IMPROVING HYPER RESOLUTION SOIL MOISTURE ESTIMATION

by

Tasnuva Rouf
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Civil, Environmental and Infrastructure Engineering

Committee:

_____	Dr. Viviana Maggioni, Dissertation Director
_____	Dr. Celso Ferreira, Committee Member
_____	Dr. Mark Houck, Committee Member
_____	Dr. Paul Houser, Committee Member
_____	Dr. Sam Salem, Department Chair
_____	Dr. Kenneth S. Ball, Dean, Volgenau School of Engineering

Date: _____ Spring Semester 2020
George Mason University
Fairfax, VA

Improving Hyper Resolution Soil Moisture Estimation

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Tasnuva Rouf
Master of Science
Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, 2015
Bachelor of Science
Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, 2012

Director: Viviana Maggioni, Assistant Professor
Civil, Environmental and Infrastructure Engineering, George Mason University

Spring Semester 2020
George Mason University
Fairfax, VA

DEDICATION

This is dedicated to my loving parents, Nafisa Quader and Abdur Rouf.

ACKNOWLEDGMENTS

I want to express my gratitude toward a very long list of people without their direct or indirect support and encouragement this thesis might not come into reality. The first and foremost, I want to thank my advisor Dr. Viviana Maggioni for believing me and providing the incredible help, support, and advice whenever needed throughout my graduate studies. I don't believe without her support I would be able to reach this step. Not only she supported for professional development, but also, she gave mental support whenever needed.

I also want to thank my co-supervisor, Dr. Paul Houser for his suggestions and help me to think critically about my research. I also want to thank my committee members Dr. Celso Ferreira and Dr. Mark Houck for their suggestions in my graduate research.

I would like to thank my colleagues and friends at George Mason University who accompanied me in this long path. A special thanks go to my colleague Dr. Yiwen Mei who provided valuable suggestions and helped me technically at every step in the research. I would like to thank all of our Maggioni's research lab group members, their valuable comments in every group meeting and their technical help me to accelerate the process of research.

I would like to take a moment to thank the NASA Science Utilization of the Soil Moisture Active-Passive Mission Program for the funding that helped to offer me a research assistantship and the Mason Provost's Office for the 2020 Dissertation Completion Grant, which facilitated the timely completion of this dissertation.

Finally, I like to express my heedful gratitude to my parents and family members for their tremendous support and encouragement in difficult situations that have enabled me to attain this level. Especially, my incredible mother who took immense responsibility for bringing me at this level and taught me to keep dreaming. She is always just one call away throughout my life. I want to thank my younger sister for helping to relieve my stressed conditions. I especially thank my husband Dr. Shah Toufiqur Rahman for maintaining long-time long-distance relationship and give me support at tough times.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF SYMBOLS

2m air temperature ..... T
2m dew point temperature ..... T_d
Dew point temperature lapse rate ..... Γ_d
Longwave radiation L
Normalized Difference Vegetation Index ..... NDVI
Pressure ..... P
Shortwave radiation ..... S
Specific Humidity ..... q
SRTM terrain elevation ..... Z
Surface roughness ..... z_0
Temperature lapse rate ..... Γ
Vapor pressure ..... E
Wind Speed ..... W
Zero-plane displacement height ..... h_0

ABSTRACT

IMPROVING HYPER RESOLUTION SOIL MOISTURE ESTIMATION

Tasnuva Rouf, Ph.D.

George Mason University, 2020

Dissertation Director: Dr. Viviana Maggioni

The need for improved accuracy of terrestrial hydrological variables across different landscapes is driven by the development of hyper-resolution land surface modeling. The goal of this work is to propose a new framework to estimate surface and root zone soil moisture at resolutions that are useful for decision making and water resources management. In order to achieve this goal, a hyper-resolution atmospheric forcing dataset (temperature, pressure, humidity, wind speed, incident longwave and shortwave radiation) is developed from coarse resolution products using a physically-based downscaling approach. These downscaling techniques rely on correlations with landscape variables, such as topography, temperature lapse rate corrections, surface roughness, and land cover. A proof-of-concept has been implemented over the Oklahoma domain, where high-resolution observations are available for validation purposes. Hourly NLDAS-2 (North America Land Data Assimilation System) atmospheric variables at 0.125° have been downscaled to 500 m over the study area during 2015. Results show that correlation

coefficients between the downscaled forcing dataset and ground observations are consistently higher, and biases are lower than the ones between the NLDAS-2 forcing dataset at their native resolution and ground observations. Results are therefore encouraging as they demonstrate that the 500 m forcing dataset has a good agreement with ground-based information and can be adapted to force a land surface model for soil moisture estimation. A land surface model is then forced with both the native resolution NLDAS-2 dataset and the downscaled one to simulate surface and root zone soil moisture. Model outputs are compared with in situ soil moisture observations at different spatial resolutions. Results show that the hyper-resolution simulation is able to bring modeled surface and root zone soil moisture closer to in situ observations. This is particularly evident in drier than usual cases, due to the improved ability of the downscaled precipitation to detect missed events and no-rain cases. In summary, finer resolution forcings have the potential to improve simulations of soil moisture, and the resolution of precipitation plays a critical role in improving the time series of soil moisture standard-normal deviates. Then, a land data assimilation system is adopted to merge the satellite soil moisture products into the land surface model. Satellite products offer a unique look at global soil moisture variability and have the potential to correct model biases. This work offers a radical improvement over current state-of-the-art forcing data and soil moisture estimates and will move us into the era of hyper-resolution land modeling.

EXTENDED ABSTRACT

Soil moisture plays a significant role in numerous hydrological-related processes (Koster et al. 2004). Surface and root zone soil moisture control the partitioning of the available energy incident on the land surface. Being a storage component for precipitation and radiation, soil water content influences cloud coverage, precipitation, runoff, and evapotranspiration. Moreover, soil moisture is involved in several feedbacks at the local, regional, and global scale. In particular, soil moisture-temperature and soil moisture-precipitation feedbacks have a significant impact on climate-change projections. Therefore, understanding and predicting both weather and climate depend critically on the realistic characterization of soil moisture (Robock et al. 2000).

As the availability of in situ surface measurements is extremely scarce in several regions of the world, satellite retrievals and model simulations are a valuable alternative for estimating soil moisture globally. In this context, the NASA Soil Moisture Active Passive (SMAP) mission, launched on 31st January 2015, measures land surface brightness temperature and radar backscatter (radar failed in July 2015), thus providing information on surface soil moisture (top 5 cm of the soil column). The usefulness of the SMAP-based soil moisture product is limited by its course resolutions (~9-36 km) and surface-only measurement. Therefore, modeling is necessary to estimate soil moisture and its variations at higher time and space resolutions and deeper in the soil column.

Current land surface models have mostly had their origins in coupled weather and climate models, with a focus more on partitioning radiation at the land surface to provide

a lower boundary condition to the atmosphere, rather than tracing the movement of carbon and water at and near the land surface. Accordingly, the spatial resolution of current land surface models has largely been dictated by the spatial resolutions of global weather and climate models: currently, at best, ~100 km for climate models and ~20 km for weather models (somewhat higher resolutions are used by regional models). Much higher resolutions, which are referred to as hyper-resolutions (100 m to 1 km globally), have recently become available and provide much more detailed information about the storage, movement, and quality of carbon and water at and near the land surface (Wood et al. 2011). Hyper-resolution land surface data are fundamental for water resources management and for making decisions related to agricultural productivity, crop yield prediction, and hydroclimatic hazards.

Developing a predictive capability for terrestrial hydrology across landscapes, with water, energy, and nutrients as the drivers of these dynamic systems, faces the challenge of scaling meter-scale process understanding to practical modeling scales. The upper limit of this scaling to achieve meaningful results is closer to 100 m than to the typical scale of current generation weather and climate models. Therefore, hyper-resolution land surface modeling would provide a framework for addressing science questions that we are not able to answer with current modeling capabilities. Further, significant social benefits would accrue because of improved ability to monitor and predict the Earth's terrestrial water, energy, and biogeochemical cycles at the decision- and management-relevant scales.

There are numerous challenges in developing a hyper-resolution land modeling system, ranging from assessing adequate model physics and computing resources to the representation of human impacts on the land surface. A current barrier is developing a global dataset required to parameterize and dynamically force these models at hyper-resolutions. Land models typically require a minimum of seven near-surface atmospheric forcing variables provided at every time step (e.g., hourly), including air temperature and humidity, wind speed, incident longwave and shortwave radiation, and precipitation. Additional pressure, precipitation type, and radiation variables may be required for some model classes but are generally easily deduced from the basic set of seven.

Currently, we do not have global sub-kilometer in situ or satellite observational capabilities from which to derive these forcing variables. Therefore, physical, dynamic, and statistical downscaling approaches have been developed that interpolate the required high-resolution fields from coarser-resolution data incorporating the interactions between the atmosphere and terrestrial surface (Cosgrove et al. 2003; Haylock et al. 2006; Liston and Elder 2006; Girotto, et al. 2014; Sunyer et al. 2015; Gaur and Simonovic 2017).

Land surface states play a crucial role in understanding the complex land-atmosphere interactions that affect climate. Soil properties like soil moisture and temperature strongly influence the atmospheric boundary layer (Sud et al. 1985; Beljaars et al. 1996; Fischer et al. 2007) and thus affect various land-atmosphere exchanges, as well as regional and global weather and climate prediction. Land surface models (LSMs) forced with in situ, or satellite-based data are the primary tools for the estimation of land

surface parameters that are used to initialize weather and regional climate models (Serpetzoglou et al. 2010).

The unavailability of direct observations to quantify heterogeneous variables like soil moisture makes LSM simulations fundamental in hydrology. Nevertheless, LSMs suffer from several uncertainties (Serpetzoglou et al. 2010; Maggioni et al. 2012; Gupta et al. 2014). Specifically, studies evaluating uncertainties in LSM simulations argue that precipitation plays a significant role in the uncertainty observed in model output variables (Gottschalck et al. 2005; Peters-Lidard et al. 2007; Zhou et al. 2012). Satellite-based precipitation products (which are often used as input to LSMs) are inherently characterized by errors of complex nature at high spatiotemporal scales (Hossain and Anagnostou 2005). When propagated through LSMs, they can influence the nonlinear land-atmosphere interaction processes by affecting soil moisture predictions.

This work builds upon past studies and goes one step further by including several novelties to the physical approaches in downscaling and developing an hourly, 500 m hyper-resolution forcing land surface weather boundary condition dataset (2m air temperature and humidity, wind speed and direction, incident longwave and shortwave radiation, and precipitation) during 2015 across Oklahoma, where high-quality and high-resolution observations are available for validation purposes. A land surface model is then forced with three combinations of input variables to simulate surface and root zone soil moisture across the study domain: 1) NLDAS-2 atmospheric forcings at their original resolution; 2) downscaled NLDAS-2 atmospheric variables (i.e., near-surface air temperature and humidity, wind speed and direction, incident longwave and shortwave

radiation, pressure) and original resolution NLDAS-2 precipitation; and 3) downscaled NLDAS-2 atmospheric variables and precipitation. As the last step, SMAP products are assimilated into the LSM to enhance the estimation of soil moisture. SMAP products offer a unique look at global soil moisture variability and have the potential to directly correct modeling errors.

The science questions we seek to answer with this research are:

1. *Is physically-based downscaling a viable approach to produce hyper-resolution atmospheric forcings?*

2. *What is the role of forcing resolution in land surface modeling? And in particular, what is the role of precipitation resolution relative to the other atmospheric forcings on simulated soil moisture?*

3. *How useful is SMAP to land surface modeling? What is the efficiency of a data assimilation system to estimate surface and root-zone soil moisture?*

Thesis Organization

The research carried out to address the aforementioned questions is presented in Chapters 1 through 3. Chapter 1 focuses on developing the methodology to downscale a set of atmospheric forcings. Chapter 2 focuses on investigating the role of the downscaled products for the estimation of soil moisture through a land surface model. Chapter 3 proposes the use of satellite-based products as observations in a land data assimilation system to improve the estimation of surface and root-zone soil moisture.

The work described in Chapters 1−3 is presented in the form of the following research articles:

Chapter 1: Rouf, T., Maggioni, V., Mei, Y., Houser, P., Noonan, M. "A physically-based downscaling technique for atmospheric forcings"; *Journal of Hydrometeorology*, 2019.

Chapter 2: Rouf, T., Maggioni, V., Mei, Y., Houser, P., "Towards hyper-resolution land-surface modeling of surface and root zone soil moisture"; *Journal of Hydrology*, under review.

Chapter 3: Rouf, T., Maggioni, V., Mei, Y., Houser, P. "Investigating the efficiency of assimilating SMAP soil moisture observations in a land surface model"; in preparation.

1. A PHYSICALLY-BASED ATMOSPHERIC VARIABLES DOWNSCALING TECHNIQUE

This chapter presents the development of a physically-based downscaling approach for a set of atmospheric variables that relies on correlations with landscape information, such as topography, surface roughness, and vegetation. A proof-of-concept has been implemented over Oklahoma, where high-resolution, high-quality observations are available for validation purposes. Hourly NLDAS-2 (North America Land Data Assimilation System, Version 2) meteorological data (i.e., near-surface air temperature, pressure, humidity, wind speed, and incident longwave and shortwave radiation) have been spatially downscaled from their original 12.5 km resolution to a 500 m grid over the study area during 2015. Results show that correlation coefficients between the downscaled products and ground observations are consistently higher than the ones between the native resolution NLDAS-2 data and ground observations. Furthermore, the downscaled variables present smaller biases than the original ones with respect to ground observations. Results are therefore encouraging toward the use of the 500 m dataset for land surface and hydrological modeling. This would be especially useful in regions where ground-based observations are sparse or not available altogether, and where downscaled global reanalysis products may be the only option for model inputs at scales that are useful for decision making.

1.1 Introduction

Hyper-resolution (100 m to 1 km globally) land surface modeling has recently become available and provides detailed information about the storage, movement, and

quality of carbon and water at and near the land surface (Wood et al. 2011). Hyper-resolution land surface data are fundamental for water resources management and for making decisions related to agricultural productivity, crop yield prediction, and hydroclimatic hazards. These hyper-resolution land surface data are expected to advance weather forecasting (Senatore et al. 2015), climate prediction (Baker et al. 2017), precise irrigation scheduling (Gibson et al. 2017), quantification of greenhouse gas fluxes (Franz et al. 2017), flood prediction (Maidment 2017), estimation of water scarcity (Zhou et al. 2016), and hydrologic simulations (Ko et al. 2019).

There are numerous challenges in developing a hyper-resolution land modeling system, ranging from assessing adequate model physics and computing resources to the representation of human impacts on the land surface. A current barrier is developing a global dataset required to parameterize and dynamically force these models at hyper-resolutions. Land surface models typically require a minimum of seven near-surface atmospheric forcing variables provided at every time step (e.g., hourly), including air temperature and humidity, wind speed, incident longwave and shortwave radiation, and precipitation. Additional pressure, precipitation type, and radiation variables may be required for some model classes but are generally easily deduced from the basic set of seven.

Currently, we do not have global sub-kilometer in situ or satellite observational capabilities from which to derive these forcing variables. Therefore, physical, dynamic, and statistical downscaling approaches have been developed that interpolate the required high-resolution fields from coarser-resolution data incorporating the interactions between

the atmosphere and terrestrial surface (Cosgrove et al. 2003; Haylock et al. 2006; Liston and Elder 2006; Girotto, Margulis, and Durand 2014; Sunyer et al. 2015; Gaur and Simonovic 2017). For precipitation downscaling, Venugopal et al. (1999) proposed dynamic space-time scaling of rainfall along with a spatial disaggregation scheme at subgrid scales. To account for orographic influences, Badas et al. (2006) considered a modulation function which superimposed to homogeneous and isotropic synthetic fields to take into account for spatial heterogeneity. Using satellite data, Zorzetto and Marani (2019) proposed a downscaling procedure to calculate the point rainfall extreme value distribution and relates it with the ground observation.

