008) and therefore requires more study. Indeed the sensitivity of SAR to soil moisture may be better exploited through the establishment of the spatial distribution pattern, allowing the soil moisture quantity to be determined by other methods (Parada and Liang, 2008). Regarding retrieval skill as being of secondary importance the major barrier remains the temporal availability of data. Distribution fitting (Liu et al. 2011) and the near future launch of SAR constellations will mitigate the problem of availability, leaving the hydrological community to study and realize the potential benefits offered by those advances to produce distributed surface soil moisture values.

B.7 References

Alavi, N., Berg, A., Warland, J.S., Parkin, G., Verseghy, D., Bartlett, P., 2010. Evaluating the impact of assimilating soil moisture variability data on latent heat flux estimation in a land surface model. Can. Wat. Res. J. 35(2), 157-172.

Altese, E., Bolognani, O., Mancini, M., Troch, P.A., 1996. Retrieving soil moisture over bare soil from ERS 1 synthetic aperture radar data: Sensitivity analysis based on a theoretical surface scattering model and field data. Water Resour. Res. 32(3), 653-661.

Álvarez-Mozos, J., Casali, J., González-Audícana, M., Verhoest, N.E.C., 2006. Assessment of the operational applicability of RADARSAT-1 data for surface soil moisture estimation. IEEE Trans.Geosci.Remote Sens. 44(4), 913-924.

Álvarez-Mozos, J., Casali, J., González-Audícana, M., Verhoest, N.E.C., 2005. Correlation between ground measured soil moisture and RADARSAT-1 derived backscattering coefficient over an agricultural catchment of Navarre (North of Spain). Biosys. Eng. 92(1) 119-133.

Álvarez-Mozos, J., González-Audícana, M., Casalí, J., 2007. Evaluation of empirical and semi-empirical backscattering models for surface soil moisture estimation. Can. J. Remote Sens. 33(3), 176-188.

Álvarez-Mozos, J., González-Audícana, M., Casalí, J., Larranaga, A., 2008. Effective versus measured correlation length for radar-based surface soil moisture retrieval. Int. J. Remote Sens. 29(17-18), 5397-5408.

Álvarez-Mozos, J., Verhoest, N.E.C., Larranaga, A., Casalí, J., González-Audícana, M., 2009. Influence of surface roughness spatial variability and temporal dynamics on the retrieval of soil moisture from SAR observations. Sensors. 9(1), 463-489.

Anderson, K., Croft, H., 2009. Remote sensing of soil surface properties. Prog. Phys. Geogr. 33(4), 457-473.

Anguela, T.P., Zribi, M., Baghdadi, N., Loumagne, C., 2010. Analysis of local variation of soil surface parameters with TerraSAR-X radar data over bare agricultural fields. IEEE Trans. Geosci. Remote Sens. 48(2), 874-881.

Aronica, G., Candela, A., 2004. A regional methodology for deriving flood frequency curves (FFC) in partially gauged catchments with uncertain knowledge of soil moisture conditions. Proceedings of the International Environmental Modelling and Software Society Conference, Osnabrück, Germany 14-17 June 2004, pp. 1147-1152.

Attema, E.P.W., Ulaby, F.T., 1978. Vegetation modeled as a water cloud. Radio Sci. 13(2), 357-364.

Attema, E., Davidson, M., Snoeij, P., Rommen, B., Floury, N., 2009. Sentinel-1 mission overview. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium July 2009, Vol. 1, pp. 36-39.

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

Aubert, M., Baghdadi, N., Zribi, M., Douaoui, A., Loumagne, C., Baup, F., El Hajj, M., Garrigues, S. 2011. Analysis of TerraSAR-X data sensitivity to bare soil moisture, roughness, composition and soil crust. Remote Sens.Environ. 115(8), 1801-1810.

Baghdadi, N., Aubert, M., Cerdan, O., Franchisteguy, L., Viel, C., Martin, E., Zribi, M., Desprats, J.F., 2007. Operational mapping of soil moisture using sythetic aperture radar data: Application to the Touch Basin (France). Sensors. 7(10), 2458-2483.

Baghdadi, N., Chaaya, J.A., Zribi, M., 2011. Semiempirical calibration of the Integral Equation Model for SAR data in C-Band and cross polarization using radar images and field measurements. IEEE Geosci. Remote Sens. Lett. 8(1), 14-18.

Baghdadi, N., Gaultier, S., King, C., 2002a. Retrieving surface roughness and soil moisture from synthetic aperture radar (SAR) data using neural networks. Can. J. Remote Sens. 28(5), 701-711.

Baghdadi, N., Gherboudj, I., Zribi, M., Sahebi, M., King, C., Bonn, F., 2004. Semi-empirical calibration of the IEM backscattering model using radar images and moisture and roughness field measurements. Int. J. Remote Sens. 25(18), 3593-3623.

Baghdadi. N., Holah, N., Zribi, M., 2006. Calibration of the Integral Equation Model for SAR data in C-band and HH and VV polarizations. Int. J. Remote Sens. 27(4), 805-816.

Baghdadi, N., King, C., Bonnifait, L., 2002b. An empirical calibration of the Integral Equation Model based on SAR data and soil parameters measurements. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium and 24th Canadian Symposium on Remote Sensing, Vols I-VI, pp. 2646-2650.

Baghdadi, N., King, C., Bourguignon, A., Remond, A., 2002c. Potential of ERS and RADARSAT data for surface roughness monitoring over bare agricultural fields: Application to catchments in Northern France. Int. J. Remote Sens. 23(17), 3427-3442.

Baghdadi, N., Cerdan, O., Zribi, M., Auzet, V., Darboux, F., El Hajj, M., Kheir, R.B., 2008. Operational performance of current synthetic aperture radar sensors in mapping soil surface characteristics in agricultural environments: application to hydrological and erosion modelling. Hydrol. Process. 22(1), 9-20.

Baghdadi, N., Zribi, M., 2006. Evaluation of radar backscatter models IEM, OH and Dubois using experimental observations. Int. J. Remote Sens. 27(18), 3831-3852.

Balenzano, A., Mattia, F., Satalino, G., Davidson, M.W.J., 2011. Dense temporal series of C- and L-band SAR data for soil moisture retrieval over agricultural crops. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 4(2), 439-450.

Barber, M., Grings, F., Perna, P., Piscitelli, M., Maas, M., Bruscantini, C., Jacobo-Berlles, J., Karszenbaum, H. 2012. Speckle noise and soil heterogeneities as error sources in a Bayesian soil moisture retrieval scheme for SAR data. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 5(3), 942-951.

Baup, F., Mougin, E., de Rosnay, P., Timouk, F., Chenerie, I., 2007. Surface soil moisture estimation over the AMMA Sahelian site in Mali using ENVISAT/ASAR data. Remote Sens. Environ. 109(4), 473-481.

Bindlish, R., Barros, A.P., 2001. Parameterization of vegetation backscatter in radar-based, soil moisture estimation. Remote Sens. Environ. 76(1), 130-137.

Bindlish, R., Barros, A.P., 2000. Multifrequency soil moisture inversion from SAR measurements with the use of IEM. Remote Sens. Environ. 71(1), 67-88.

Bindlish, R., Jackson, T., Sun, R., Cosh, M., Yueh, S., Dinardo, S., 2009. Combined passive and active microwave observations of soil moisture during CLASIC. IEEE Geosci. Remote Sens. Lett. 6(4), 644-648.

Bisselink, B., van Meijgaard, E., Dolman, A.J., de Jeu, R.A.M., 2011. Initializing a regional climate model with satellite-derived soil moisture. J. Geophys. Res.-Atmos. 116 D02121.