Cosgrove et al. (2003) proposed algorithms for developing 0.125º/hourly spatial/temporal-resolution products from nine primary forcing fields at a native resolution of 40 km/3hourly across North America. Their assumption was that the original 0.125º topography differs significantly within a 40 km grid cell, and elevation could be used as the prime factor for downscaling temperature, pressure, specific humidity, and longwave radiation. The downscaled variables were successfully validated against ground observations from the Oklahoma Mesonet network and the Atmospheric Radiation Measurement Program/cloud and radiation test bed, and Surface Radiation observation data.

Another framework was proposed by Liston and Elder (2006) who developed an intermediate-complexity, quasi–physically based, meteorological model (MicroMet) to produce high-resolution (1 km) atmospheric forcings (air temperature, relative humidity, wind speed, incoming shortwave radiation, incoming longwave radiation, surface

pressure, and precipitation). They focused on complex terrain regions in Colorado, Wyoming, Idaho, Arctic Alaska, Svalbard, central Norway, Greenland, and Antarctica. Their downscaling approach applied a temperature–elevation relationship and used meteorological stations at hourly resolution.

Fiddes and Gruber (2014) proposed another physically-based approach, TopoSCALE, to downscale coarse-grid climate variables to a fine-scale sub-grid forcing data (<100 m), primarily based on a high-resolution digital elevation model (DEM). Elevation and topography correction were estimated by normalizing geopotential heights by gravity at sea level to downscale temperature, humidity, shortwave radiation, and wind speed. This method was tested across the Swiss Alps (characterized by a large elevation gradient of 195 – 4634 m above sea-level) against a ground-based validation dataset.

More recently, Tao and Barros (2018) developed a framework to derive high-resolution long-term meteorological forcings for hydrologic modeling from mesoscale atmospheric reanalysis products, including topographic and cloud corrections and a new physical parameterization of near-surface wind fields. The downscaling methodology is applied to 3-hourly North American Regional Reanalysis (NARR) fields originally at 32 km spacing to 1 km/hourly resolution for seven years (2007–2013) over the Southeast U.S. The downscaled datasets were assessed against flux tower observations available in the region, and performance statistics, Root Mean Squared Error (RMSE) had improved for all the variables.

This work builds upon these past studies and goes one step further by including several novelties to the approaches discussed above. We derived dynamic lapse rates based on air and dew point temperature data and elevation for the downscaling of near-surface air temperature and dew point temperature. These two temperature fields are subsequently used to correct air pressure, humidity, and incident longwave radiation. For downscaling wind speed, we assume a log-wind profile and introduce the use of vegetation index for the parameterization of surface roughness and zero-plane displacement height. The downscaling shortwave radiation takes into consideration the optical air depth difference, local illumination, cast-shadowing, portion of the visible sky, and surface reflection to calculate direct, diffusive, and reflected shortwave radiation. These downscaling approaches have been applied to downscale the North American Land Data Assimilation System Phase-2 (NLDAS-2; Cosgrove et al. 2003; Mitchell 2004) dataset (original resolution of 12.5 km) to a 500 m grid across Oklahoma. The main reason for choosing 500 m as the target resolution is that most physical landscape parameters used to downscale the atmospheric variables are available at 500 m resolution. Achieving finer resolutions is possible, but the supporting landscape parameters would need to be downscaled or use higher-resolution physical parameter sources, adding the opportunity for more uncertainties and errors. Additionally, precipitation is one of the most important inputs in a hydrologic model. All the downscaled variables can be potentially used as predictands in a machine learning algorithm for downscaling precipitation, as shown by Mei et al. (2018). This manuscript

represents the first step in this direction and focuses on the methodology to downscale all atmospheric variables except precipitation.

The present manuscript is organized as follows. Section 1.2 introduces the dataset and study area. A comprehensive step-by-step description of the downscaling algorithms is presented in section 1.3. The improvement in the downscaled atmospheric forcings with respect to the original resolution NLDAS-2 is assessed using ground observations and is discussed in section 1.4. Section 1.5 summarizes our conclusions.

1.2 **Dataset**

This work focuses on a domain in Oklahoma, in the Midwestern U.S. during 2015. The region is characterized by a diverse landscape that includes the Great Plains, hills, lakes, and forests. This area is chosen because of the availability of high-quality, high-resolution ground-based data to validate the proposed downscaling approach.

1.2.1 **North American Land Data Assimilation System Phase 2**

The NLDAS-2 dataset is available at hourly temporal resolution and 1/8th-degree grid spatial resolution for a period ranging from 1st January 1979 to present over the contiguous United States. The NLDAS-2 datasets are primarily derived from NARR (Mesinger et al. 2006) interpolated from the 32 km horizontal resolution NARR grid to the 1/8th degree NLDAS-2 grid, adjusted for elevation differences and temporally disaggregated from three-hourly to hourly time scales (Cosgrove et al. 2003; Mitchell 2004). Surface downward shortwave radiation is derived by bias correcting NARR output using Geostationary Operational Environmental Satellite data (Pinker et al. 2003). The downscaling scheme proposed in this work is applied to the NLDAS-2 2m air

temperature, surface pressure, 2m specific humidity, downward longwave and shortwave radiation, and wind speed at 1/8th-degree spatial resolution and hourly temporal resolution. Surface albedo is required for the shortwave downscaling, and it is collected from the NLDAS-2 Noah land surface model. All NLDAS-2 variables are projected to 12.5 km under the USA Contiguous Lambert Conformal Conic coordinate system, approximating the original 1/8th-degree resolution.

1.2.2 Modern-Era Retrospective analysis for Research and Applications, Version 2

The Modern-Era Retrospective analysis for Research and Applications, Version 2, MERRA-2, (Rienecker et al. 2011; Gelaro et al. 2017) provides data since 1980. The advancement in the assimilation system replaces the original MERRA dataset that enables the merging of modern hyperspectral radiance and microwave observations, along with GPS-Radio Occultation dataset. Resolution of MERRA-2 is 0.5°/0.625° in the latitudinal/longitudinal and hourly. Surface roughness and zero-plane displacement height from the MERRA-2 dataset are used to support the downscaling scheme, as described in detail in section 1.3. These data have been interpolated to match the NLDAS-2 resolution of 12.5 km.

1.2.3 Shuttle RADAR Topography Mission Data

The Shuttle RADAR Topography Mission (SRTM; "SRTM Data – CGIAR-CSI SRTM" n.d.) digital elevation model (DEM) data at 90m spatial resolution, originally produced by National Aeronautics and Space Administration (NASA), are a breakthrough in digital mapping and provide high-quality elevation data at the global scale (Survey 2006). The SRTM data are upscaled to 500 m, and NLDAS-2 (12.5 km)

resolution using the average pooling operation. The SRTM elevation data play a major role in the proposed downscaling technique, as shown in the methodology section.

1.2.4 Moderate Resolution Imaging Spectroradiometer

The Moderate Resolution Imaging Spectroradiometer (MODIS) vegetation indices products, MOD13Q1 and MYD13Q1 version 6, at 250m/16-daily resolution are used in the downscaling framework. By considering both MOD13Q1 and MYD13Q1, we obtain Normalized Difference Vegetation Index (*NDVI*) for every 8 days. The 8-daily *NDVI* is upscaled to 500 m and 12.5 km representing the target and original resolution.

1.2.5 Ground observations

The installation of the Oklahoma Mesonet network began in the 1980s, as the result of an ongoing collaboration between Oklahoma State University and the University of Oklahoma to develop a near real-time, extremely reliable source of surface observational data on local weather conditions across the state ("Oral History Interview with Fred V. Brock" n.d., "Oral History Interview with Ken Crawford" n.d.). As a result, a statewide network of 121 automated environmental monitoring stations was officially launched in March 1994. These provide regular measurements of air and soil temperature, barometric pressure, rainfall, relative humidity, solar radiation, wind speed and direction, and soil moisture, both direct and calculated, including instrumentation data, all of which are regularly transmitted every 5 minutes to the Oklahoma Climatological Survey, where the data quality is verified (Martens et al. 2017). Temperature, pressure, specific humidity, longwave radiation and shortwave radiation,

wind speed and direction observations at several stations across Oklahoma are used for validation purposes in this study (Figure 1; “Mesonet | Home Page” n.d.).

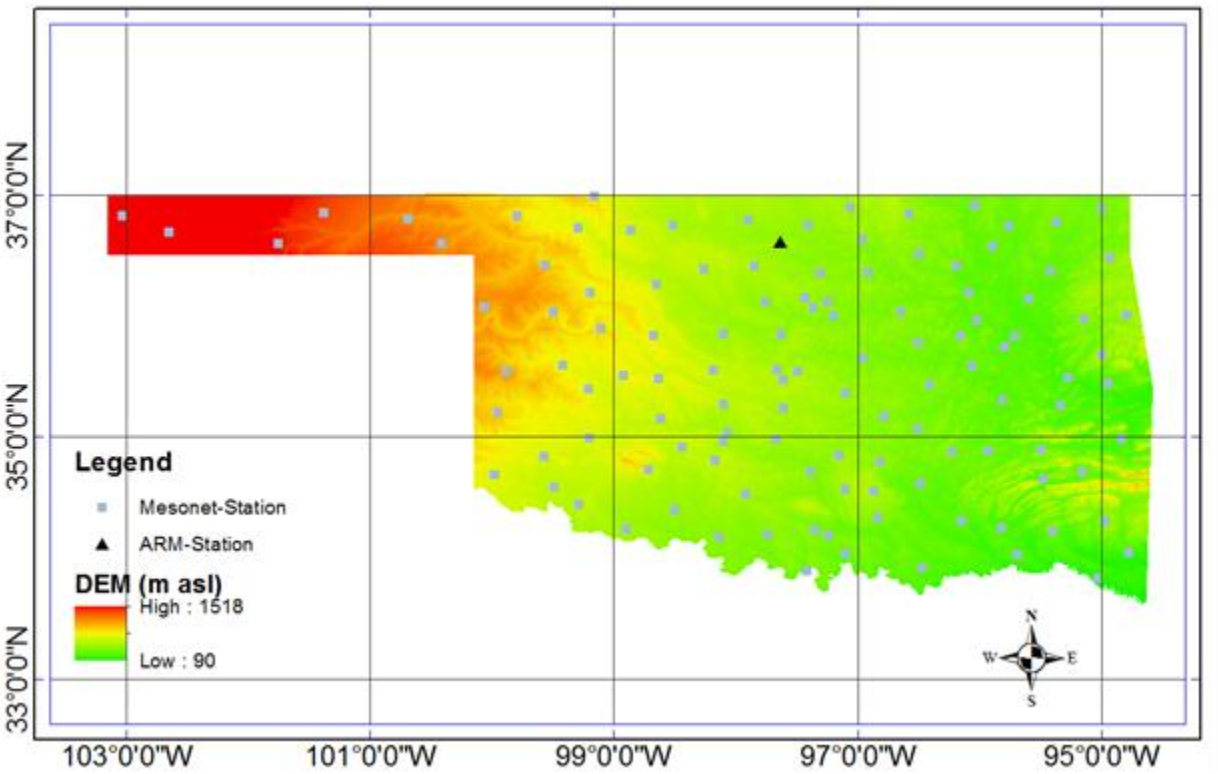

Figure 1. Study region and location of the Mesonet and ARM stations.

The Atmospheric Radiation Measurement (ARM) Program, supported by the U.S. Department of Energy, had an intention to improve the understanding of processes that affect atmospheric radiation and the description of these processes in climate models. To achieve this goal, the ARM Program employed five highly instrumented primary measurement sites at land and ocean locations for up to 10 years, from the Tropics to the Arctic. These provide a measurement of longwave and shortwave radiation, along with

several meteorological variables including wind velocity, precipitation rate, surface moisture, temperature, and fluxes of sensible and latent heat (Stokes and Schwartz, 1994). In 2015, for the launch of the NOAA Geostationary Operational Environmental Satellite R-Series satellite, the sites were improved with new capabilities (Michalsky et al. 2016). In this study, we have used longwave radiation data at the Oklahoma site, which are available at hourly resolution. The ground observations used in this study are from the Oklahoma Mesonet network and one ARM site, which are independent of both NLDAS and NARR data (Xia et al. 2012).

1.3 **Methodology**

The overarching idea behind the proposed methodology is that topography and vegetation cover have a large impact on surface atmospheric conditions. Therefore, we propose to use high-resolution topography and vegetation information to spatially downscale NLDAS-2 2m air temperature, surface pressure, 2m humidity, incident longwave/shortwave radiation, and 10m wind speed from 12.5 km to a 500 m regular Cartesian grid across Oklahoma. The downscaling approach comprises of two main steps: 1) a statistical interpolation (bilinear interpolation) of all NLDAS-2 variables to match the downscaled terrain resolution of 500 m, and 2) deterministic downscaling rules to account for the local scale effects.

1.3.1 **Air and Dew Point Temperature**

The downscaling of air and dew point temperature is based on a lapse rate correction that accounts for the strong temperature–elevation relationship. These two variables are used to derive the 500 m pressure, humidity, and downward longwave

radiation (that will be introduced in section 1.3.2). While 2m air temperature (T) is available from NLDAS-2, dew point temperature at the same altitude needs to be derived first. The 2m dew point temperature (T_d) is calculated (in K) from vapor pressure using the (Lawrence 2005) method:

$$T_d = \frac{C_3 \ln\left(\frac{E}{C_1}\right)}{C_2 - \ln\left(\frac{E}{C_1}\right)} \quad (1)$$

where C_1, C_2, and C_3 are constant, and their values are 611.21Pa/611.15Pa, 17.368/22.452, and 238.88°C/272.55°C for water/ice, adopted from Buck (1981). E stands for vapor pressure (Pa) and is derived from the NLDAS-2 pressure and specific humidity by reorganizing the following equation that $E = \frac{P}{0.622 \mp 0.378q}$, where P and q represent the NLDAS-2 air pressure (Pa) and specific humidity (kg/kg), respectively.

The lapse rate corrections of air and dew point temperature are defined as follows:

$$\hat{T} = T + \Gamma\left(\hat{Z} - Z\right) \quad (2)$$

$$\widehat{T_d} = T_d + \Gamma_d\left(\hat{Z} - Z\right) \quad (3)$$

where a variable with/without "^" denotes the downscaled/original resolution one. T (K) is the NLDAS-2 2m air temperature. Z and $\hat{Z}$ (m asl) are SRTM terrain elevation at the 12.5 km resolution and at 500 m resolution, respectively. Γ and Γ_d (K/m) correspond to the temperature and dew point temperature lapse rates. The elevation and air temperature differences between a target grid cell and its eight nearest neighbors at each time step are calculated, and a line is fitted to describe the T-Z relationship. The slope of the fitted line

is taken as the estimate of Γ for the target grid cell. This process is repeated for all grid cells and time steps and similarly for Γ_d. As shown by maps of the lapse rate means and standard deviations, the average Γ is not a constant (although commonly assumed to be -6.5 K/km) and the standard deviation is not null over time and across the study domain (Figure 2). Γ and Γ_d are computed based on NLDAS-2 because the goal is to produce a robust methodology which will generate high-resolution data from any coarse resolution product only with the help of elevation data and MODIS vegetation cover data, which are available all over the world at the target resolution.

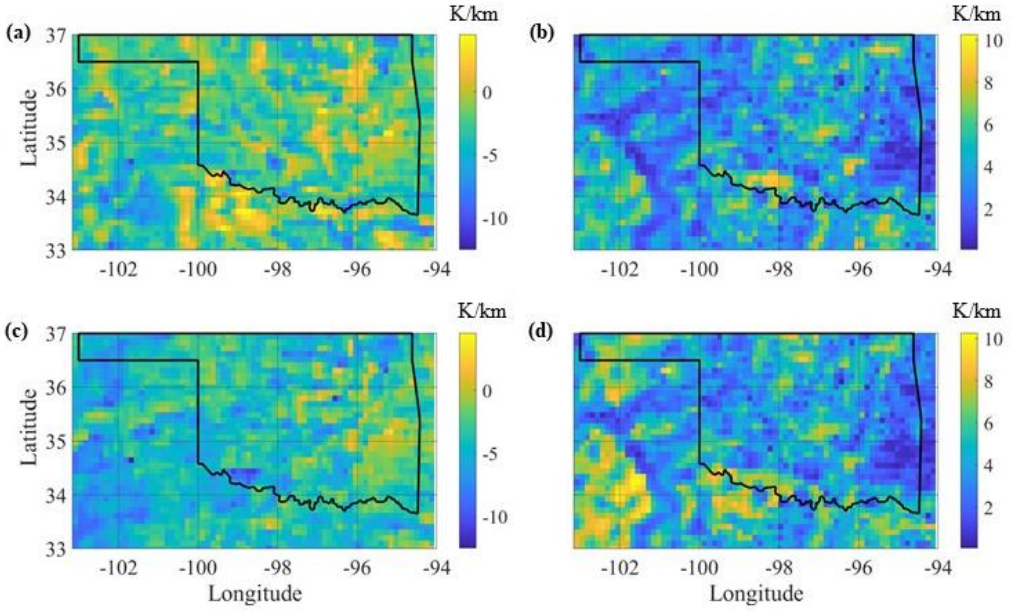

Figure 2. (a) Mean and (b) standard deviation of air temperature-based lapse rates and (c) mean and (d) standard deviation of dew point temperature-based lapse rates across Oklahoma during 2015.

Figure 3 shows a comparison between the air temperature downscaling technique proposed here that adopts a dynamic lapse rate and a constant lapse rate. Correlation

coefficients are also computed between the downscaled air temperature NLDAS-2 and the corresponding MESONET observations across Oklahoma during 2015. Firstly, the correlation coefficient for the dynamic lapse rate case (0.97) is higher than the one obtained with a constant lapse rate (0.93), which is very close to the correlation between the original resolution NLDAS-2 and the in situ observations (0.94, as shown in Table 1), demonstrating minimal improvement due to the downscaling process in the latter case. Secondly, the downscaled temperature obtained assuming a constant lapse rate is consistently higher than the observed temperature, showing an overestimation, especially at low temperatures. Thirdly, the deviation around the 1:1 line is largely reduced when the dynamic lapse rate is introduced.

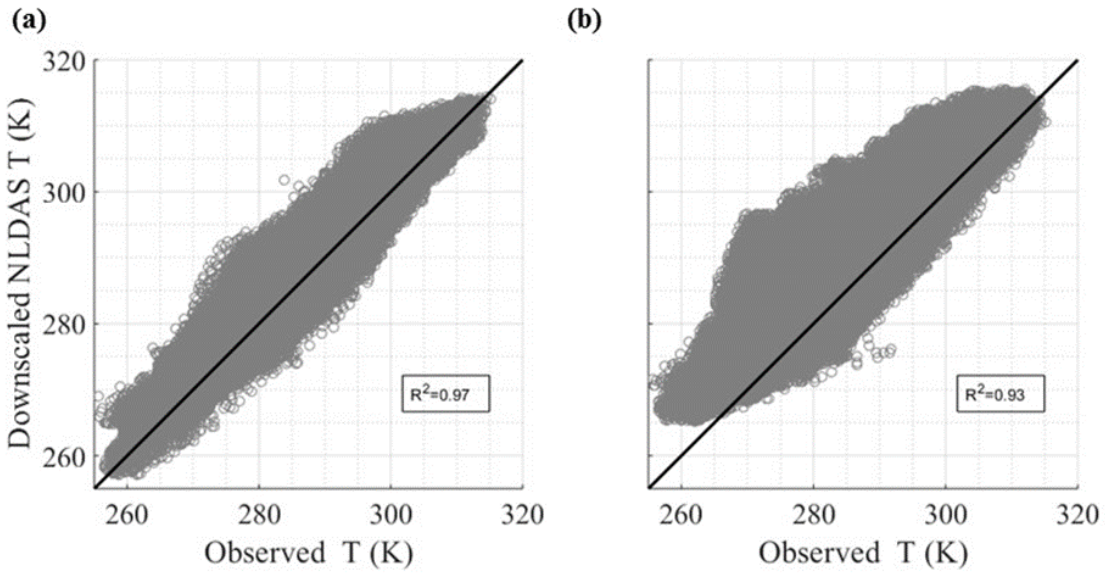

Figure 3. Density scatter plots of hourly NLDAS-2 2m air temperature downscaled to 500 m resolution, using (a) a dynamic lapse rate and (b) a constant lapse rate against the corresponding Mesonet ground observations across Oklahoma during 2015.

1.3.2 **Pressure, Humidity, and Incident Longwave Radiation**

The downscaled 2m air temperature is used to adjust the surface pressure using the hydrostatic approximation and the ideal gas law (following the methodology developed by Cosgrove et al. 2003):

$$\hat{P} = Pe^{-\frac{g(\hat{Z}-Z)}{RT_m}} (4) \quad (4)$$

where R is the ideal gas constant (287J·kg^{-1}·K^{-1}), g is the gravitational acceleration (9.81m/s^2), and T_m (K) is the mean air temperature between the T and $\hat{T}$.

The specific humidity is calculated with $\hat{P}$ and $\hat{E}$ as follows:

$$\hat{q} = \frac{0.622\hat{E}}{\hat{P} - 0.378\hat{E}} \quad (5)$$

$\hat{E}$ is calculated based on the Magnus formula with $\hat{T}_d$ ($\hat{E} = C_1 e^{\frac{C_2\hat{T}_d}{\hat{T}_d + C_3}}$).

The downscaling of downward longwave radiation is based on the Cosgrove et al. (2003) method, which is derived from the Stefan-Boltzmann law:

$$\hat{L} = \frac{\hat{\varepsilon}}{\varepsilon}\left(\frac{\hat{T}}{T}\right)^4 L \quad (6)$$

where L is the NLDAS-2 incident longwave radiation (W/m^2), and ε is emissivity. The empirical relationship suggested by Cosgrove et al. (2003) is applied here to calculate ε and $\hat{\varepsilon}$.

1.3.3 **Incident Shortwave Radiation**

The shortwave radiation downscaling technique comprises four steps from different methodologies. First, the global shortwave radiation (S) is partitioned into direct (S_b) and (S_d) diffuse radiation based on the regression model proposed by Ruiz-Arias et al. (2010). This model estimates the solar transmissivity of the atmospheric column by defining the ratio of incident shortwave radiation between the surface level and the top of the atmosphere as the clearness index. Larger values of the clearness index indicate larger weighting for the direct shortwave. As a second step, S_b is adjusted for optical air depth difference, local illumination, and cast-shadowing and S_d is adjusted for sky obstruction (Fiddes and Gruber 2014; Tao and Barros 2018). In the third step, a reflected radiation component is estimated with the adjusted S_b and S_d, albedo, and a terrain configuration factor (Tao and Barros 2018). Fourth, the downscaled global shortwave radiation ($\hat{S}$) results from the summation of the three components:

$$\hat{S} = \underbrace{\delta\, cos(\theta)\, e^{k(\hat{P}-P)} S_b}_{Beam} + \underbrace{F_v S_d}_{Diffuse} + \underbrace{A F_t \left[\widehat{S_b} + (1 - F_v)\widehat{S_d}\right]}_{Reflected} \tag{7}$$

where δ is a binary shadowing mask indicating whether the location is blocked by the shadow of nearby terrain. The presence of shadow implies that the local horizontal angle on the solar azimuth direction is higher than the solar altitude and vice-versa. *cos(θ)* is the cosine of the solar illumination angle, which indicates if the sun is below or above the local horizon:

$$cos(\theta) = cos\,\theta_z\, cos\,\beta + \sin\theta_z \sin\beta\, cos(\theta_a - \alpha) \tag{8}$$

where θ_z and θ_a are the solar zenith and azimuth, respectively; α and β stand for the terrain aspect and slope, respectively; and k is the broadband attenuation coefficient (in Pa^{-1}) calculated according to Sen Gupta and Tarboton (2016):

$$k = -\frac{ln(S_T) - ln(S)}{P} \tag{9}$$

where S_T is the top-of-atmosphere incident shortwave radiation, computed based on $\boldsymbol{S_T = S^* \cos \theta_z}$, with S^* being the solar constant (S^* = 1370 W/m^2; Dingman 2015). F_v is the sky-view factor indicating the portion of the visible sky of a location. F_t is the terrain configuration factor, which is a function of both sky-view factor and slope. Both F_v and F_t are calculated using the SAGA-GIS Sky View Factor Module (Häntzschel et al. 2005). A is the surface albedo from NLDAS-2. This study is conducted over Oklahoma which is a region characterized by neither frequent snowfalls nor sharp topography. Therefore, the effects of subgrid surface heterogeneity on the albedo are minor. In areas where the frequency of snow is high, albedo at finer resolution should be considered to account for such heterogeneity.

1.3.4 **Wind Speed**

The downscaling of wind speed (W, m/s) is based on an adjustment for friction velocity, terrain slope, aspect, and curvature (Tao and Barros 2018; Liston and Elder 2006). We start by assuming a logarithmic wind profile:

$$\widehat{W} = \frac{\widehat{u_*}}{\kappa} ln \frac{H - \widehat{h_0}}{\widehat{z_0}} \tag{10}$$

where $\hat{W}$ is wind speed adjusted for the friction velocity $\hat{u}_*$ (m/s), κ is the Von Kármán constant (~0.41), H is the measurement height (m above ground), $\hat{z}_0$ is the surface roughness (m), and $\hat{h}_0$ is the zero-plane displacement height (m). The friction velocity at high resolution is calculated by taking advantage of the dependence of the geostrophic drag coefficient on surface roughness and the assumption that the geostrophic wind component remains constant at different scales (Tao and Barros 2018):

$$\widehat{u_*} = u_* \left(\frac{\widehat{z_0}}{z_0}\right)^{0.09} \tag{11}$$

u_* is calculated from the NLDAS-2 wind speed (W), MERRA-2 surface roughness (z_0), and zero-plane displacement height (h_0) using Eq.(10). Surface roughness and zero-plane displacement height are related to the type and height of vegetation (Allen et al. 2007; Bastiaanssen 2000; Dong et al. 2001). There are some other factors on which surface roughness and displacement height depend, e.g., land cover and snow cover. However, our analysis over Oklahoma found the vegetation index as the most important factor to downscale surface roughness. If other regions of the world are considered, the dependency of surface roughness and displacement height on such factors should be investigated. Given the difficulties in measuring vegetation characteristics, *NDVI* is often used as a proxy to estimate z_0 and h_0 (Bastiaanssen 2000). In this study, a customized function for z_0 based on the MODIS *NDVI* is developed for every time step:

$$z_0 = e^{a_1 NDVI + a_2} \tag{12}$$

where the coefficient a_1 and a_2 are evaluated by log-transferring z_0 and *NDVI* at the NLDAS-2 resolution. Coefficients a_1 and a_2 have been calibrated for each time step over

the entire region (40x72 pixels); however, future applications of this methodology in more heterogeneous areas should consider a spatial analysis of these two parameters. Figure 4 shows a histogram of the coefficients of determination (R^2) between log-transferred z_0 and *NDVI*. Most of the R^2 values lie within 0.5 to 0.8, which reveals the strong correlation between the two variables. $\hat{h_0}$ is estimated using a transfer function evaluated between z_0 and h_0 Nicholas and Lewis (1980):

$$log_{10}\, z_0 = a_3\, log_{10}\, h_0 + a_4 \tag{13}$$

The coefficient a_3 and a_4 are evaluated at a coarse resolution for every time step, and then $\hat{z_0}$ is substituted in for $\hat{h_0}$.

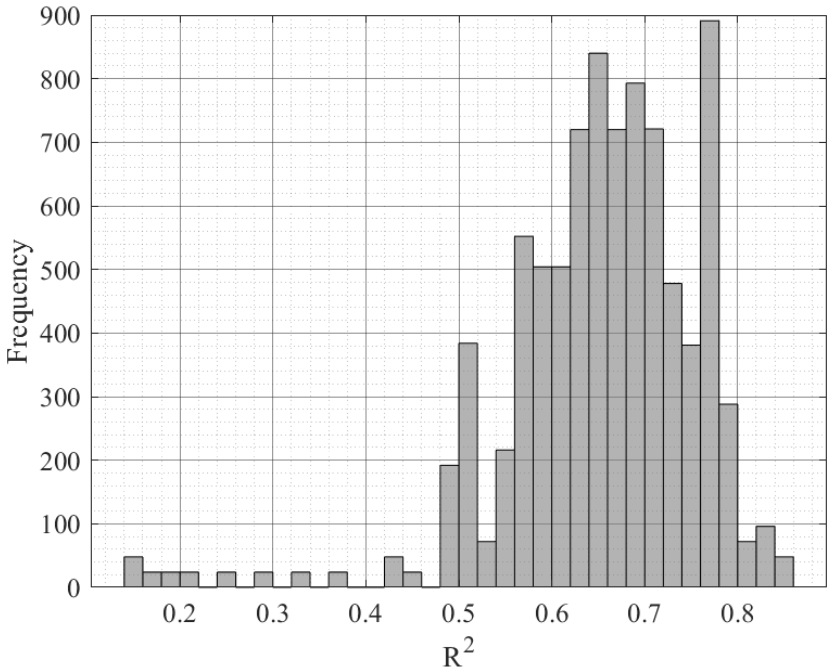

Figure 4. Frequency distribution of R^2 for the linear regression between log-transferred surface roughness and MODIS *NDVI*.

The downscaled wind speed is further adjusted by the terrain slope and curvature following Liston and Elder (2006):

$$\widehat{W_t} = \widehat{W}\left(1 + \frac{\Omega_s}{2} + \frac{\Omega_c}{2}\right) \quad (14)$$

where $\hat{W}_t$ is the final 500 m wind speed adjusted for the friction velocity and terrain characteristics (m/s). Ω_s and Ω_c are the slope in the wind direction and curvature of terrain, respectively. $\Omega_s$and Ω_c are available at 500 m resolution – note that no hat is used for these symbols since they are not available at the NLDAS-2 native resolution of 12.5 km.

1.4 **Results**

The annual average for the six atmospheric variables over Oklahoma during 2015 is presented in Figure 5 and Figure 6 at the original NLDAS-2 resolution (left panels) and at the downscaled 500 m resolution (right panels). Specifically, Figure 5 presents maps of 2m air temperature, surface pressure, and 2m specific humidity, whereas Figure 6 shows maps of downward longwave radiation, downward shortwave radiation, and 10m wind speed. These figures highlight how the spatial patterns of the original resolution NLDAS-2 variables are retained in the downscaled maps. For instance, the temperature gradient (i.e., from colder NW regions to the warmer SE plains) is evident at both resolutions (Figure 5a and 5b) and is consistent with the topography gradient illustrated in Figure 1. However, the downscaled resolution variables are able to capture more detail, thanks to the inclusion of physiological feature information (e.g., orography, vegetation) in the downscaling techniques.

Five downscaled variables – air temperature, pressure, specific humidity, downward shortwave radiation, and wind speed – are then validated against ground observations at 118 Mesonet stations, whereas downward longwave radiation is validated against data collected at the ARM station. Specific humidity is converted to relative humidity for comparison with the measurements recorded by the Mesonet network. Firstly, we present scatterplots of NLDAS-2 data (both at their native resolution and the downscaled products) against ground observations (Figure 7 and Figure 8). Secondly, we analyze maps of correlation coefficients across all sites in the study region to investigate the spatial variability of the downscaling technique performance (Figure 9 and Figure 10). Lastly, we compute overall statistics, including average correlation coefficient, additive bias, RMSE, and Nash Sutcliffe model efficiency (NSE) to summarize the performance of the proposed approaches for each variable across Oklahoma (Table 1).