Blöschl, G., Sivapalan, M., 1995. Scale issues in hydrological modeling: A Review. Hydrol. Process. 9(3-4), 251-290.

Blumberg, D.G., Freilikher, V., Lyalko, I.V., Vulfson, L.D., Kotlyar, A.L., Shevchenko, V.N., Ryabokonenko, A.D., 2000. Soil moisture (water-content) assessment by an airborne scatterometer: The Chernobyl disaster area and the Negev desert. Remote Sens. Environ. 71(3), 309-319.

Boisvert, J.B., Gwyn, Q.H.J., Chanzy, A., Major, D.J., Brisco, B., Brown, R.J., 1997. Effect of surface soil moisture gradients on modelling radar backscattering from bare fields. Int. J. Remote Sens. 18(1), 153-170.

Bolten, J.D., Crow, W.T., Zhan, X., Jackson, T.J., Reynolds, C.A., 2010. Evaluating the utility of remotely sensed soil moisture retrievals for operational agricultural drought monitoring. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 3(1), 57-66.

Bolton, J.D., Lakshmi, V., Njoku, E.G., 2003. Soil moisture retrieval using the passive/active L- and S-band radar/radiometer. IEEE Trans. Geosci. Remote Sens. 41(12), 2792-2801.

Bosch, D.D., Lakshmi, V., Jackson, T.J., Choi, M., Jacobs, J.M., 2006. Large scale measurements of soil moisture for validation of remotely sensed data: Georgia soil moisture experiment of 2003. J. Hydrol. 323(1-4), 120-137.

Bourgeau-Chavez, L.L., Kasischke, E.S., Riordan, K., Brunzell, S., Nolan, M., Hyer, E., Slawski, J., Medvecz, M., Walters, T., Ames, S., 2007. Remote monitoring of spatial and temporal surface soil moisture in fire disturbed boreal forest ecosystems with ERS SAR imagery. Int. J. Remote Sens. 28(10), 2133-2162.

Bracaglia, M., Ferrazzoli, P., Guerriero, L., 1995. A fully polarimetric multiple scattering model for crops. Remote Sens. Environ. 54(3), 170-179.

Brocca, L., Melone, F., Moramarco, T., Morbidelli, R., 2010. Spatial-temporal variability of soil moisture and its estimation across scales. Water Resour. Res. 46 W02516.

Brocca, L., Melone, F., Moramarco, T., Singh, V.P., 2009. Assimilation of observed soil moisture data in storm rainfall-runoff modeling. J. Hydrol. Eng. 14(2), 153-165.

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

Bronstert, A., Bardossy, A., 1999. The role of spatial variability of soil moisture for modelling surface runoff generation at the small catchment scale. Hydrol. Earth Sys. Sci. 3(4), 505-516.

Brown, S.C.M., Quegan, S., Morrison, K., Bennett, J.C., Cookmartin, G., 2003. High-resolution measurements of scattering in wheat canopies - Implications for crop parameter retrieval. IEEE Trans. Geosci. Remote Sens. 41(7), 1602-1610.

Bryant, R., Moran, M.S., Thoma, D.P., Collins, C.D.H., Skirvin, S., Rahman, M., Slocum, K., Starks, P., Bosch, D., González-Dugo, M.P., 2007. Measuring surface roughness height to parameterize radar backscatter models for retrieval of surface soil moisture. IEEE Geosci. Remote Sens. Lett. 4(1), 137-141.

Callens, M., Verhoest, N.E.C., Davidson, M.W.J., 2006. Parameterization of tillage-induced single-scale soil roughness from 4-m profiles. IEEE Trans. Geosci. Remote Sens. 44(4), 878-888.

Calvet, J.C., Noilhan, J., 2000. From near-surface to root-zone soil moisture using year-round data. J. Hydrometeorol. 1(5), 393-411.

Chen, K.S., Wu, T.D., Tsang, L., Li, Q., Shi, J.C., Fung, A.K., 2003. Emission of rough surfaces calculated by the integral equation method with comparison to three-dimensional moment method Simulations. IEEE Trans. Geosci. Remote Sens. 41(1), 90-101.

Chen, K.S., Wu, T.D., Tsay, M.K., Fung, A.K., 2000. A note on the multiple scattering in an IEM model. IEEE Trans. Geosci. Remote Sens. 38(1), 249-256.

Choi, M., Jacobs, J.M., 2007. Soil moisture variability of root zone profiles within SMEX02 remote sensing footprints. Adv. Water Resour. 30(4), 883-896.

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

Crow, W.T., van den Berg, M.J., 2010. An improved approach for estimating observation and model error parameters in soil moisture data assimilation. Water Resour. Res. 46 W12519.

Crow, W.T., Wood, E.F., 2002a. The value of coarse-scale soil moisture observations for regional surface energy balance modeling. J. Hydrometeorol. 3(4), 467-482.

Crow, W.T., Wood, E.F., 2002b. Impact of soil moisture aggregation on surface energy flux prediction during SGP'97. Geophys. Res. Lett. 29(1), 1008-1012.

Crow, W.T., Wagner, W., Naeimi, V., 2010. The impact of radar incidence angle on soil-moisture-retrieval skill. IEEE Geosci. Remote Sens. Lett. 7(3), 501-505.

Dall'Amico, J.T., Schlenz, F., Loew, A., Mauser, W., 2010. SMOS soil moisture validation: Status at the Upper Danube cal/val site eight months after launch. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium 2008, pp.3801-3804.

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

Das, N.N., Mohanty, B.P., 2008. Temporal dynamics of PSR-based soil moisture across spatial scales in an agricultural landscape during SMEX02: A wavelet approach. Remote Sens. Environ. 112(2), 522-534.

Davidson, M.W.J., Le Toan, T., Mattia, F., Satalino, G., Manninen, T., Borgeaud, M., 2000. On the characterization of agricultural soil roughness for radar remote sensing studies. IEEE Trans. Geosci. Remote Sens. 38(2), 630-640.

Davidson, M.W.J., Le Toan, T., Mattia, F., Satalino, G., Verhoest, N.E.C., Borgeaud, M., 2001. Improving soil moisture retrieval by incorporating a priori information on roughness parameters. Proceedings of the IEEE International Symposium on Geoscience and Remote Sensing, Jul 9-13, 2001; pp.1306-1308.

Davidson, M.W.J., Mattia, F., Satalino, G., Verhoest, N.E.C., Le Toan, T., Borgeaud, M., Louis, J.M.B., Attema, E., 2003. Joint statistical properties of RMS height and correlation length derived from multisite 1-m roughness measurements. IEEE Trans. Geosci. Remote Sens. 41(7), 1651-1658.

De Keyser, E., Vernieuwe, H., Álvarez-Mozos, J., De Baets, B., Verhoest, N.E.C., 2012. Assessment of SAR-retrieved soil moisture uncertainty induced by uncertainty on modeled soil surface roughness. Int. J. App. Earth Obs. and Geoinf. 18, 176-192.

De Roo, R.D., Du, Y., Ulaby, F.T., Dobson, M.C., 2001. A semi-empirical backscattering model at L-band and C-band for a soybean canopy with soil moisture inversion. IEEE Trans. Geosci. Remote Sens. 39(4), 864-872.

Della Vecchia, A., Ferrazzoli, P., Guerriero, L., Ninivaggi, L., Strozzi, T., Wegmueller, U., 2008. Observing and modeling multifrequency scattering of maize during the whole growth cycle. IEEE Trans. Geosci. Remote Sens. 46(11), 3709-3718.

Dobson, M.C., Pierce, L., Sarabandi, K., Ulaby, F.T., Sharik, T., 1992. Preliminary-Analysis of ERS-1 SAR for forest ecosystem studies. IEEE Trans. Geosci. Remote Sens. 30(2), 203-211.