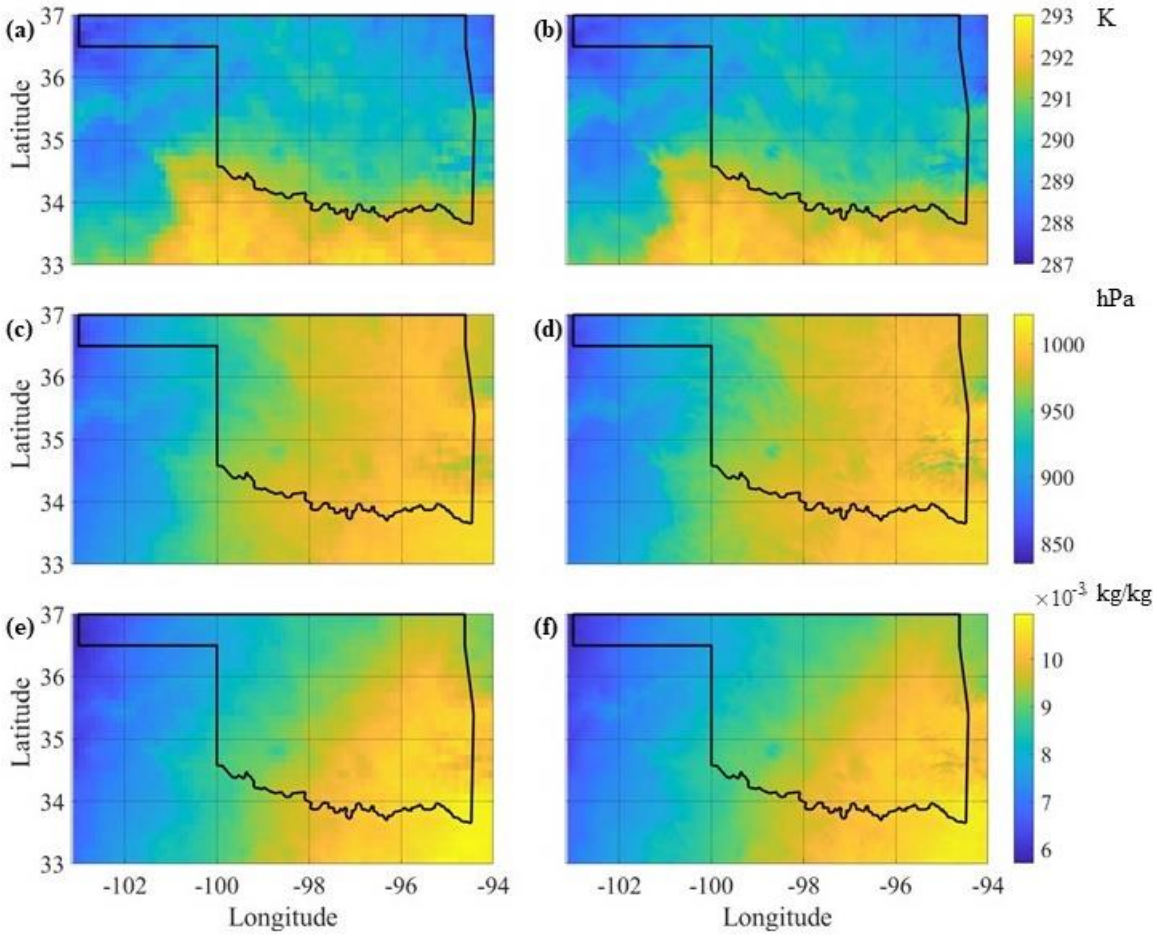

Figure 5. Annual average maps of (a, b) 2m air temperature, (c, d) surface pressure, and (e, f) 2m specific humidity at the original NLDAS-2 resolution (left panels) and at the downscaled 500 m resolution (right panels).

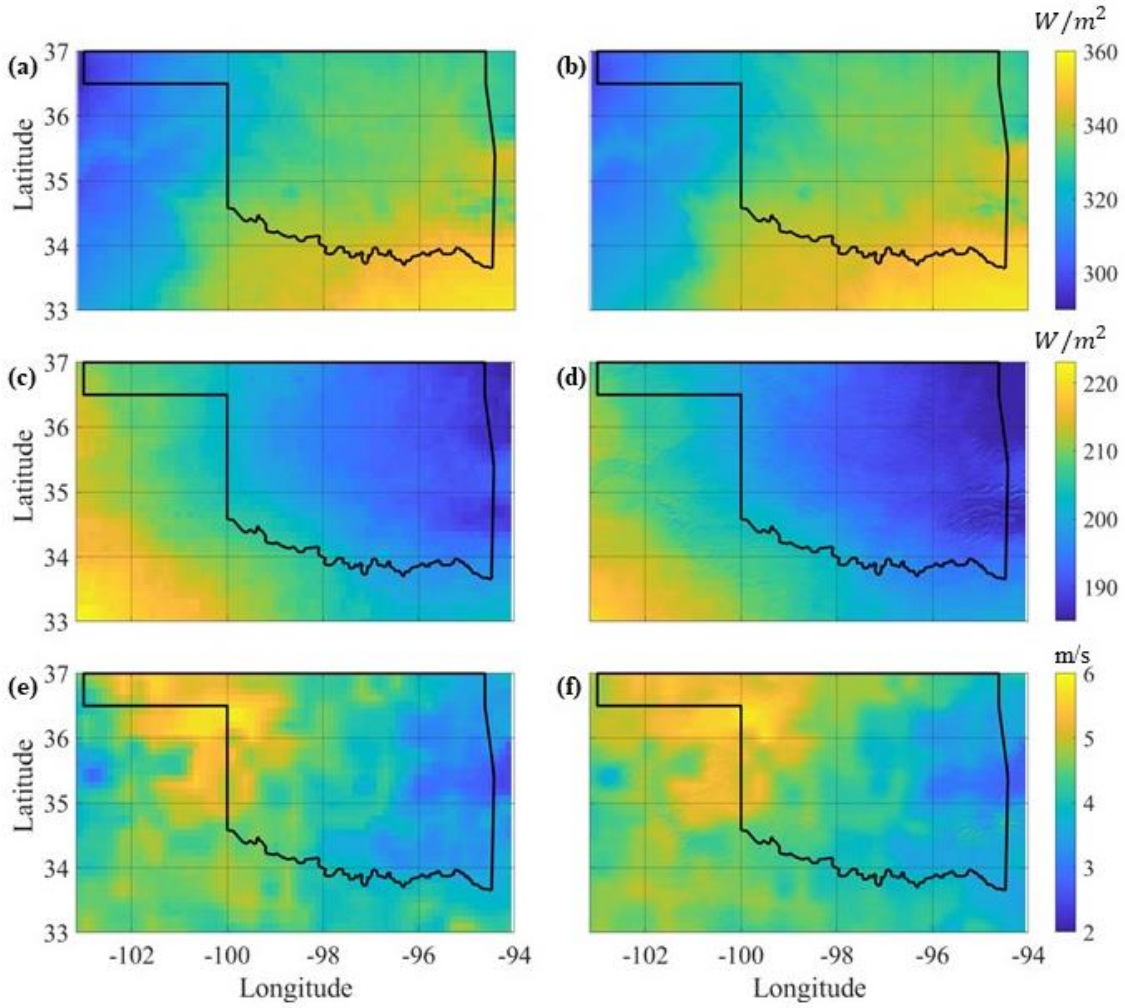

Figure 6. Annual average maps of (a, b) downward longwave radiation, (c, d) downward shortwave radiation, and (e, f) 10m wind speed at the original NLDAS-2 resolution (left panels) and at the downscaled 500 m resolution (right panels).

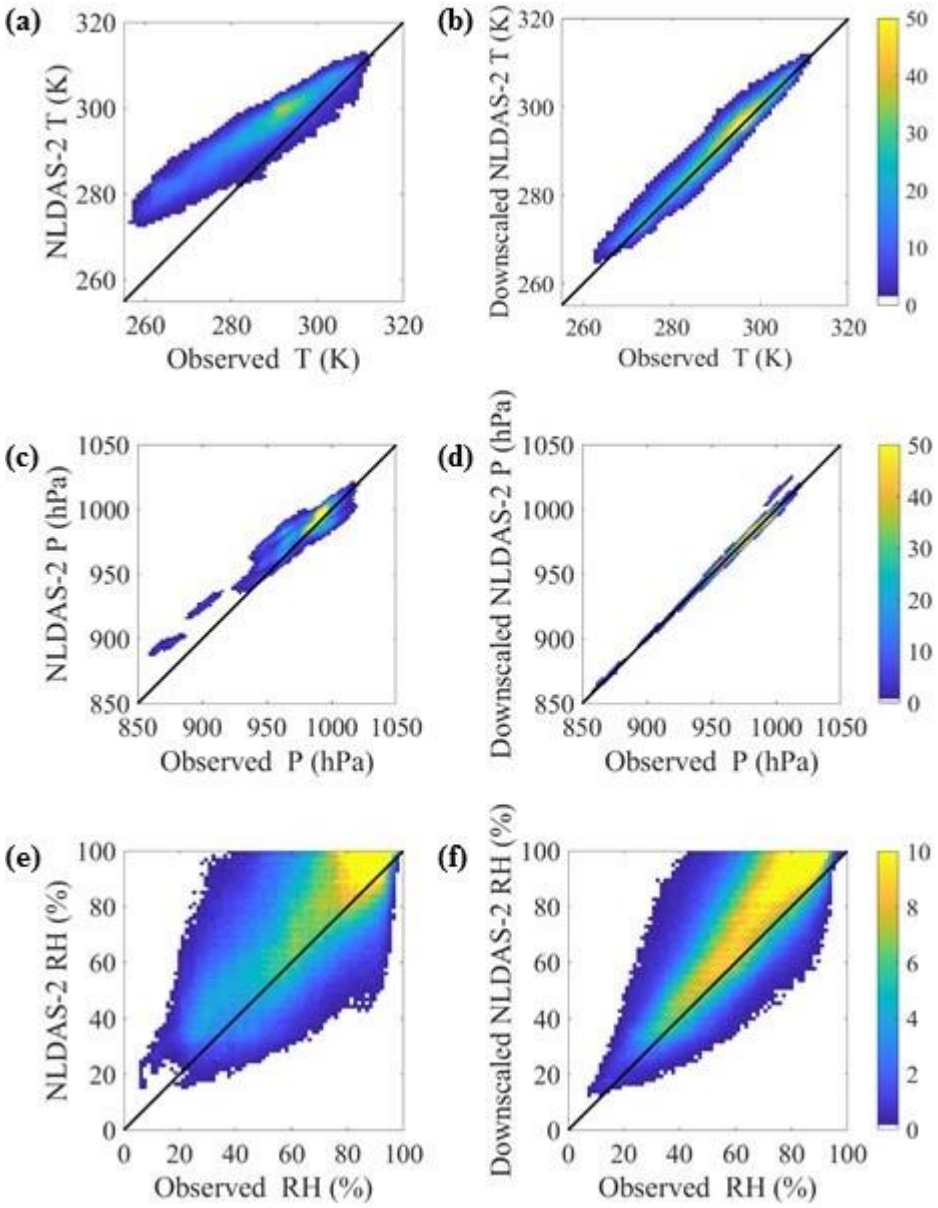

Figure 7. Density scatter plots of NLDAS-2 variables at their original resolution (left panels) and at the downscaled 500 m resolution against the corresponding Mesonet ground observations for (a, b) 2m air temperature, (c, d) surface pressure, (e, f) 2m relative humidity (RH) at hourly resolution.

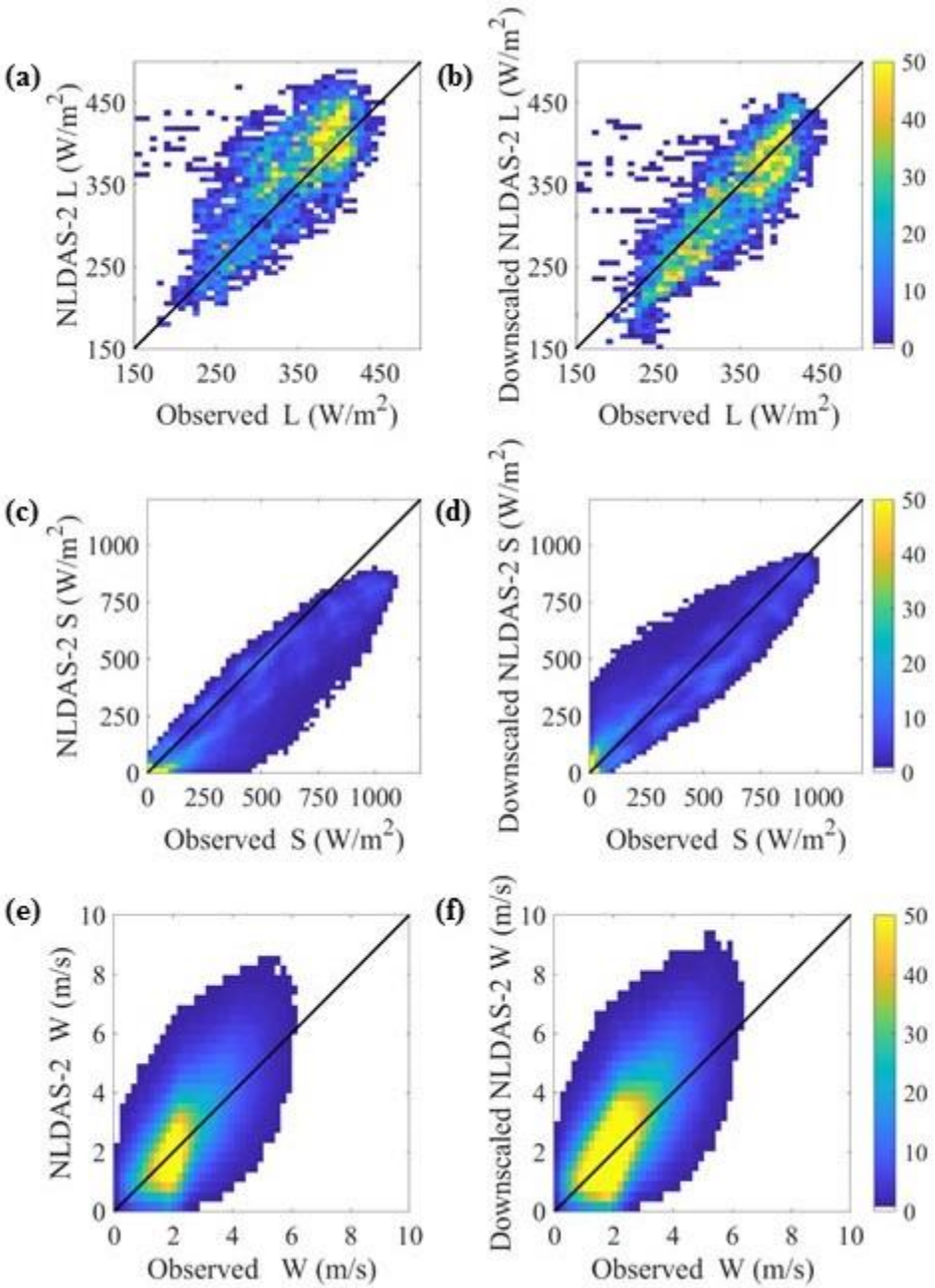

Figure 8. Density scatter plots of NLDAS-2 variables at their original resolution (left panels) and at the downscaled 500 m resolution against the corresponding Mesonet ground observations for (a, b) downward longwave radiation (L), (c, d) downward shortwave radiation (S), and (e, f) 10m wind speed (W) at hourly resolution.

The downscaled temperature and pressure products align more closely with the Mesonet observations than the original resolution NLDAS-2 data, as shown by the density

scatterplots in Figure 7. The dynamic lapse-rate correction largely improves both the air temperature and pressure estimates, moving them closer to the 1:1 line. The correlation coefficient maps presented in Figure 9 a-d corroborate this point: although correlations between the original resolution NLDAS-2 estimates and observations are already high, they are even higher when the downscaled temperature and pressure products are considered. These maps also show that there is no spatial bias in the downscaling algorithm performance since the improvement in the correlation coefficient is evident across the 118 Mesonet stations. For temperature, the correlation coefficient improves from 0.94 for the original resolution to 0.97 for the downscaled product, whereas for pressure, the correlation coefficient goes from 0.94 to 1.00 (Table 1). Although ground observations are quite representative of the coarse grid, the improvement in the correlation coefficients demonstrates that finer resolutions would be even more representative of the in situ observations.

The downscaled relative humidity slightly deviates from the observed data; however, the proposed methodology shows a large improvement over the original resolution NLDAS-2 data (Figure 7 e-f). Overall the correlation improves from 0.69 to 0.92. Similarly, to what observed for temperature and pressure, correlations improve consistently at all stations across the study domain (Figure 9 e-f).

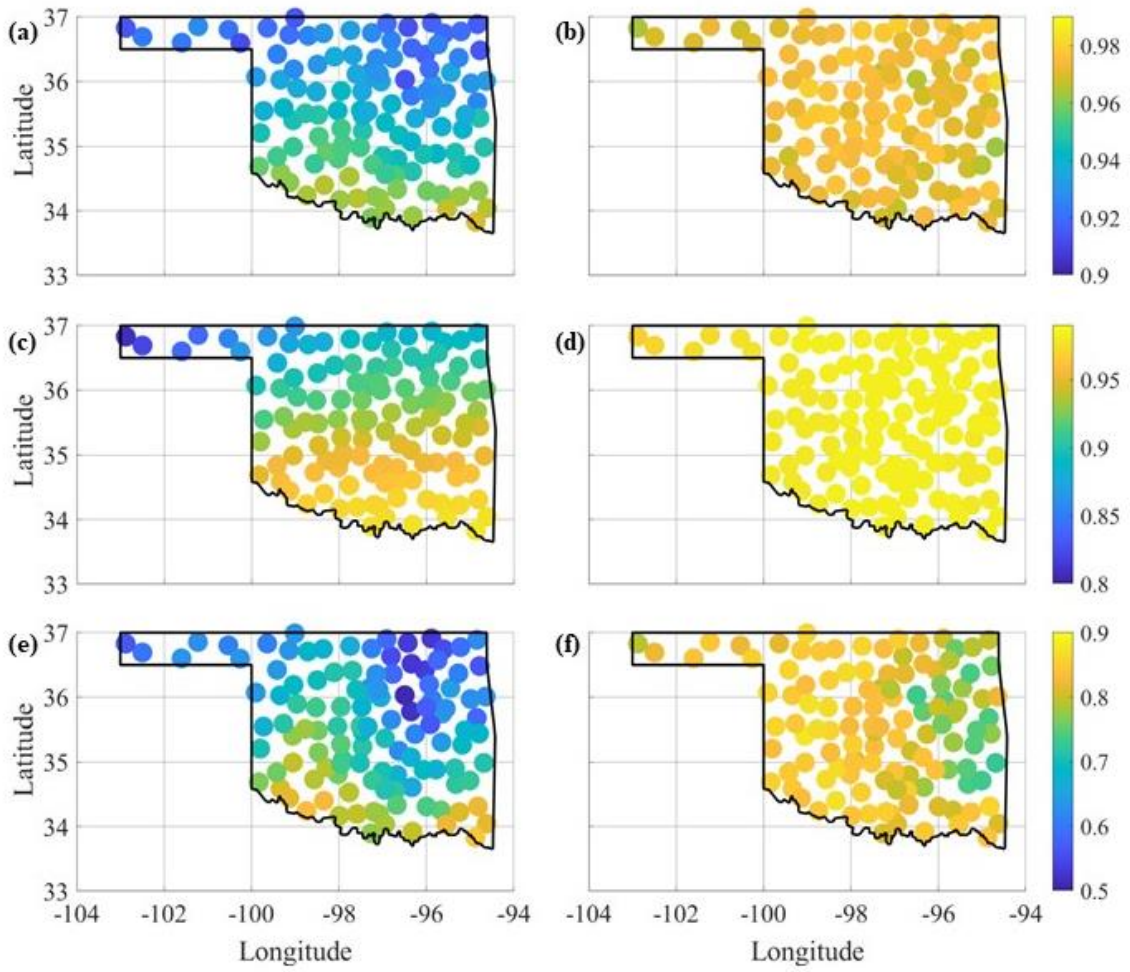

Figure 9. Maps of correlation coefficients of NLDAS-2 variables at their original resolution (left panels) and at the downscaled 500 m resolution against the corresponding Mesonet ground observations for (a, b) 2m air temperature, (c, d) surface pressure, (e, f) 2m relative humidity.

The ARM site is used to validate the downscaled downward longwave radiation product. The sample size is smaller compared to the Mesonet stations, but still significant since data collected at the ARM sites are at hourly resolution and one year of data is considered, with a total of 8,627 data points. Although the improvement in this variable is not as great as the previous ones, both the scatterplots (Figure 8 a-b) and the correlation maps (Figure 10 a-b) reveal an improvement in the downscaled product. Specifically, the

downscaling scheme shifts the 500 m product closer to the ground observations, and the correlation improves from 0.69 to 0.82. Analogously to longwave radiation, the improvement in the downscaled shortwave radiation is not obvious from the scatterplot (Figure 8 c-d), but Figure 9 c-d show higher correlations in the downscaled product over the original NLDAS-2 data.

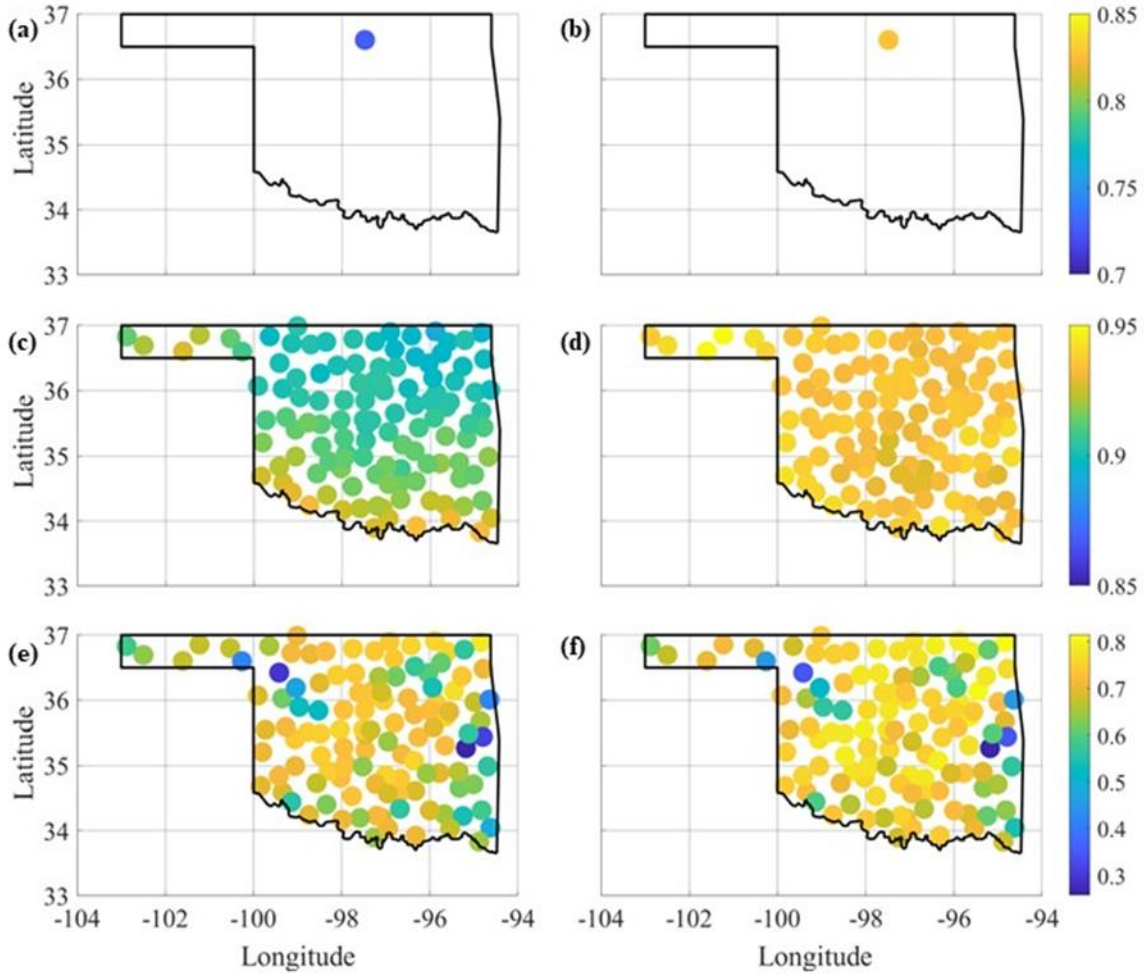

Figure 10. Maps of correlation coefficients of NLDAS-2 variables at their original resolution (left panels) and at the downscaled 500 m resolution against the corresponding Mesonet ground observations for (a, b) downward longwave radiation, (c, d) downward shortwave radiation, and (e, f) 10m wind speed.

For wind speed, we do not notice a clear improvement in the downscaled product when analyzing the scatterplots in Figures 8 e-f. Similarly, the correlation values presented in Figures 10 e-f are very similar to each other at most of the Mesonet stations.

Table 1: Statistical metrics for each NLDAS-2 atmospheric forcing at their original and at downscaled resolution (in italic) with respect to ground observations.

Native Resolution/Downscaled	**Correlation**	**Bias**	**RMSE**	**NSE**	**Sample Size**
Temperature (K)	0.94/0.97	3.94/1.95	5.09/2.69	0.74/0.93	1,028,409
Pressure (hPa)	0.94/1.00	11.6/1.52	13.8/2.75	0.56/0.99	1,030,356
Relative Humidity (%)	0.69/0.92	13.2/9.71	17.1/12.6	0.25/0.53	1,028,328
Longwave Radiation (W/m2)	0.73/0.82	37.7/28.1	49.4/36.9	0.36/0.67	8,627
Shortwave Radiation (W/m2)	0.89/0.93	63.7/52.3	119/101	0.81/0.86	1,029,358
Wind Speed (m/s)	0.67/0.70	2.62/1.75	3.04/2.19	-0.27/-0.24	1,020,335

The overall statistics presented in Table 1 confirm the improved performance of the downscaled variables with respect to the original NLDAS-2, in terms of additive bias, RMSE, and NSE, computed against the hourly ground observations recorded at 118 stations (and for the ARM site for longwave radiation). RMSE of pressure is reduced from 13.8hPa for the original resolution NLDAS-2 to 2.75hPa for the downscaled product. Moreover, for the downscaled pressure, NSE gets close to 1, showing an almost perfect match with the corresponding ground observations. For longwave radiation, both bias (from 13.18 to 9.63W/m^2) and RMSE (from 17.1 to 12.6W/m^2) are lower thanks to

the downscaling procedure. For shortwave radiation, bias and RMSE reduce from 63.73 to 52.30W/m^2 and from 119 to 100.68W/m^2, respectively, while the correlation coefficient and the NSE increase from 0.89 to 0.93 and from 0.81 to 0.86, respectively. For wind speed, the overall statistical analysis shows slightly improved results in terms of RMSE for the downscaled product (2.19m/s) with respect to the original resolution NLDAS-2 wind speed (3.04m/s).

1.5 **Conclusions**

This study presented a downscaling methodology for six atmospheric variables that are commonly used as land surface model inputs. The proposed downscaling framework comprises a statistical interpolation and a set of deterministic physical rules and is tested on the NLDAS-2 dataset across Oklahoma during 2015. The novelties introduced in this study with respect to previous work can be summarized as follows. First, we adopt a dynamic lapse rate of air and dew point temperature that is not constant in space and time. This step is extremely important since i) lapse rate showed high variability across Oklahoma, and ii) the dynamic lapse rate is used to downscale air temperature, which is then used to downscale air pressure and air humidity, thus, critically modifying the downscaling of all three variables. Second, the method adopted for downscaling shortwave radiation considers multiple factors, including a clearness index, local illumination, cast-shadowing, sky obstruction, and topographic configuration. Third, spatially downscaled surface roughness and zero-plane displacement height are introduced in the downscaling technique adopted for wind speed. In summary, this work combines different

methodologies that have been previously proposed and modifies some of them to account for more factors.

To assess the improvement with respect to the original resolution dataset (at 12.5 km resolution), the 500 m downscaled products were compared to high-quality, high-resolution ground observations. The goal of this work is to prove the viability of the proposed methodology where high-quality, high-resolution continuous ground-based data are available, and the Oklahoma Mesonet network offers such dataset.

For all the atmospheric variables, an improvement was observed in terms of several performance metrics (bias, correlation coefficient, RMSE, and NSE). Except for wind speed, the improvement in the other five downscaled products with respect to the original NLDAS-2 dataset was found substantial in terms of RMSE, correlation, and NSE. Thanks to the adopted dynamic lapse-rate correction, the downscaled temperature and pressure matched almost perfectly with the ground observations. Although the improvements in specific humidity, longwave radiation, and shortwave radiation were not as obvious, the statistical analysis revealed the pointedly better performance of the downscaled products with respect to their native resolution version. Overall, the downscaled atmospheric products captured more spatial variability compared to the original NLDAS-2.

During the past few decades, land surface models have become popular for simulating surface energy and water fluxes in response to near-surface atmospheric forcing. The downscaled near-surface atmospheric forcing dataset developed in this study would facilitate the transition of these models to hyper-resolution and improve our ability to monitor and predict the Earth's terrestrial water, energy, and biogeochemical cycles at

scales that are relevant for decision making. Nevertheless, this study only tested the proposed methodology over Oklahoma and during one year. More validation exercises using quality-controlled observations and longer time series should be conducted to understand the potential of such an approach in different regions of the world. Although the downscaling approach here was applied only to the NLDAS-2 dataset, the proposed methodology could be applied to any other dataset. The methodology proposed in this study can be replicated in other regions of the world, by applying the downscaling technique to global datasets, such as the Modern-Era Retrospective Analysis for Research and Applications Version 2 (MERRA-2; Rienecker et al. 2011, Gelaro et al. 2017) and the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis 5th Generation (ERA-5; Dee et al. 2011), as presented in Mei et al. (2018). Moreover, the downscaling spatial resolution is limited to 500 m for this work. This method can potentially be applied to finer resolutions if all parameters used in the proposed framework are made available at higher resolutions in the future.

2. TOWARDS HYPER-RESOLUTION LAND-SURFACE MODELING OF SURFACE AND ROOT ZONE SOIL MOISTURE

The goal of this chapter is to estimate surface and root zone soil moisture at resolutions that are useful for decision making and water resources management. A 500 m atmospheric forcing dataset is developed from the 12.5 km NLDAS-2 (North America Land Data Assimilation System) products across Oklahoma, where high-quality observations are available for validation purposes. A land surface model is then forced with three combinations of input variables to simulate surface and root zone soil moisture: 1) NLDAS-2 atmospheric forcings at their original resolution; 2) downscaled NLDAS-2 atmospheric variables (i.e., near-surface air temperature and humidity, wind speed and direction, incident longwave and shortwave radiation, pressure) and original resolution NLDAS-2 precipitation; and 3) downscaled NLDAS-2 atmospheric variables and precipitation. Results show that the third simulation is able to bring modeled standard-normal deviates of both surface and root zone soil moisture closer to in situ observations, whereas the second simulation only shows slight improvements with respect to one forced with original resolution NLDAS-2 data. This is particularly evident for negative values of standard-normal deviates, which correspond to drier than usual cases, due to the improved ability of the downscaled precipitation to detect missed events and no-rain cases. In summary, finer resolution forcings have the potential to improve simulations of soil moisture, and the resolution of precipitation plays a critical role in improving the time series of soil moisture standard-normal deviates.