Dobson, M.C., Ulaby, F.T., 1986. Preliminary evaluation of the SIR-B response to soil moisture, surface roughness, and crop canopy cover. IEEE Trans. Geosci. Remote Sens. 24(4), 517-526.

Dobson, M.C., Ulaby, F.T., 1981. Microwave backscatter dependence on surface roughness, soil moisture, and soil texture .3: Soil tension. IEEE Trans. Geosci. Remote Sens. 19(1), 51-61.

Dobson, M.C., Ulaby, F.T., Hallikainen, M.T., Elrayes, M.A., 1985. Microwave dielectric behavior of wet soil .2: Dielectric mixing models. IEEE Trans. Geosci. Remote Sens. 23(1), 35-46.

Doubkova, M., van Dijk, A.I.J.M., Sabel, D., Wagner, W., Blöschl, G., 2012. Evaluation of the predicted error of the soil moisture retrieval from C-band SAR by comparison against modelled soil moisture estimates over Australia. Remote Sens. Environ. 120, 188-196.

Draper, C., Mahfouf, J., Calvet, J., Martin, E., Wagner, W., 2011. Assimilation of ASCAT near-surface soil moisture into the SIM hydrological model over France. Hydrol. Earth Sys. Sci. 15(12), 3829-3841.

Draper, C.S., Reichle, R.H., De Lannoy, G.J.M., Liu, Q., 2012. Assimilation of passive and active microwave soil moisture retrievals. Geophys. Res. Lett. 39 L04401.

Du, J., Shi, J., Sun, R., 2010. The development of HJ SAR soil moisture retrieval algorithm. Int. J. Remote Sens. 31(14), 3691-3705.

Dubois, P.C., van Zyl, J., Engman, T., 1995a. Measuring soil moisture with imaging radars. IEEE Trans. Geosci. Remote Sens. 33(4), 915-926.

Dubois, P.C., van Zyl, J., Engman, T., 1995b. Corrections to “Measuring soil moisture with imaging radars” IEEE Trans. Geosci. Remote Sens. 33(6), 1340-1340.

D'Urso, G., Minacapilli, M., 2006. A semi-empirical approach for surface soil water content estimation from radar data without a-priori information on surface roughness. J. Hydrol. 321(1-4), 297-310.

Entekhabi, D., Njoku, E.G., O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J., Goodman, S., Jackson, T., Johnson, J., Kimball, J., Piepmeier, J., Koster, R.D., Martin, N., McDonald, K.C. Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., van Zyl, J., 2010. The Soil Moisture Active Passive (SMAP) Mission. Proc. IEEE 98(5), 704-716.

Entin, J.K., Robock, A., Vinnikov, K.Y., Hollinger, S.E., Liu, S.X., Namkhai, A., 2000. Temporal and spatial scales of observed soil moisture variations in the extratropics. J. Geophys. Res. –Atmospheres. 105(D9), 11865-11877.

Famiglietti, J.S., Devereaux, J.A., Laymon, C.A., Tsegaye, T., Houser, P.R., Jackson, T.J., Graham, S.T., Rodell, M., van Oevelen, P.J., 1999. Ground-based investigation of soil moisture variability within remote sensing footprints during the Southern Great Plains 1997 (SGP97) Hydrology Experiment. Water Resour. Res. 35(6), 1839-1851.

Famiglietti, J.S., Rudnicki, J.W., Rodell, M., 1998. Variability in surface moisture content along a hillslope transect: Rattlesnake Hill, Texas. J. Hydrol. 210(1-4), 259-281.

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

Famiglietti, J.S., Ryu, D., Berg, A.A., Rodell, M., Jackson, T.J., 2008. Field observations of soil moisture variability across scales. Water Resour. Res. 44(1), W01423.

Fernández-Gálvez, J. 2008. Errors in soil moisture content estimates induced by uncertainties in the effective soil dielectric constant. Int. J. Remote Sens. 29(11), 3317-3323.

Flett, D., Crevier, Y., Girard, R., 2009. The RADARSAT Constellation Mission: Meeting the Government of Canada's needs and requirements. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Vol 2. Pp. 910-913.

Fung, A.K., 1994. Microwave Scattering and Emission Models and Their Applications. Norwood, MA: Artech House Inc.

Fung, A.K., Chen, K.S., 2010. Microwave Scattering and Emission Models for Users. Norwood, MA: Artech House Inc.

Fung, A.K., Chen, K.S., 2004. An update on the IEM surface backscattering model. IEEE Geosci. Remote Sens. Lett. 1(2), 75-77.

Fung, A.K., Li, Z.Q., Chen, K.S., 1992. Backscattering from a Randomly Rough Dielectric Surface. IEEE Trans. Geosci. Remote Sens. 30(2), 356-369.

Fung, A.K., Tjuatja, S., 1993. Backscattering and emission signatures of randomly rough surfaces based on IEM. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium pp. 1006-1008.

Gherboudj, I., Magagi, R., Berg, A.A., Toth, B., 2011. Soil moisture retrieval over agricultural fields from multi-polarized and multi-angular RADARSAT-2 SAR data. Remote Sens. Environ. 115(1), 33-43.

Gherboudj, I., Magagi, R., Berg, A., Toth, B., 2009. Use of RADARSAT-2 images to develop a scaling method of soil moisture over an agricultural area. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Vol 3 pp. 510-514.

Glenn, N.F., Carr, J.R., 2004. Establishing a relationship between soil moisture and RADARSAT-1 SAR data obtained over the Great Basin, Nevada, USA. Can. J. Remote Sens. 30(2), 176-181.

Graham, A.J., Harris, R., 2003. Extracting biophysical parameters from remotely sensed radar data: A review of the water cloud model. Prog. Phys. Geogr. 27(2), 217-229.

Hallikainen, M.T., Ulaby, F.T., Dobson, M.C., Elrayes, M.A., Wu, L.K., 1985. Microwave dielectric behavior of wet soil .1: Empirical models and experimental observations. IEEE Trans. Geosci. Remote Sens. 23(1), 25-34.

Heathman, G.C., Larose, M., Cosh, M.H., 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.

Hoeben, R., Troch, P.A., 2000. Assimilation of active microwave observation data for soil moisture profile estimation. Water Resour. Res. 36(10), 2805-2819.

Hoeben, R., Troch, P.A., Su, Z.B., Mancini, M., Chen, K.S., 1997. Sensitivity of radar backscattering to soil surface parameters: A comparison between theoretical analysis and experimental evidence. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium pp.1368-1370.

Holah, N., Baghdadi, N., Zribi, M., Bruand, A., King, C., 2005. Potential of ASAR/ENVISAT for the characterization of soil surface parameters over bare agricultural fields. Remote Sens. Environ. 96(1), 78-86.

Hsieh, C.Y., Fung, A.K., Nesti, G., Sieber, A.J., Coppo, P., 1997. A further study of the IEM surface scattering model. IEEE Trans. Geosci. Remote Sens. 35(4), 901-909.

Huang, S., Tsang, L., Njoku, E.G., Chan, K.S., 2010. Backscattering coefficients, coherent reflectivities, and emissivities of randomly rough soil surfaces at L-Band for SMAP applications based on numerical solutions of Maxwell Equations in three-dimensional simulations. IEEE Trans. Geosci. Remote Sens. 48(6), 2557-2568.

Iodice, A., Natale, A., Riccio, D., 2011. Retrieval of soil surface parameters via a polarimetric two-scale model. IEEE Trans. Geosci. Remote Sens. 49(7), 2531-2547.
Hupet, F., Vanclooster, M., 2002. Intraseasonal dynamics of soil moisture variability within a small agricultural maize cropped field. J. Hydrol. 261(1-4), 86-101.