2.1 **Introduction**

By controlling the partitioning of available energy incident on the land surface, soil moisture is a key variable in land–atmosphere interactions that impact local weather, such as cloud coverage and precipitation, and hydrological parameters, such as runoff and evapotranspiration (Seneviratne et al. 2010). Soil moisture is also involved in several feedbacks, e.g., soil moisture-temperature and soil moisture-precipitation, that may be significant not only at the local scale but also at the regional and global scale. In particular, root zone soil moisture plays a prime role in the regulation of water and energy budgets at the soil–vegetation–atmosphere interface through evaporation processes of the surface soil layer and plant transpiration (Shukla and Mintz 1982). If the initialization of root-zone soil moisture is not accurate, it may cause drifts of the temporal evolution of the surface state variables and degrade weather forecasts (Beljaars et al. 1996; Dirmeyer 2000; Koster and Suarez 2003). Thus, realistic estimates of surface and root-zone soil moisture can improve weather and climate prediction, hazard mitigation (floods and droughts), agricultural planning, and water resources management.

As the global availability of in situ high-quality/high-resolution soil moisture measurements is limited, satellite missions, such as the Soil Moisture Active Passive (SMAP; Entekhabi et al. 2010a) and the Soil Moisture Ocean Salinity (SMOS; Kerr et al. 2001), represent a valid alternative. Satellite-based instruments measure land surface brightness temperature and radar backscatter, thus providing information on surface soil moisture (top 5 cm of the soil column). The significance of these products depends critically on the resolution at which they are available, which often limits their use in

decision making. Therefore, modeling is necessary to estimate soil moisture and its variations over time, space, and with depth in the soil column.

This need for improved accuracy and resolution of soil moisture across different landscapes drove the development of hyper-resolution land surface modeling. Models at high resolution, often referred to as hyper-resolution (100 m to 1 km globally), have recently been able to provide detailed information about the storage, movement, and quality of carbon and water at and near the land surface (Wood et al. 2011). However, developing such modeling systems is challenging, as adequate model physics should be assessed, computing resources must be available, and atmospheric variables at such resolutions are needed as input to force the model (Beven et al. 2015; Garnaud et al. 2015; Singh et al. 2015; Ji et al. 2017).

Land surface models (LSMs) generally require a set of near-surface atmospheric forcing variables at every time step: air temperature, air humidity, wind speed and direction, incident longwave radiation, shortwave radiation, and precipitation. Physical, dynamic, and statistical downscaling approaches have been developed in the past to interpolate coarser resolution atmospheric variables to the required resolution (Cosgrove et al. 2003; Haylock et al. 2006; Liston and Elder 2006; Girotto et al. 2014; Sunyer et al. 2015; Gaur and Simonovic 2017). Rouf et al. (2019) recently developed a downscaling approach for a set of atmospheric variables based on correlations with landscape information, such as topography, surface roughness, and vegetation. Their proof-of-concept over Oklahoma showed that correlation coefficients between the downscaled products (at 500 m) and ground observations were consistently higher (and biases

smaller) than the ones between the native resolution data (at 12.5 km) and ground observations. Moreover, the random forest (RF) framework developed by Mei et al. (2020) for downscaling precipitation showed promises. Specifically, they downscaled the Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2) precipitation product to 1 km resolution using the RF classification and regression algorithm across High Mountain Asia. Their results suggest improvements with respect to the original resolution MERRA-2 and comparable performance with several satellite products and ground-based observations, both in terms of precipitation magnitude and variability.

This work investigates the potential of applying these downscaling techniques to a set of atmospheric forcings and subsequently force an LSM to estimate surface and root zone soil moisture at 500 m across Oklahoma. Particular attention is given to the resolution of precipitation. The reason is twofold. First, being precipitation the main driving forcing variable in land surface modeling, it controls several hydrological and biogeochemical processes, including runoff, evaporation, transpiration, groundwater recharge, and soil moisture (Hazra et al. 2019). Second, precipitation largely varies both in space and time, and its resolution may play a fundamental role in the estimation of soil moisture. Nevertheless, soil moisture temporally integrates prior precipitation and is subject to lower and upper limits, and the variability of errors in soil moisture is typically smaller than that of errors in precipitation (Maggioni et al. 2012). Hence, the impact of atmospheric forcings and their associated resolutions on soil moisture simulated by an LSM is not straightforward.

This works seeks therefore to answer the following main research questions: *do finer resolution forcing data improve estimates of surface and root-zone soil moisture simulated by an LSM? is the resolution of precipitation more important than the one of others?* In order to answer such questions, we propose three different simulations that use different combinations of LSM input data (at different resolutions), as described in the Methodology section. The coarse resolution dataset comes from Phase 2 of the North American Land Data Assimilation System (NLDAS-2; Cosgrove et al. 2003). The downscaling approaches developed by Rouf et al. (2019) and Mei et al. (2020), which consider local topography to downscale a set of atmospheric variables, are applied to the NLDAS-2 dataset.

2.2 **Methodology**

This work focuses on a domain in Oklahoma, in the Midwestern United States during 2015. This area is chosen because of the availability of a dense network of hydrometeorological stations, the Oklahoma Mesonet (Brock et al. 1994; McPherson et al. 2007; Figure 11), which is fundamental to evaluate the proposed modeling approach. Oklahoma is characterized by a gentle topography that spans from an altitude of 88 MAMSL (meters above mean sea level) in the southeastern corner to a height of 1515 MAMSL in the northwestern corner and by a continental climate with cold winters and hot summers. The western region is drier compared to the wetter eastern half, as shown in Figure 11.

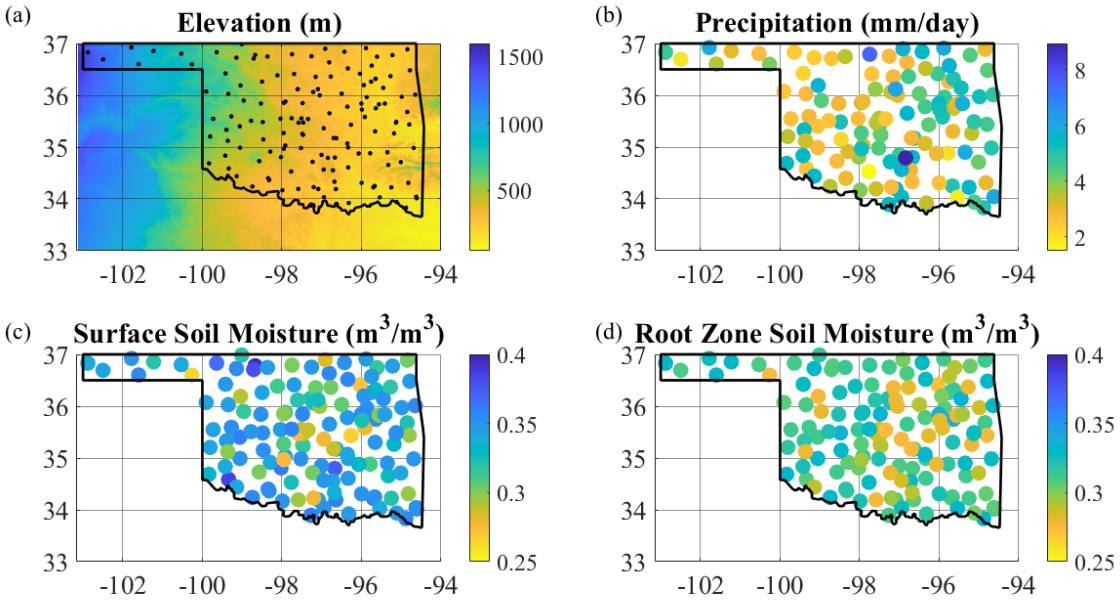

Figure 11. Maps of the study region showing (a) elevation (from the Shuttle Radar Topography Mission (SRTM; "SRTM Data – CGIAR-CSI SRTM" n.d.) and location of the Mesonet stations, (b) average precipitation in 2015 measured by the Mesonet rain gauges, and average (c) surface soil moisture and (d) root zone soil moisture observed at the Mesonet stations in 2015.

The Noah-MP is selected as the land surface model (see section 2.2.2). The meteorological forcing used for Noah-MP comes from NLDAS-2 (Cosgrove et al. 2003; Mitchell 2004). The non-precipitation land-surface forcing fields for NLDAS-2 are derived from the NCEP North American Regional Reanalysis (NARR; Mesinger et al. 2006), while NLDAS uses the Eta Data Assimilation System for precipitation (EDAS; Rogers et al. 1996). NLDAS-2 has a spatial resolution of 1/8 degree and temporal resolution hourly for the period of 1979-present.

The original NLDAS-2 variables (near-surface air temperature and humidity, wind speed and direction, incident longwave and shortwave radiation, pressure and precipitation) are downscaled to 500 m using the combination of physically-based

methods (for all variables but precipitation) and a RF-based framework (for precipitation), as described in the next section. Then, both the original resolution NLDAS-2 and the downscaled variables are used to force the Noah-MP land surface model to produce high-resolution soil moisture estimates. Three different simulations are proposed in this work, in which Noah-MP is forced with different combinations of input data:1) NLDAS-2 atmospheric variables and precipitation at their original resolution (hereinafter, simulation 1); 2) NLDAS-2 downscaled atmospheric variables (i.e., near-surface air temperature and humidity, wind speed and direction, incident longwave and shortwave radiation, pressure) and original resolution NLDAS-2 precipitation (hereinafter, simulation 2); and 3) NLDAS-2 downscaled atmospheric variables and precipitation (hereinafter, simulation 3).

2.2.1 The Downscaled Forcing Dataset

The downscaling framework proposed by Rouf et al. (2019) is based on a set of deterministic physical rules and was tested on the NLDAS-2 dataset across Oklahoma during 2015. Their methodology assumes that topography and vegetation cover have a large impact on surface atmospheric conditions, and, thus, uses high-resolution topography and vegetation information to spatially downscale 12.5 km NLDAS-2 2m air temperature, surface pressure, 2m humidity, incident longwave/shortwave radiation, and 10-m wind speed to 500 m. Moreover, this approach adopts i) a dynamic lapse rate of air and dewpoint temperature to downscale air temperature, pressure, and humidity, ii) multiple factors, including a clearness index, local illumination, cast- shadowing, sky obstruction, and topographic configuration for downscaling shortwave radiation, and iii)

a spatially downscaled surface roughness and zero-plane displacement height for downscaling wind speed.

In Rouf et al. (2019), the 500 m downscaled products were compared to high-quality, high-resolution ground observations collected by the Oklahoma Mesonet network. For all the atmospheric variables, an improvement was observed in terms of (bias, correlation coefficient, root mean square error, and Nash–Sutcliffe efficiency). Additionally, the downscaled atmospheric products were able to overall capture more spatial variability than the original NLDAS-2. Since the dataset adopted in this study is exactly the same as the one developed and validated by Rouf et al. (2019) over the same domain, no additional evaluation is proposed here and the reader is referred to their study for supplementary information and more detail on the algorithms and their performance.

The 500 m precipitation dataset is obtained through the framework developed by Mei et al. (2020). The core of this method is RF classification and regression. First, a recursive feature elimination procedure is used to select relevant predictors, which include variables representing atmospheric, geographic, and vegetation cover information (i.e., air temperature, dew point temperature, air pressure, specific humidity, relative humidity, incident longwave radiation, incident shortwave radiation, wind speed, 30-day lagged vegetation index, 60-day lagged vegetation index, distance to the closest dry grid cell, climate class, latitude, longitude, and day of year). The eight variables selected as relevant predictors are dew point temperature, air pressure, specific humidity, incident longwave radiation, incident shortwave radiation, wind speed, 60-day lagged vegetation index, and day of year. A binary precipitation mask is produced based on a daily

cumulative precipitation rate greater than 0 mm, and an RF classification model is trained to the daily precipitation mask. Then, the RF regression model is applied to provide the 500 m precipitation field using the downscaled predictors over rainy pixels only. The obtained product is validated across the study area using observations collected at the Mesonet rain gauges, as discussed in section 2.2.3.

The RF-based framework provides downscaled precipitation at 500 m/daily resolution. This downscaled precipitation is further disaggregated to hourly for the land surface model following a method similar to (López López et al. 2018). Specifically, the fraction of precipitation per hour derived from the 12.5 km NLDAS-2 dataset is used to disaggregate daily downscaled precipitation at 500 m resolution to the hourly time scale, following these steps:

(i) Precipitation at 12.5 km resolution that occurs during the i^{th} hour (Pi) is divided by the daily total precipitation to obtain the fraction of precipitation F for hour i:

$$F_i = \frac{\boldsymbol{P_i}}{\sum_{\boldsymbol{i=1}}^{\boldsymbol{24}} \boldsymbol{P_i}} \quad \textbf{(15)}$$

where 24 is the number of hours in a day. Note that F_i is defined for pixel with non-zero daily precipitation accumulation only.

(ii) The fractions of precipitation obtained for the 12.5 km grid, F_i, are bilinearly interpolated to the 500 m grid, obtaining FD_i.

(iii) Daily downscaled precipitation is multiplied by the corresponding fractions FD_i to obtain hourly precipitation at each 500 m grid cell. Daily precipitation is disaggregating evenly for pixels whose FD_i is not defined.

2.2.2 **The Noah-MP Land Surface Model Simulations**

The LSM adopted in this study is the Community Noah Land Surface Model with Multi- Parameterization Options (Noah-MP) (version 3.6; Niu et al. 2011; Yang et al. 2011). The Noah-MP is based on the original Noah LSM, with improved physical processes, such as the separation of the vegetation canopy from the ground surface and the inclusion of a multi-layer snow model. Noah-MP has been shown to successfully simulate land-atmosphere energy, water, and carbon exchanges and hydrologic states, such as surface runoff, soil moisture, snow depth, snow water equivalent, and terrestrial water storage both over the United States (e.g., Cai et al. 2014; Chen et al. 2014; Ma et al. 2017) and in other complex regions, such as High Mountain Asia (Xue et al. 2019).

Noah-MP is run within the NASA Land Information System (LIS; Peters-Lidard et al. 2007), a software for high-performance land surface modeling and data assimilation that provides a common framework capable of ensemble land surface modeling on points, regions or the globe. In order to initialize the offline LSM, a multi-year spin-up loop approach is adopted. The model is first spun up for 20 years, looping over a 10-year period from 2005 to 2014, reaching quasi-equilibrium for both surface and subsurface temperature states and then run at a 15 min time step to produce hourly output for 2015 on a regular 1 km and 12.5 km spatial grid. Model parameters (e.g., land cover, land mask, soil texture, elevation, slope, aspect, greenness data, albedo, snow albedo, bottom temperature) are used at 1 km for the high-resolution runs and averaged to the 12.5 km grid for the coarser resolution simulations. Quasi-equilibrium is reached when the

difference in soil moisture values between the last two runs was less than 0.1 %, following the approach in Rodell et al. (2005) and Cai et al. (2016).

Table 2. Noah-MP model runtime options and parameters used in this study.

Option	**Definition**
Vegetation model option	Dynamic vegetation
Canopy stomatal resistance option	Ball-Berry type (Ball et al. 1987)
Soil moisture factor for stomatal resistance option	Original Noah (Chen and Dudhia 2001)
Runoff and groundwater option	TOPMODEL with groundwater (Niu et al. 2007)
Surface layer drag coefficient option	Monin-Obukhov (Brutsaert 1982)
Supercooled liquid water option	No iteration (Niu and Yang 2006)
Frozen soil permeability option	Linear effects, more permeable (Niu and Yang 2006)
Radiation transfer option	Modified two-stream
Snow surface albedo option	CLASS (Verseghy 1991)
Rainfall and snowfall option	Jordan (Jordan 1991)
Lower boundary of soil temperature option	Noah
Snow and soil temperature time scheme	Semi-implicit
Parameter	**Value**
Number of soil layers	4
Each soil layer thickness (from top to bottom)	0.1, 0.3, 0.6, 1.0 meter
Soil color index	4: for medium dark color soil

Within LIS, the NLDAS-2 forcing fields (i.e., near surface air temperature and specific humidity, downward longwave and shortwave radiation, eastward and northward wind, surface pressure, and total precipitation) with an hourly temporal resolution and 12.5 km and 1 km spatial resolution onto the same model time step. The static input data for Noah-MP are obtained from the National Center for Atmospheric Research Application Laboratory website (https://ral.ucar.edu/solutions/products/noah-multiparameterization-land-surface-model-noah-mp-lsm), which are preprocessed onto the same model grid using the NASA Land surface Data Toolkit (LDT) public release of

version 7.2 (Arsenault et al. 2018). The Noah-MP is developed based on the original Noah land surface model, with multiple options available for surface water infiltration, runoff, groundwater transfer and storage, dynamic vegetation, canopy resistance, and frozen soil physics (Niu et al. 2011). Table 2 summarizes all Noah-MP options and parameters used in this study.