Ivanov, V.Y., Fatichi, S., Jenerette, G.D., Espeleta, J.F., Troch, P.A., Huxman, T.E., 2010. Hysteresis of soil moisture spatial heterogeneity and the "homogenizing" effect of vegetation. Water Resour. Res. 46, W09521.

Jackson, T.J., Chen, D.Y., Cosh, M., Li, F.Q., Anderson, M., Walthall, C., Doriaswamy, P., Hunt, E.R., 2004. Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans. Remote Sens. Environ. 92(4), 475-482.

Jackson, T.J., Le Vine, D.M., Hsu, A.Y., Oldak, A., Starks, P.J., Swift, C.T., Isham, J.D., Haken, M., 1999. Soil moisture mapping at regional scales using microwave radiometry: The Southern Great Plains Hydrology Experiment. IEEE Trans. Geosci. Remote Sens. 37(5), 2136-2151.

Jackson, T.J., McNairn, H., Weltz, M.A., Brisco, B., Brown, R., 1997. First order surface roughness correction of active microwave observations for estimating soil moisture. IEEE Trans. Geosci. Remote Sens. 35(4), 1065-1069.

Jackson, T.J., Schmugge, T.J., 1989. Passive microwave remote-sensing system for soil moisture: Some supporting research. IEEE Trans. Geosci. Remote Sens. 27(2), 225-235.

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

James, A.L., Roulet, N.T., 2009. Antecedent moisture conditions and catchment morphology as controls on spatial patterns of runoff generation in small forest catchments. J. Hydrol. 377(3-4), 351-366.

Joseph, A.T., van der Velde, R., O'Neill, P.E., Lang, R., Gish, T., 2010. Effects of corn on C- and L-band radar backscatter: A correction method for soil moisture retrieval. Remote Sens. Environ. 114(11), 2417-2430.

Joseph, A.T., van der Velde, R., O'Neill, P.E., Lang, R.H., Gish, T., 2008. Soil moisture retrieval during a corn growth cycle using L-band (1.6 GHz) radar observations. IEEE Trans. Geosci. Remote Sens. 46(8), 2365-2374.

Karam, M.A., Fung, A.K., Lang, R.H., Chauhan, N.S., 1992. A microwave scattering model for layered vegetation. IEEE Trans. Geosci. Remote Sens. 30(4), 767-784.

Kasischke, E.S., Tanase, M.A., Bourgeau-Chavez, L.L., Borr, M., 2011. Soil moisture limitations on monitoring boreal forest regrowth using spaceborne L-band SAR data. Remote Sens. Environ. 115(1), 227-232.

Kellndorfer, J.M., Pierce, L.E., Dobson, M.C., Ulaby, F.T., 1998. Toward consistent regional-to-global scale vegetation characterization using orbital SAR systems. IEEE Trans. Geosci. Remote Sens. 36(5), 1396-1411.

Kerr, Y.H., Wadlteufel, P., Wigneron, J.P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.J., Font, J., Reul, N., Gruhier, C., Juglea, S.E., Drinkwater, M.R., Hahne, A., Martín-Neira, M., Mecklenburg, S., 2010. The SMOS Mission: New tool for monitoring key elements of the global water cycle.; Proc. IEEE. 98 (5), 666-687.

Kim, Y., van Zyl, J.J., 2009. A time-series approach to estimate soil moisture using polarimetric radar data. IEEE Trans. Geosci. Remote Sens. 47(8), 2519-2527.

Komma, J., Blöschl, G., Reszler, C., 2008. Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting. J. Hydrol. 357(3-4), 228-242.

Kong, X., Dorling, S.R., 2008. Near-surface soil moisture retrieval from ASAR wide swath imagery using a principal component analysis. Int. J. Remote Sens. 29(10), 2925-2942.

Koren, V., Moreda, F., Smith, M., 2008. Use of soil moisture observations to improve parameter consistency in watershed calibration. Phys. Chem. Earth 33(17-18), 1068-1080.

Koster, R.D., Mahanama, S.P.P., Yamada, T.J., Balsamo, G., Berg, A.A., Boisserie, M., Dirmeyer, P.A., Doblas,-Reyes, F.J., Drewitt, G., Gordon, C.T., Guo, Z., Jeong, J.-H., Lawrence, D.M., Lee, W.-S., Li, Z., Luo, L., Malyshev, S., Merryfield, W.J., Seneviratne, S.I., Stanelle, T., van der Hurk, M.J.J.M., Vitart, F., Wood, E.F., 2010a. Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. Geophys. Res. Lett. 37, L02402.

Koster, R.D., Mahanama, S.P.P., Livneh, B., Lettenmaier, D.P., Reichle, R.H., 2010b. Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. Nature Geoscience 3(9), 613-616.

Kumar, S.V., Reichle, R.H., Peters-Lidard, C.D., Koster, R.D., Zhan, X., Crow, W.T., Eylander, J.B., Houser, P.R., 2008. A land surface data assimilation framework using the land information system: Description and applications. Adv. Water Resour. 31(11), 1419-1432.

Kurucu, Y., Sanli, F.B., Esetlili, M.T., Bolca, M., Goksel, C., 2009. Contribution of SAR images to determination of surface moisture on the Menemen Plain, Turkey. Int. J. Remote Sens. 30(7), 1805-1817.

Lakhankar, T., Ghedira, H., Khanbilvardi, R., 2006. Soil moisture retrieval from RADARSAT data: A Neuro-Fuzzy approach. 2006. Proceedings of the IEEE International Symposium on Geoscience and Remote Sensing; Jul 31-Aug 4, 2006; pp. 2328-2331.

Lakhankar, T., Ghedira, H., Temimi, M., Azar, A.E., Khanbilvardi, R., 2009. Effect of land cover heterogeneity on soil moisture retrieval using active microwave remote sensing data. Remote Sens. 1, 80-91.

Lawrence, H., Demontoux, F., Wigneron, J., Paillou, P., Wu, T., Kerr, Y.H., 2011. Evaluation of a numerical modeling approach based on the finite-element method for calculating the rough surface scattering and emission of a soil layer. IEEE Geosci. Remote Sens. Lett. 8(5), 953-957.

Le Hégarat-Mascle, S., Zribi, M., Alem, F., Weisse, A., Loumagne, C., 2002. Soil moisture estimation from ERS/SAR data: Toward an operational methodology. IEEE Trans. Geosci. Remote Sens. 40(12), 2647-2658.

Lee, J., Grunes, M., Kwok, R., 1994. Classification of multi-look polarimetric SAR imagery based on complex Wishart Distribution Int. J. Remote Sens. 15(11), 2299-2311.

Lievens, H., Verhoest, N.E.C., De Keyser, E., Vernieuwe, H., Matgen, P., Álvarez-Mozos J., De Baets, B., 2011. Effective roughness modelling as a tool for soil moisture retrieval from C-and L-band SAR. Hydrol. Earth Syst. Sci. 15(1), 151-162.

Lievens, H., Verhoest, N.E.C., 2012. Spatial and temporal soil moisture estimation from RADARSAT-2 imagery over Flevoland, The Netherlands. J. Hydrol. 456-457, 44-56.

Lievens, H., Verhoest, N.E.C., 2011. On the retrieval of soil moisture in wheat fields from L-Band SAR based on water cloud modeling, the IEM, and effective roughness parameters. IEEE Geosci. Remote Sens. Lett. 8(4), 740-744.

Lievens, H., Vernieuwe, H., Álvarez-Mozos, J., De Baets, B., Verhoest, N.E.C., 2009. Error in radar-derived soil moisture due to roughness parameterization: An analysis based on synthetical surface profiles. Sensors 9(2), 1067-1093.

Lin, H., Chen, J., Pei, Z., Zhang, S., Hu, X., 2009. Monitoring sugarcane growth using ENVISAT ASAR data. IEEE Trans. Geosci. Remote Sens. 47(8), 2572-2580.

Liu, W.Y., Chen, K.S., Tsay, M.K., Wu, T.D., 2003. A re-examination of the IEM model for microwave scattering from randomly rough boundary. J. Chin. Inst. Eng. 26(3), 271-277.

Liu, Y.Y., Parinussa, R.M., Dorigo, W.A., De Jeu, R.A.M., Wagner, W., van Dijk, A.I.J.M., McCabe, M.F., Evans, J.P., 2011. Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. Hydrol. Earth Sys. Sci. 15(2), 425-436.

Loew, A., Mauser, W., 2006. A semiempirical surface backscattering model for bare soil surfaces based on a generalized power law spectrum approach. IEEE Trans. Geosci. Remote Sens. 44(4), 1022-1035.

Loew, A., Schwank, M., Schlenz, F., 2009. Assimilation of an L-Band microwave soil moisture proxy to compensate for uncertainties in precipitation data. IEEE Trans. Geosci. Remote Sens. 47(8), 2606-2616.

Lucas, R., Armston, J., Fairfax, R., Fensham, R., Accad, A., Carreiras, J., Kelley, J., Bunting, P., Clewley, D., Bray, S., Metcalfe, D., Dwyer, J., Bowen, M., Eyre, T., Laidlaw, M., Shimada, M., 2010. An evaluation of the ALOS PALSAR L-Band backscatter-above ground biomass relationship in Queensland, Australia: Impacts of surface moisture condition and vegetation structure. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 3(4), 576-593.

Mahanama, S.P.P., Koster, R.D., Reichle, R.H., Zubair, L., 2008. The role of soil moisture initialization in subseasonal and seasonal streamflow prediction - A case study in Sri Lanka. Adv. Water Resour. 31(10), 1333-1343.

Mahanama, S., Livneh, B., Koster, R., Lettenmaier, D., Reichle, R., 2012. Soil moisture, snow, and seasonal streamflow forecasts in the United States. J. Hydrometeorol. 13(1), 189-203.

Mandelbrot, B., B., 1983. The fractal geometry of nature. New York: W.H. Freeman.

Manninen, T., Stenberg, P., Rautiainen, M., Voipio, P., Smolander, H., 2005. Leaf area index estimation of boreal forest using ENVISAT ASAR. IEEE Trans. Geosci. Remote Sens. 43(11), 2627-2635.

Marzhan, L., Ludwig, R., 2009. On the derivation of soil surface roughness from multi-polarimetric PolSAR data and its potential for hydrological modeling. Hydrol. Earth Sys. Sci. 13(3), 381-394.

Mattia, F., Davidson, M.W.J., Le Toan, T., D'Haese, C.M.F., Verhoest, N.E.C., Gatti, A.M., Borgeaud, M., 2003a. A comparison between soil roughness statistics used in surface scattering models derived from mechanical and laser profilers. IEEE Trans. Geosci. Remote Sens. 41(7), 1659-1671.

Mattia, F., Le Toan, T., 1999. Backscattering properties of multi-scale rough surfaces. J. Electromagn. Waves Appl. 13(4), 493-527.
Mattia, F., Le Toan, T., Davidson, M., Borderies, P., Bachelier, E., Borgeaud, M., 2000. On the effect of multi scale surface roughness on SAR data. CEOS SAR Workshop; 450, 569-574.

Mattia, F., Le Toan, T., Picard, G., Posa, F.I., D'Alessio, A., Notarnicola, C., Gatti, A.M., Rinaldi, M., Satalino, G., Pasquariello, G., 2003b. Multitemporal C-band radar measurements on wheat fields. IEEE Trans. Geosci. Remote Sens. 41(7), 1551-1560.

Mattia, F., Satalino, G., Dente, L., Pasquariello, G., 2006. Using a priori information to improve soil moisture retrieval from ENVISAT ASAR AP data in semiarid regions. IEEE Trans. Geosci. Remote Sens. 44(4), 900-912.

Mattia, F., Satalino, G., Pauwels, V.R.N., Loew, A., 2009. Soil moisture retrieval through a merging of multi-temporal L-band SAR data and hydrologic modelling. Hydrol. Earth Sys. Sci. 13(3), 343-356.

Mauser, W., Schadlich, S., 1998. Modelling the spatial distribution of evapotranspiration on different scales using remote sensing data. J. Hydrol. 213(1-4), 250-267.

McNairn, H., Brisco, B., 2004. The application of C-band polarimetric SAR for agriculture: a review. Can. J. Remote Sens. 30(3), 525-542.

Merlin, O., Chehbouni, A., Kerr, Y., Goodrich, D., 2006. A downscaling method for distributing surface soil moisture within a microwave pixel: Application to the Monsoon '90 data. Remote Sens. Environ. 101(3), 379-389.

Merz, R., Parajka, J., Blöschl, G., 2011. Time stability of catchment model parameters: Implications for climate impact analyses. Water Resour. Res. 47, W02531.

Merzouki, A., McNairn, H., Pacheco, A., 2010. Evaluation of the Dubois, Oh, and IEM radar backscatter models over agricultural fields using C-band RADARSAT-2 SAR image data. Can. J. Remote Sens. 36, 274-286.

Merzouki, A., McNairn, H., Pacheco, A., 2011. Mapping Soil Moisture Using RADARSAT-2 Data and Local Autocorrelation Statistics. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 4(1), 128-137.

Minet, J., Laloy, E., Lambot, S., Vanclooster, M., 2011. Effect of high-resolution spatial soil moisture variability on simulated runoff response using a distributed hydrologic model. Hydrol. Earth Sys. Sci. 15(4), 1323-1338.

Mironov, V.L., Dobson, M.C., Kaupp, V.H., Komarov, S.A., Kleshchenko, V.N., 2004. Generalized refractive mixing dielectric model for moist soils. IEEE Trans. Geosci. Remote Sens. 42(4), 773-785.

Mironov, V.L., Dobson, M.C., Kaupp, V.H., Komarov, S.A., Kleshchenko, V.N., 2002. Generalized refractive mixing dielectric model for moist soils. Proceedings of the IEEE International Geoscience and Remotes Sensing Symposium, Vol 4, pp. 3556-3558.

Mironov, V.L., Kosolapova, L.G., Fomin, S.V., 2009. Physically and Mineralogically Based Spectroscopic Dielectric Model for Moist Soils. IEEE Trans. Geosci. Remote Sens. 47(7), 2059-2070.

Moghaddam, M., Saatchi, S., Cuenca, R.H., 2000. Estimating subcanopy soil moisture with radar. J. Geophys. Res.-Atmos. 105, 14899-14911.

Mohanty, B.P., Famiglietti, J.S., Skaggs, T.H., 2000b. Evolution of soil moisture spatial structure in a mixed vegetation pixel during the Southern Great Plains 1997 (SGP97) Hydrology Experiment. Water Resour. Res. 36(12), 3675-3686.

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

Mohanty, B.P., Skaggs, T.H., Famiglietti, J.S., 2000a. Analysis and mapping of field-scale soil moisture variability using high-resolution, ground-based data during the Southern Great Plains 1997 (SGP97) Hydrology Experiment. Water Resour. Res. 36(4), 1023-1031.

Mohr, K.I., Baker, R.D., Tao, W.K., Famiglietti, J.S., 2003. The sensitivity of West African convective line water budgets to land cover. J. Hydrometeorol. 4(1), 62-76.

Mohr, K.I., Famiglietti, J.S., Boone, A., Starks, P.J., 2000. Modeling soil moisture and surface flux variability with an untuned land surface scheme: A case study from the Southern Great Plains 1997 Hydrology Experiment. J. Hydrometeorol. 1(2), 154-169.