2.2.3 **Validation Dataset and Performance Metrics**

The Oklahoma Mesonet network, used as reference for validating the model estimates (both in terms of precipitation and soil moisture) is a statewide network designed and implemented by the University of Oklahoma and Oklahoma State University (Figure 11). It consists of 120 automated environmental monitoring stations that provide regular measurements of several atmospheric and hydrologic variables (Martens et al. 2017; "Mesonet | Home Page," n.d.). The measurements are taken every 5 minutes and transmitted to a central facility that performs quality control.

The total amount of rainfall is measured by tipping bucket rain gauge sensors (approximately 0.01 inch per tip, or 0.254 mm). The gauges are located 0.6-m above the ground, have a 30.5 cm diameter opening, and are surrounded by a 121 cm shield to minimize wind effects. Mesonet uses unheated gauges because of their cost and power efficiency, which may cause underestimation in case of snow or freezing rain events.

Soil moisture at each Mesonet site is measured using thermocouples at four different depths (5, 25, 60, and 75 cm). For each measurement, temperature is measured before and after a short heat pulse (21 seconds). Each sensor measurement is calibrated with its own coefficient first; then a linear regression is applied to normalize the response

and reduce sensor-specific variances. Using a set of empirical coefficients measured from different retention curves, the volumetric water content is determined from the soil matric potential, which is a measure of the capillary force needed to retain water in the soil, calculated using the calibrated temperature values (Van Genuchten 1980; Arya and Paris 1981; Illston et al. 2008).

In this work, we focus on anomaly time series of precipitation and soil moisture, i.e., standard-normal deviates, computed by subtracting the 2015 yearly mean and dividing by the corresponding standard deviation. This is because we are well aware of systematic differences between in situ observations and model estimates due to the point-scale character of the first versus the distributed nature of the latter. In the case of soil moisture, there is also a mismatch in the available measurement depths and the vertical resolution of the land surface model. Specifically, surface soil moisture in this manuscript refers to the Mesonet soil moisture measured at 5-cm depth and the 0–10-cm soil moisture output from the Noah-MP model, whereas root zone soil moisture is defined as the Mesonet observation at 60cm depth and the 40–100-cm (the mid-point of the layer is 70cm) output from Noah-MP. Standard-normal deviates capture the phase correspondence between model estimates and in situ measurements, regardless of potential mean biases or differences in dynamic range (Maggioni et al. 2011, Entekhabi et al. 2010b).

To validate the downscaled precipitation dataset against the Mesonet in situ observations, we adopt both continuous metrics like correlation coefficient (R) and root mean square error (RMSE) of time series of standard-normal deviates. Additionally, we

investigate categorical metrics like hit rate, probability of detection, threat score, probability of false detection, and false alarm rate, defined based on a contingency table (Table 3). Specifically, the hit rate (HR) measures the number of events that were correctly detected – either both products detect rain (hits) or neither does (correct no-rain) – divided by the total number of events (N). The probability of detection (POD) assesses the chances that the model estimates (in our case, either NLDAS-2 or the downscaled product) correctly detect rain (hits) when the in situ stations measure rain (hits + misses). The threat score (TS) is equal to the total number of correctly detected events (hits) divided by the total number of estimated events plus the number of misses (hits + false alarms + misses). The probability of false detection (POFD) measures the number of false alarms per the total number of non-events (false alarms + correct zeros), whereas the false alarm rate (FAR) is computed as the number of false alarms per the total number of estimated events (hits + false alarms).

For soil moisture, we analyze the time series of station-averaged standard-normal deviates of surface and root zone soil moisture simulated by Noah-MP for the three experiments described above and the corresponding Mesonet observations to study the temporal correspondence between model prediction and in situ measurements. Similar to precipitation, soil moisture standard-normal deviate time series is computed by subtracting the 2015 yearly mean and dividing by the corresponding standard deviation. Next, the cumulative distribution function (CDF) of each dataset is computed to investigate whether the model performs particularly well (or poorly) for certain value ranges of surface/root zone soil moisture. We also analyze the scatterplots of simulated

soil moisture deviates versus their corresponding reference values to visually assess the proximity of these two datasets for each proposed simulation. As the last step, we summarize what observed through the time series, CDFs, and scatterplots in boxplots of correlation coefficients and RMSEs computed between the model estimates and the Mesonet observations.

2.3 **Results**

2.3.1 **Precipitation**

In order to validate the downscaled precipitation product, we investigate maps of precipitation standard-normal deviates and compute correlation coefficients and root mean square errors between the original resolution NLDAS-2 (and downscaled product) deviates and the corresponding Mesonet deviates at each site across the study region (Figure 12). The annual average for precipitation (mm/day) over Oklahoma during 2015 is presented in Figure 12(a) at the original NLDAS-2 resolution and in Figure 12(b) at the downscaled 500 m resolution. The overall statistics, shown on the plots to summarize the product's performance, prove that R does not change and RMSE decreases thanks to downscaling. In general, R and RMSE only slightly improve the precipitation anomalies after downscaling. Some locations show poor performance in the original NLDAS-2 dataset, and the downscaling algorithm only marginally improves some of them, e.g., the yellow (dark blue) dot located in the panhandle in panel 12c (12e) turns into brownish (light blue) in panel 12d (12f), demonstrating a slightly higher correlation and smaller root mean square error. A larger improvement is noticed in RMSE compared to R at several stations.

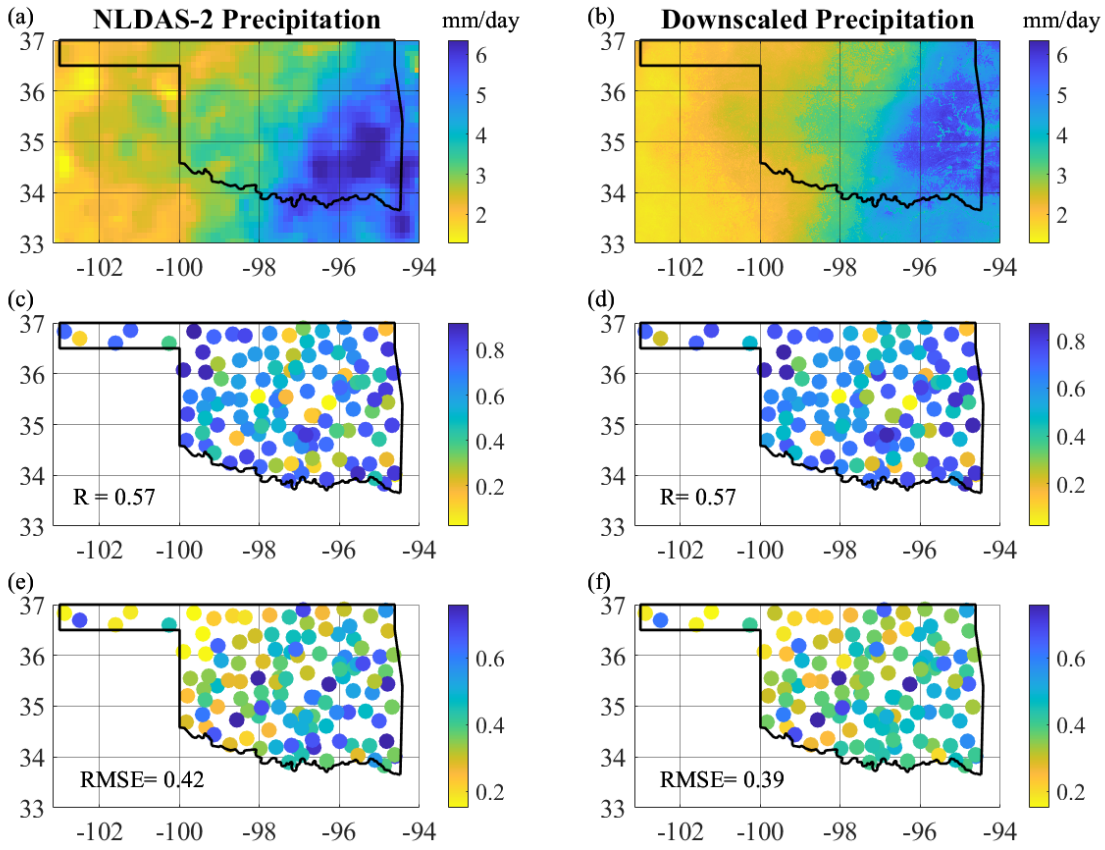

Figure 12. Maps of standard-normal deviates of the original resolution NLDAS-2 (a) and downscaled product (b) averaged during 2015, correlation coefficients (c, d) and root mean square errors (e, f) between Mesonet and NLDAS-2 at the original resolution (left column) and downscaled (right column).

Although the improvement in the statistical metrics is not substantial, the level of detail that the 500 m map (1a) presents compared to the native NLDAS-2 resolution one (1b) is certainly noteworthy. Figure 12 also demonstrates that there is no systematic spatial bias introduced by the downscaling algorithm since the improvement in the RMSE is evident at most Mesonet stations, whereas R values are consistently similar to the ones observed in the original dataset.

Table 3. Contingency table to compute categorical statistics for the original NLDAS-2 precipitation dataset and the downscaled products against Mesonet observations. The rain/no-rain threshold (*th*) is set to 0.025 cm/day (which corresponds to 0.01 inch/day).

N = 43,628		**Observation**	
		> th	≤ th
Estimate (NLDAS-2/ *Downscaled*)	> th	Hits 10,513/*10,380*	False alarms 2,564/*3,697*
	≤ th	Misses 5,133/*3,585*	Correct zeros 25,418/*25,966*

As shown by the contingency table in Table 3 and the categorical statistics in Table 4, the downscaled precipitation product is able to improve (decrease) the number of misses, and therefore correctly detect more no-rain events. This results in a slightly larger HR and TS (i.e., 1% higher) and in an enhanced POD (from 72% to 79%). The downside is an increased false alarm rate (from 18% to 23%) and a slightly higher probability of false detection (i.e., 1% higher), although still below 10%. False alarms are particularly critical in flood prediction, but in agriculture management, missed events are also essential, for instance, to avoid irrigating crops where and when there is no need.

Table 4. Categorical statistics of the original NLDAS-2 precipitation dataset and the downscaled products against Mesonet observations computed during 2015 across all stations and corresponding pixels.

Categorical Statistic	**Definition**	**NLDAS-2/ Downscaled**
Hit Rate (HR)	(Hits + Correct zeros)/N	0.86/*0.87*
Threat Score (TS)	Hits/ (Hits +False alarms + Misses)	0.63/*0.64*
Probability of Detection (POD)	Hits/ (Hits + Misses)	0.72/*0.79*
Probability of False Detection (POFD)	False alarms/ (False alarms + Correct zeros)	0.07/*0.09*
False Alarm Ratio (FAR)	False alarms/ (False alarms + Hits)	0.18/*0.23*

2.3.2 **Soil Moisture**

Figure 13

Figure 13. Maps of average surface (left column) and root zone (right column) soil moisture during 2015 from Noah-MP simulation 1 (a, b), simulation 2 (b, d), and simulation 3 (e, f) across the study region.presents maps of 12.5 km output soil moisture obtained from the Noah-MP model forced with the native resolution NLDAS-2 atmospheric variables (simulation 1) and output at 500 m resolution from simulations 2 and 3, forced with the downscaled NLDAS-2 data. As a reminder, in simulation 2 only select atmospheric variables (i.e., near-surface air temperature and humidity, wind speed and direction, incident longwave and shortwave radiation, pressure) are downscaled and the original resolution NLDAS-2 precipitation is used, whereas in simulation 3 all NLDAS-2 atmospheric variables, including precipitation, are downscaled to 500 m. The spatial distribution of soil wetness across the study region is maintained when moving the coarser 12.5 km to the finer grid. Nevertheless, the level of detail added by the downscaling procedure is evident when comparing maps in the first row with the remainder.

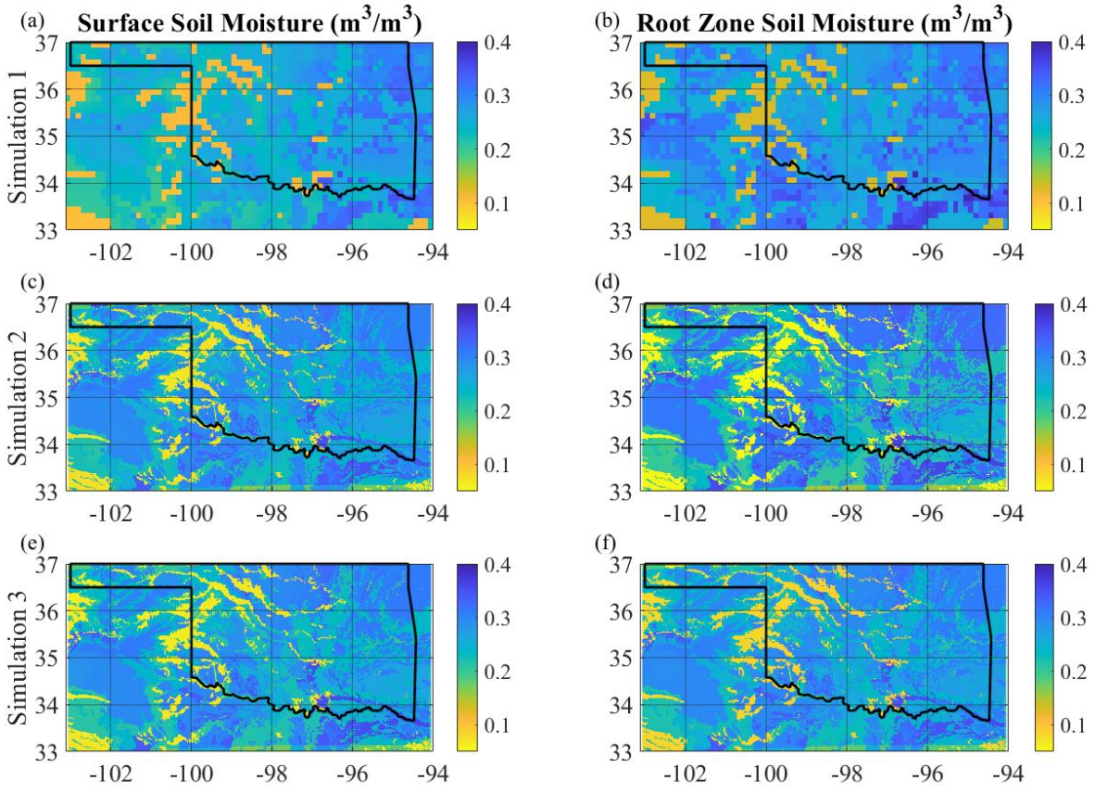

Figure 13. Maps of average surface (left column) and root zone (right column) soil moisture during 2015 from Noah-MP simulation 1 (a, b), simulation 2 (b, d), and simulation 3 (e, f) across the study region.

In order to investigate the temporal variability of soil moisture, Figure 14 presents station-average standard-normal deviate daily time series of model-predicted surface and root zone soil moisture for the three simulations and corresponding Mesonet observations during the study period. First off, standard-normal deviate time series are consistent between surface and root zone soil moisture. Second, variations of the modeled soil moisture in all three simulations are consistent with the Mesonet observations. Third, simulation 3 (which uses as input all the downscaled atmospheric variables, including precipitation) is able to bring the model closer to the in situ time series, especially when

the soil is either particularly wet (winter) or dry (summer). This is noticeable both in surface and root zone soil moisture time series.