Moran, M.S., Hymer, D.C., Qi, J.G., Sano, E.E., 2000. Soil moisture evaluation using multi-temporal synthetic aperture radar (SAR) in semiarid rangeland. Agric. For. Meteorol. 105(1-3), 69-80.

Narayan, U., Lakshmi, V., Jackson, T.J., 2006. High-resolution change estimation of soil moisture using L-band radiometer and radar observations made during the SMEX02 experiments. IEEE Trans. Geosci. Remote Sens. 44(6), 1545-1554.

Njoku, E.G., Jackson, T.J., Lakshmi, V., Chan, T.K., Nghiem, S.V., 2003. Soil moisture retrieval from AMSR-E. IEEE Trans. Geosci. Remote Sens. 41(2), 215-229.

Notarnicola, C., Angiulli, M., Posa, F., 2006. Use of radar and optical remotely sensed data for soil moisture retrieval over vegetated areas. IEEE Trans. Geosci. Remote Sens. 44(4), 925-935.

Notarnicola, C., Angiulli, M., Posa, F., 2008. Soil moisture retrieval from remotely sensed data: Neural network approach versus Bayesian method. IEEE Trans. Geosci. Remote Sens. 46(2), 547-557.

Oh, Y., 2004. Quantitative retrieval of soil moisture content and surface roughness from multipolarized radar observations of bare soil surfaces. IEEE Trans. Geosci. Remote Sens. 42(3), 596-601.

Oh, Y., Kay, Y.C., 1998. Condition for precise measurement of soil surface roughness. IEEE Trans. Geosci. Remote Sens. 36(2), 691-695.

Oh, Y., Sarabandi, K., Ulaby, F.T., 1992. An empirical model and an inversion technique for radar scattering from bare soil surfaces. IEEE Trans. Geosci. Remote Sens. 30(2), 370-381.

Oh, Y., Sarabandi, K., Ulaby, F., 2002. Semi-empirical model of the ensemble-averaged differential Mueller matrix for microwave backscattering from bare soil surfaces. IEEE Trans. Geosci. Remote Sens. 40(6), 1348-1355.

Oh, Y., Sarabandi, K., Ulaby, F., 1994. An inversion algorithm for retrieving soil moisture and surface roughness from polarimetric radar observation. Proceedings of the IEEE International Symposium on Geoscience and Remote Sensing. pp. 1582-1584.

Onier, C., Chanzy, A., Chambarel, A., Rouveure, R., Chanet, M., Bolvin, H., 2011. Impact of soil structure on microwave volume scattering evaluated by a two-dimensional numerical model. IEEE Trans. Geosci. Remote Sens. 49(1), 415-425.

Paloscia, S., 2002. A summary of experimental results to assess the contribution of SAR for mapping vegetation biomass and soil moisture. Can. J. Remote Sens. 28(2), 246-261.

Pan, M., Wood, E.F., 2010. Impact of accuracy, spatial availability, and revisit time of satellite-derived surface soil moisture in a multiscale ensemble data assimilation system. IEEE J. Sel. Top. App. Earth Observ. Remote Sens. 3(1), 49-56.

Parada, L.M., Liang, X., 2008. Impacts of spatial resolutions and data quality on soil moisture data assimilation. J. Geo. Res.-Atmos. 113(D10), D10101.

Patel, P., Srivastava, H.S., Panigrahy, S., Parihar, J.S., 2006. Comparative evaluation of the sensitivity of multi-polarized multi-frequency SAR backscatter to plant density. Int. J. Remote Sens. 27(2), 293-305.

Pathe, C., Wagner, W., Sabel, D., Doubkova, M., Basara, J.B., 2009. Using ENVISAT ASAR Global Mode data for surface soil moisture retrieval over Oklahoma, USA. IEEE Trans. Geosci. Remote Sens. 47(2), 468-480.

Pauwels, V.R.N., Hoeben, R., Verhoest, N.E.C., De Troch, F.P., Troch, P.A., 2002. Improvement of TOPLATS-based discharge predictions through assimilation of ERS-based remotely sensed soil moisture values. Hydrol. Process. 16(5), 995-1013.
Peel, M.C., Blöschl, G., 2011. Hydrological modelling in a changing world. Prog. Phys. Geogr. 35(2), 249-261.

Penna, D., Tromp-van Meerveld, H.J., Gobbi, A., Borga, M., Dalla Fontana, G., 2011. The influence of soil moisture on threshold runoff generation processes in an alpine headwater catchment. Hydrol. Earth Sys. Sci. 15(3), 689-702.

Peplinski, N.R., Ulaby, F.T., Dobson, M.C., 1995. Dielectric properties of soils in the 0.3-1.3-GHz range. IEEE Trans. Geosci. Remote Sens. 33(6), 1340-1340.

Peters-Lidard, C.D., Pan, F., Wood, E.F., 2001. A re-examination of modeled and measured soil moisture spatial variability and its implications for land surface modeling. Adv. Water Resour. 24(9-10), 1069-1083.

Pierdicca, N., Castracane, P., Pulvirenti, L., 2008. Inversion of electromagnetic models for bare soil parameter estimation from multifrequency polarimetric SAR data. Sensors 8(12), 8181-8200.

Pierdicca, N., Pulvirenti, L., Bignami, C., 2010. Soil moisture estimation over vegetated terrains using multitemporal remote sensing data. Remote Sens. Environ. 114(2), 440-448.

Piles, M., Entekhabi, D., Camps, A., 2009. A change detection algorithm for retrieving high-resolution soil moisture from SMAP radar and radiometer observations. IEEE Trans. Geosci. Remote Sens. 47(12), 4125-4131.

Quesney, A., Le Hégarat-Mascle, S., Taconet, O., Vidal-Madjar, D., Wigneron, J.P., Loumagne, C., Normand, M., 2000. Estimation of watershed soil moisture index from ERS/SAR data. Remote Sens. Environ. 72(3), 290-303.

Rabus, B., When, H., Nolan, M., 2010. The importance of soil moisture and soil structure for InSAR phase and backscatter, as determined by FDTD modeling. IEEE Trans. Geosci. Remote Sens. 48(5), 2421-2429.

Rahman, M.M., Moran, M.S., Thoma, D.P., Bryant, R., Collins, C.D.H., Jackson, T., Orr, B.J., Tischler, M., 2008. Mapping surface roughness and soil moisture using multi-angle radar imagery without ancillary data. Remote Sens. Environ. 112(2), 391-402.

Rahman, M.M., Moran, M.S., Thoma, D.P., Bryant, R., Sano, E.E., Collins, C.D.H., Skirvin, S., Kershner, C., Orr, B.J., 2007. A derivation of roughness correlation length for parameterizing radar backscatter models. Int. J. Remote Sens. 28(18), 3995-4012.

Reichle, R.H., Crow, W.T., Koster, R.D., Sharif, H.O., Mahanama, S.P.P., 2008. Contribution of soil moisture retrievals to land data assimilation products. Geophys. Res. Lett. 35(1), L01404.

Reichle, R.H., Koster, R.D., 2005. Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. Geophys. Res. Lett. 32(2), L02404.

Romshoo, S., Koike, M., Onaka, S., Oki, T., Musiake, K., 2002. Influence of surface and vegetation characteristics on C-band radar measurements for soil moisture content. J. Indian Soc. Remote Sens. 30(4), 229-244.

Ryu, D., Famiglietti, J.S., 2006. Multi-scale spatial correlation and scaling behavior of surface soil moisture. Geophys. Res. Lett. 33(8), L08404.

Ryu, D., Famiglietti, J.S., 2005. Characterization of footprint-scale surface soil moisture variability using Gaussian and beta distribution functions during the Southern Great Plains 1997 (SGP97) hydrology experiment. Water Resour. Res. 41(12), W12433.

Sahebi, M., Angles, J., 2010. An inversion method based on multi-angular approaches for estimating bare soil surface parameters from RADARSAT-1. Hydrol. Earth Sys. Sci. 14(11), 2355-2366.

Sahebi, M.R., Angles, J., Bonn, F., 2002. A comparison of multi-polarization and multi-angular approaches for estimating bare soil surface roughness from spaceborne radar data. Can. J. Remote Sens. 28(5), 641-652.

Said, S., Kothyari, U.C., Arora, M.K., 2008. ANN-based soil moisture retrieval over bare and vegetated areas using ERS-2 SAR data. J. Hydrol. Eng. 13(6), 461-475.

Salama, M.S., van der Velde, R., van der Woerd, H.J., Kromkamp, J.C., Philippart, C.J.M., Joseph, A.T., O'Neill, P.E., Lang, R.H., Gish, T., Wedell, P.J., Su, Z., 2012. Technical note: Calibration and validation of geophysical observation models. Biogeosci. 9, 2195-2201.

Satalino, G., Mattia, F., Davidson, M.W.J., Le Toan, T., Pasquariello, G., Borgeaud, M., 2002. On current limits of soil moisture retrieval from ERS-SAR data. IEEE Trans. Geosci. Remote Sens. 40(11), 2438-2447.

Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowski, B., Teuling, A.J., 2010. Investigating soil moisture-climate interactions in a changing climate: A review. Earth-Sci. Rev. 99(3-4), 125-161.

Shi, J., Chen, K.S., Kim, Y., van Zyl, J.J., Sun, G., O'Neill, P., Jackson, T., Entekhabi, D., 2004. Estimation of soil moisture with L-band multi-polarization radar. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium pp. 815-818.

Shi, J.C., Wang, J., Hsu, A.Y., O'Neill, P.E., Engman, E.T., 1997. Estimation of bare surface soil moisture and surface roughness parameter using L-band SAR image data. IEEE Trans. Geosci. Remote Sens. 35(5), 1254-1266.

Singh, D., Dubey, V., 2007. Microwave bistatic polarization measurements for retrieval of soil moisture using an incidence angle approach. J. Geophys. Eng. 4(1), 75-82.

Sivapalan, M., Wood, E.F., 1986. Spatial heterogeneity and scale in the infiltration response of catchments. In: Gupta, V.K., Rodriguez-Iturbe, I., Wood, E.F., editors. Scale Problems in Hydrology, Norwell, Mass. D. Reidel Publisher. pp. 81-106.

Snoeij, P., Attema, E., Davidson, M., Duesmann, B., Floury, N., Levrini, G., Rommen, B., Rosich, B., 2010. Sentinel-l Radar Mission: Status and Performance. IEEE Aerosp. Electron. Syst. Mag. 25(8), 32-39.

Soria-Ruiz, J., Fernandez-Ordonez, Y., McNairn, H., 2009. Corn monitoring and crop yield using optical and microwave remote sensing. In: Pei-Gee, P.H., editor. Geoscience and Remote Sensing: InTech. p. 405.

Srivastava, H.S., Patel, P., Manchanda, M.L., Adiga, S., 2003. Use of multi-incidence angle RADARSAT-1 SAR data to incorporate the effect of surface roughness in soil moisture estimation. IEEE Trans. Geosci. Remote Sens. 41(7), 1638-1640.

Srivastava, H.S., Patel, P., Sharma, Y., Navalgund, R.R., 2008. Retrieval of surface roughness using multi-polarized ENVISAT ASAR data. Geocarto Int. 23(1), 67-77.
Srivastava, H.S., Patel, P., Navalgund, R.R., 2006. Incorporating soil texture in soil moisture estimation from extended low-1 beam mode RADARSAT-1 SAR data. Int. J. Remote Sens. 27(12), 2587-2598.

Stiles, J.M., Sarabandi, K., 2000. Electromagnetic scattering from grassland Part I: A fully phase-coherent scattering model. IEEE Trans. Geosci. Remote Sens. 38(1), 339-348.

Stiles, J.M., Sarabandi, K., Ulaby, F.T., 2000. Electromagnetic scattering from grassland Part II: Measurement and modeling results. IEEE Trans. Geosci. Remote Sens. 38(1), 349-356.

Su, Z., Troch, P.A., De Troch, F.P., 1997. Remote sensing of bare surface soil moisture using EMAC/ESAR data. EMAC 94/95 Final Results - Workshop Proceedings (136), 53-70.

Taflove, A., Hagness, S.C., 2005. Computational electrodynamics : the finite-difference time-domain method. Boston, Mass. Artech House, p. 1006.

Thoma, D.P., Moran, M.S., Bryant, R., Rahman, M., Holifield-Collins, C.D., Skirvin, S., Sano, E.E., Slocum, K., 2006. Comparison of four models to determine surface soil moisture from C-band radar imagery in a sparsely vegetated semiarid landscape. Water Resour. Res. 42(1), W01418.

Thoma, D.P., Moran, M.S., Bryant, R., Rahman, M.M., Collins, C.D.H., Keefer, T.O., Noriega, R., Osman, I., Skrivin, S.M., Tischler, M.A., Bosch, D.D., Starks, P.J., Peters-Lidard, C., 2008. Appropriate scale of soil moisture retrieval from high resolution radar imagery for bare and minimally vegetated soils. Remote Sens. Environ. 112(2), 403-414.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., Traver, I.N. Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., Rostan, F., 2012. GMES Sentinel-1 mission. Remote Sens. Environ. 120, 9-24.

Ulaby, F., 1974. Radar measurement of soil moisture content. IEEE Trans. Antennas Prop. 22(2), 257-265.

Ulaby, F.T., Batlivala, P.P., 1976. Optimum radar parameters for mapping soil moisture. IEEE Trans. Geosci. Remote Sens. 14(2), 81-93.

Ulaby, F.T., Moore, R.K., Fung, A.K., 1986. Microwave Remote Sensing: Active and Passive, Volume 3. Reading, Mass.: Addison-Wesley Publishers.

Ulaby, F.T., Moore, R.K., Fung, A.K., 1982. Microwave Remote Sensing: Active and Passive, Volume 2. Reading, Mass.: Addison-Wesley Publishers.

Ulaby, F.T., Moore, R.K., Fung, A.K., 1981. Microwave Remote Sensing: Active and Passive, Volume 1. Reading, Mass.: Addison-Wesley Publishers.

Ulaby, F.T., Sarabandi, K., McDonald, K., Whitt, M., Dobson, M.C., 1990. Michigan Microwave Canopy Scattering Model. Int. J. Remote Sens. 11(7), 1223-1253.
Ulaby, F., Dubois, P., van Zyl, J., 1996. Radar mapping of surface soil moisture. J. Hydrol. 184(1-2), 57-84.

Vachaud, G., Silans, P.D., Balabanis, P., Vauclin, M., 1985. Temporal stability of spatially measured soil water probability density function. Soil Sci. Soc. Am. J. 49(4),822-828.

van der Velde, R., Su, Z., van Oevelen, P., Wen, J., Ma, Y., Salama, MS., 2012. Soil moisture mapping over the central part of the Tibetan Plateau using a series of ASAR WS images. Remote Sens. Environ. 120, 175-187.

van Oevelen, P.J., Hoekman, D.H., 1999. Radar backscatter inversion techniques for estimation of surface soil moisture: EFEDA-Spain and HAPEX-Sahel case studies. IEEE Trans. Geosci. Remote Sens. 37(1), 113-123.

Verhoest, N.E.C., De Baets, B., Vernieuwe, H., 2007a. A Takagi-Sugeno fuzzy rule-based model for soil moisture retrieval from SAR under soil roughness uncertainty. IEEE Trans. Geosci. Remote Sens. 45(5), 1351-1360.

Verhoest, N.E.C., De Baets, B., Mattia, F., Satalino, G., Lucau, C., Defourny, P., 2007b. A possibilistic approach to soil moisture retrieval from ERS synthetic aperture radar backscattering under soil roughness uncertainty. Water Resour. Res. 43(7), W07435.

Verhoest, N.E.C., Lievens, H., Wagner, W., Álvarez-Mozos, J., Moran, M.S., Mattia, F., 2008. On the soil roughness parameterization problem in soil moisture retrieval of bare surfaces from synthetic aperture radar. Sensors 8(7), 4213-4248.

Vernieuwe, H., Verhoest, N.E.C., Lievens, H., De Baets, B., 2011. Possibilistic soil roughness identification for uncertainty reduction on SAR retrieved soil moisture. IEEE Trans. Geosci. Remote Sensing 49(2), 628-638.

Vivoni, E.R., Gebremichael, M., Watts, C.J., Bindlish, R., Jackson, T.J., 2008. Comparison of ground-based and remotely-sensed surface soil moisture estimates over complex terrain during SMEX04. Remote Sens. Environ. 112(2), 314-325.

Wagner, W., Blöschl, G., Pampaloni, P., Calvet, J., Bizzarri, B., Wigneron, J., Kerr, Y., 2007. Operational readiness of microwave remote sensing of soil moisture for hydrologic applications. Nordic Hydrol. 38(1), 1-20.

Wagner, W., Lemoine, G., Rott, H., 1999. A method for estimating soil moisture from ERS scatterometer and soil data. Rem. Sens. Environ. 70, 191-207.

Wagner, W., Pathe, C., Doubkova, M., Sabel, D., Bartsch, A., Hasenauer, S., Blöschl, G., Scipal, K., Martínez-Fernández, J., Löw, A., 2008. Temporal stability of soil moisture and radar backscatter observed by the advanced Synthetic Aperture Radar (ASAR). Sensors 8(2), 1174-1197.

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

Walker, J.P., Houser, P.R., Willgoose, G.R., 2004. Active microwave remote sensing for soil moisture measurement: a field evaluation using ERS-2. Hydrol. Process. 18(11), 1975-1997.

Wang, J.R., Hsu, A., Shi, J.C., O'Neill, P.E., Engman, E.T., 1997. A comparison of soil moisture retrieval models using SIR-C measurements over the Little Washita River watershed. Remote Sens. Environ. 59(2), 308-320.

Wang, S.G., Li, X., Han, X.J., Jin, R., 2011. Estimation of surface soil moisture and roughness from multi-angular ASAR imagery in the Watershed Allied Telemetry Experimental Research (WATER). Hydrol. Earth Syst. Sci. 15(5), 1415-1426.

Warnick, K.F., Chew, W.C., 2001. Numerical simulation methods for rough surface scattering. Waves in Random Media 11(1), R1-R30.

Weisse, A., Michel, C., Aubert, D., Loumagne, C., 2001. Assimilation of soil moisture in a hydrological model for flood forecasting. Proceedings of the IAHS Symposium on Soil-Vegetation-Atmosphere Transfer Schemes and Large-Scale Hydrological Models (270), 249-256.

Western, A.W., Grayson, R.B., 1998. The Tarrawarra data set: Soil moisture patterns, soil characteristics, and hydrological flux measurements. Water Resour. Res. 34(10), 2765-2768.

Western, A.W., Zhou, S.L., Grayson, R.B., McMahon, T.A., Blöschl, G., Wilson, D.J., 2004. Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes. J. Hydrol. 286(1-4), 113-134.

Wickel, A.J., Jackson, T.J., Wood, E.F., 2001. Multitemporal monitoring of soil moisture with RADARSAT SAR during the 1997 Southern Great Plains hydrology experiment. Int. J. Remote Sens. 22(8), 1571-1583.

Wigneron, J.P., Parde, M., Waldteufel, P., Chanzy, A., Kerr, Y., Schmidl, S., Skou, N., 2004. Characterizing the dependence of vegetation model parameters on crop structure, incidence angle, and polarization at L-band. IEEE Trans. Geosci. Remote Sens. 42(2), 416-425.

Wilson, D.J., Western, A.W., Grayson, R.B., Berg, A.A., Lear, M.S., Rodell, M., Famiglietti, J.S., Woods, R.A., McMahon, T.A., 2003. Spatial distribution of soil moisture over 6 and 30 cm depth, Mahurangi river catchment, New Zealand. J. Hydrol. 276(1-4), 254-274.

Wood, E.F., Roundy, J.K., Troy, T.J., van Beek, L.P.H., Bierkens, M.F.P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P.R., Kollet, S., Lehner, B., Lettenmaier, D.P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., Whitehead, P., 2011. Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. Water Resour. Res. 47, W05301.

Wooldridge, S.A., Kalma, J.D., Walker, J.P., 2003. Importance of soil moisture measurements for inferring parameters in hydrologic models of low-yielding ephemeral catchments. Envir. Mod. Software 18(1), 35-48.

Wu, T.D., Chen, K.S., 2004. A reappraisal of the validity of the IEM model for backscattering from rough surfaces. IEEE Trans. Geosci. Remote Sens. 42(4), 743-753.

Wu, T., Chen, K.S., Shi, J., Fung, A.K., 2001. A transition model for the reflection coefficient in surface scattering. IEEE Trans. Geosci. Remote Sens. 39(9), 2040-2050.

Yee, K.S., 1966. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. IEEE Trans. Antennas Propagat. 14, 302-307.

Zehe, E., Becker, R., Bardossy, A., Plate, E., 2005. Uncertainty of simulated catchment runoff response in the presence of threshold processes: Role of initial soil moisture and precipitation. J. Hydrol. 315(1-4), 183-202.

Zehe, E., Blöschl, G., 2004. Predictability of hydrologic response at the plot and catchment scales: Role of initial conditions. Water Resour. Res. 40(10), W10202.

Zhan, X.W., Houser, P.R., Walker, J.P., Crow, W.T., 2006. A method for retrieving high-resolution surface soil moisture from hydros L-band radiometer and radar observations. IEEE Trans. Geosci. Remote Sens. 44(6), 1534-1544.

Zhang, Y., Wei, H., Nearing, M.A., 2011. Effects of antecedent soil moisture on runoff modeling in small semiarid watersheds of southeastern Arizona. Hydrol. Earth Syst. Sci. 15(10),3171-3179.

Zienkiewicz, O.C., Taylor, R.L., 1989. The finite element method. New York, NY. McGraw-Hill.

Zribi, M., Baghdadi, N., Holah, N., Fafin, O., 2005. New methodology for soil surface moisture estimation and its application to ENVISAT-ASAR multi-incidence data inversion. Remote Sens. Environ. 96(3-4), 485-496.

Zribi, M., Ciarletti, V., Taconet, O., 2000. Validation of a rough surface model based on fractional Brownian geometry with SIRC and ERASME radar data over Orgeval. Remote Sens. Environ. 73(1), 65-72.

Zribi, M., Dechambre, M., 2002. A new empirical model to inverse soil moisture and roughness using two radar configurations. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium and 24th Canadian Symposium on Remote Sensing 2223-2225.

Zribi, M., Saux-Picart, S., André, C., Descroix, L., Ottlé, C., Kallel, A., 2007. Soil moisture mapping based on ASAR/ENVISAT radar data over a Sahelian region. Int. J. Remote Sens. 28(16), 3547-3565.