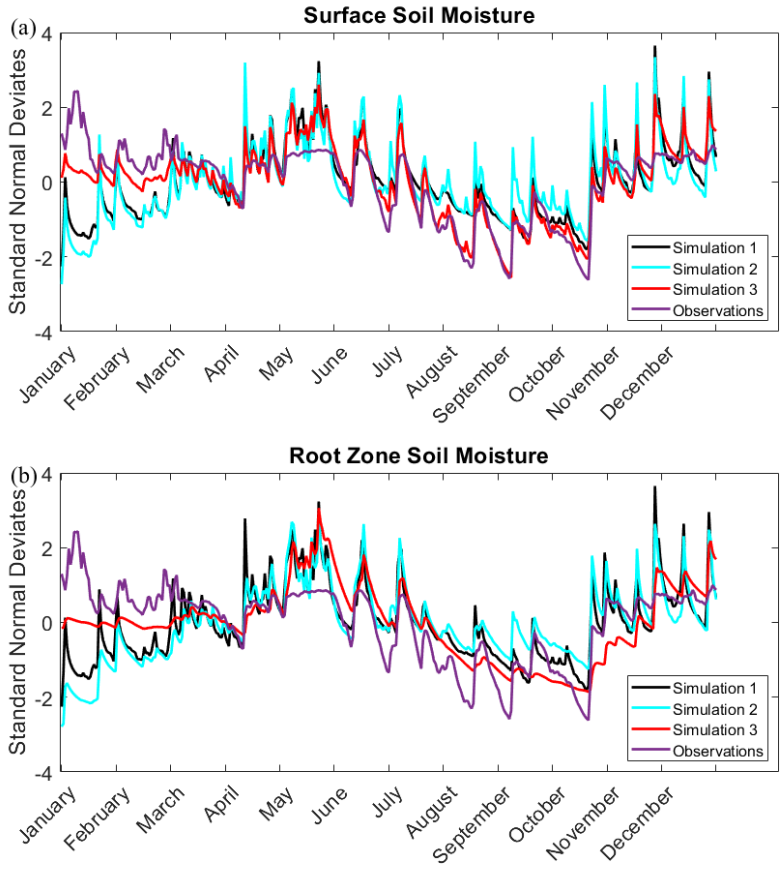

Figure 14. Time series of standard-normal deviates of (a) surface and (b) root zone soil moisture averaged across the study area, simulated by Noah-MP (3 simulations) and observed by the Masonet network during 2015

Cumulative distribution functions of the standard-normal deviates of soil moisture estimates and observations are constructed based on all model pixels where both a Mesonet measurement is available. CDFs for each simulation and for the Mesonet observations are shown in Figure 15. First off, simulations 1 and 2 are very closed to each other, as also demonstrated by the time series in Figure 14, with a slight difference for low values of standard-normal deviates (between -2.5 and -1.5) and being more

pronounced in the root zone soil moisture distributions. This stresses how the resolution in the input atmospheric variables only has a minor impact on modeled soil moisture anomalies. The story is very different when actual values of soil moisture are considered (rather than anomalies), as simulation 2 is able to remove the bias between the original resolution Noah-MP (i.e., simulation 1) and the reference dataset. However, since this comparison is not completely fair, as discussed in Section 2.3.2, such analysis is not discussed here.

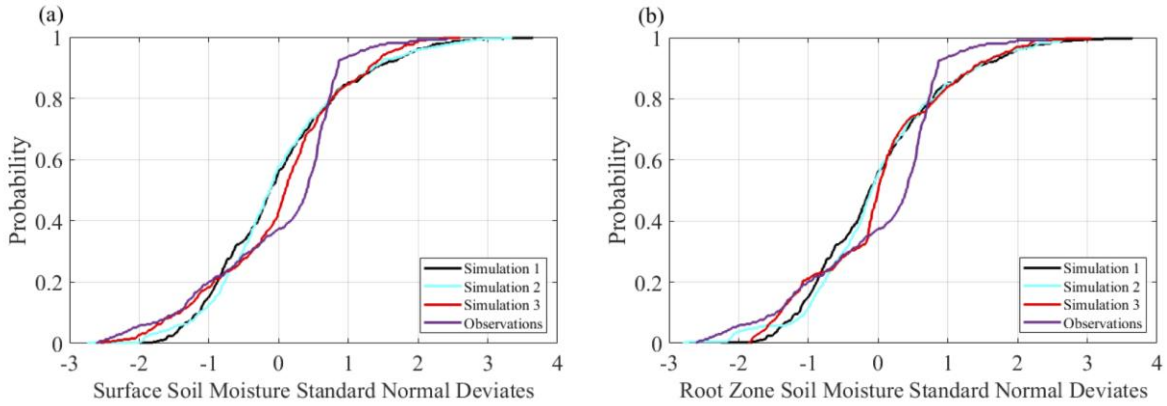

Figure 15. Cumulative distribution functions of standard-normal deviates of soil moisture from the three Noah-MP simulations and the Mesonet observations

The CDF from simulation 3 is much closer to the ground observations both in terms of surface and root zone moisture and particularly for negative standard-normal deviates, i.e. when the soil is drier than the average value. These results suggest that the higher resolution precipitation product improves the accuracy of the land surface model under dry condition, which is linked to the better performance of the downscaled

precipitation product with respect to the original resolution NLDAS-2 in terms of number of misses and correct detection of no-rain events (as shown in Table 3 and Table 4).

What previously shown in the time series and CDFs is corroborated by the density scatterplots in

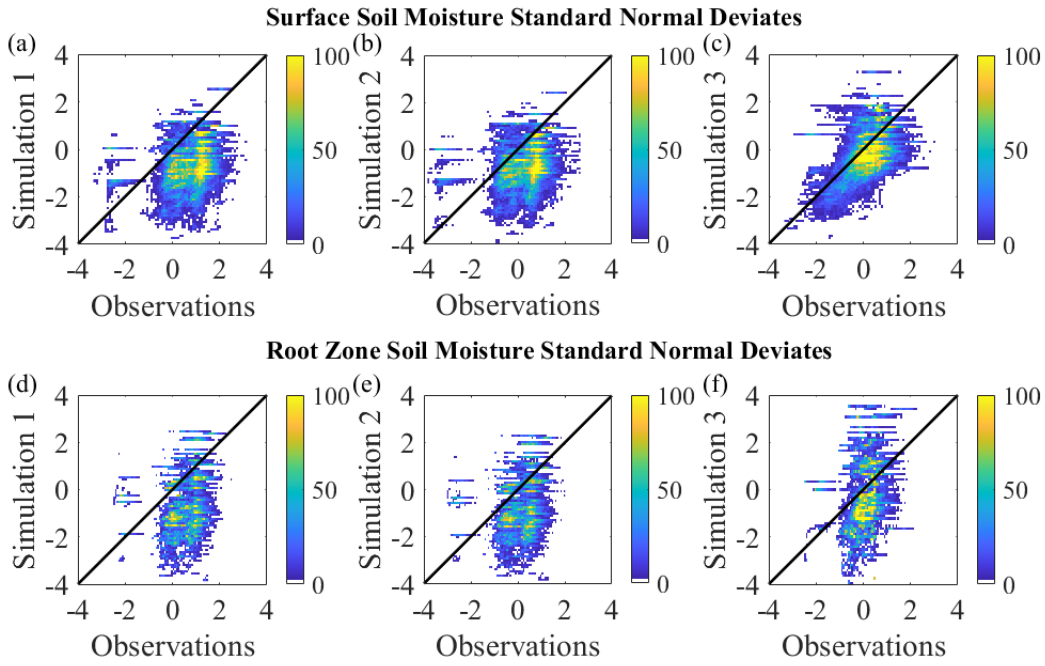

Figure 16 that present soil moisture predicted by Noah-MP in the three different simulations versus the corresponding Mesonet observations. Simulation 3 largely improves surface soil moisture estimates, moving them closer to the 1:1 line with respect to the native resolution simulation (simulation 1) and the one with only the atmospheric variables downscaled to 500 m, but not the precipitation input (simulation 2). Once again, this is particularly evident at negative values of standard-normal deviates, which are pushed closer to the reference (Mesonet) values. In the case of root zone soil moisture, the improvement is not as obvious as in surface soil moisture, which is why Rs and RMSEs are investigated next.

In order to quantify what discussed above, we analyze the correlation coefficients and RMSEs between simulated and observed standard-normal deviates of soil moisture for each station and corresponding model pixel. Boxplots in Figure 17 present the median, the 25^{th} and 75^{th} percentiles, and the minimum and maximum of such values of R and RMSE. The median value of R slightly increases throughout the different simulations for both surface and root zone soil moisture. Even more interestingly, the variability around the median (i.e., the difference between 75^{th} and 25^{th} percentile and the difference between minimum and maximum value) also shrinks thanks to the higher resolution input dataset. Once again, simulation 3 improves the performance of both simulations 1 and 2, showing the importance of precipitation resolution in modeling soil moisture.

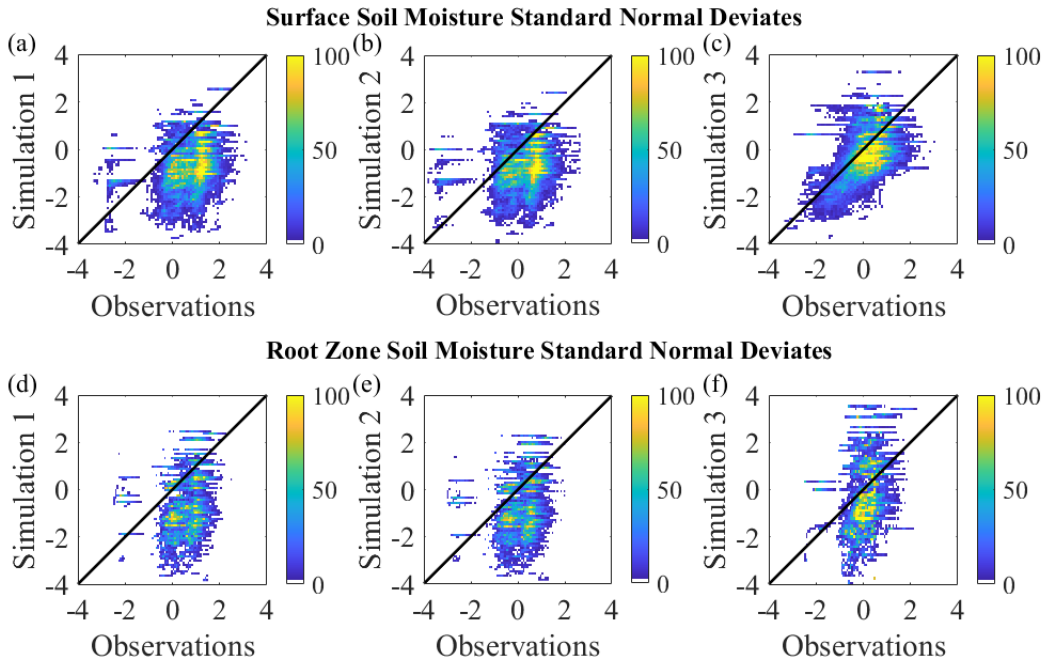

Figure 16. Density scatterplots of model simulated standard-normal deviates of surface (a, b, c) and root zone (d, e, f) soil moisture versus the corresponding Mesonet observations

Similarly, the median RMSE decreases when downscaled NLDAS-2 atmospheric variables and precipitation are used to force Noah-MP to simulate both surface and root zone soil moisture (with respect to simulation 1, in which the native resolution NLDAS-2 dataset is used). Although the difference between the RMSE in simulation 1 and 2 is minimal for surface soil moisture, it is more substantial in predictions of root zone soil moisture, showing some impact due to the change in resolution of most atmospheric variables (excluding precipitation). The variability around the median RMSE is noticeably narrower for surface and root zone soil moisture in simulation 3 with respect to the corresponding simulations 1.

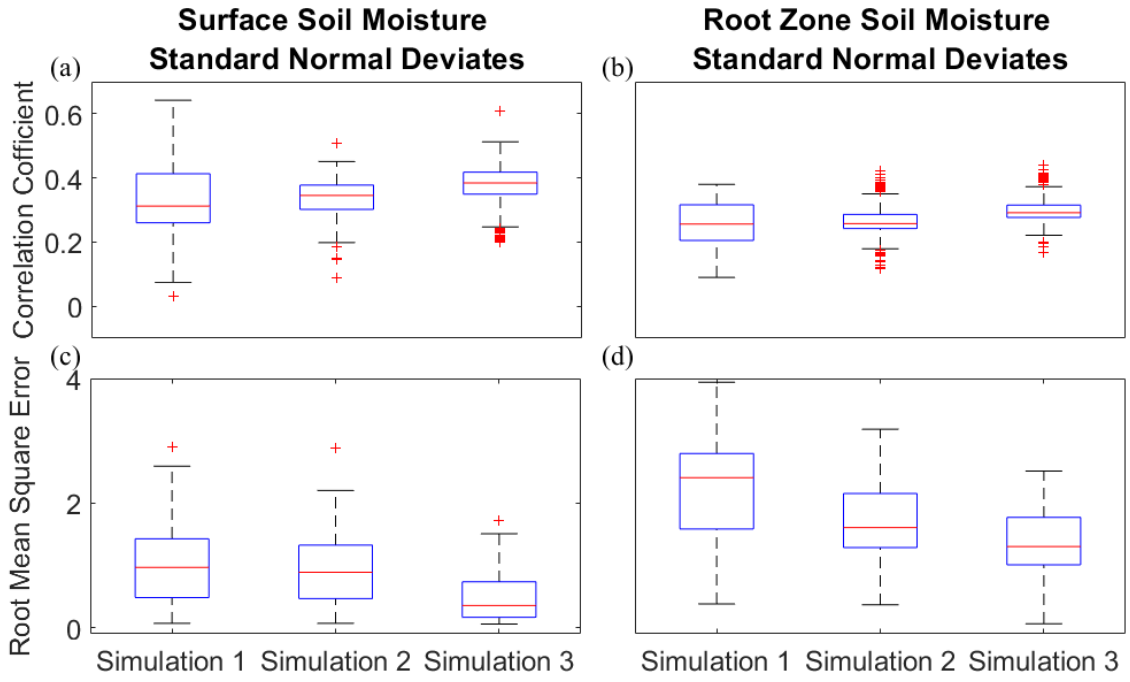

Figure 17. Boxplots of correlation coefficients (a, b) and RMSEs (c, d) of standard-normal deviates of surface (a, c) and root zone (b, d) soil moisture. In each box, the central mark indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively, the whiskers extend to the most extreme points not considered outliers, which are plotted individually as red crosses.

2.4 **Conclusions**

One of the top priorities and fundamental input for hydrologic process models is surface weather at the land-atmosphere boundary. As hydrologic models move toward higher spatial and temporal resolution and larger spatial domains, this work contributes an assessment of a land surface model forced with 500 m atmospheric variables that simulates surface and root zone soil moisture.

The proposed downscaled precipitation product is shown to only marginally improve the original resolution NLDAS-2 dataset in terms of correlation coefficients and root mean square errors of standard-normal deviates of precipitation with respect to the

Oklahoma Mesonet observations. In terms of categorical statistics, the downscaled product is able to improve the number of missed events and correctly detect more no-rain events, but showed larger false alarm rates and a slightly higher probability of false detection. Nevertheless, the level of detail that the 500 m precipitation field presents compared to the 12.5 km resolution is noteworthy, which is also translated into maps of soil moisture simulated by Noah-MP.

Simulation 3 (which used as input all the downscaled atmospheric variables, including precipitation) is proven to be the most successful at pushing the model closer to the Mesonet soil moisture observations, for standard-normal deviates of both surface and root zone soil moisture. Simulation 2 (forced with downscaled atmospheric variables but original resolution NLDAS-2 precipitation) only shows slight improvements with respect to Simulation 1 (forced with original resolution NLDAS-2 data), which stresses the importance of precipitation resolution on modeled soil moisture anomalies. This is more evident for negative values of standard-normal deviates, which correspond to drier than usual cases and may be related to the ability of the downscaled precipitation product to better detect missed events and improve the detection of no-rain cases.

In summary, to answer the research questions posed in the introductory section, results show that 1) finer resolution forcing data have the potential to improve estimates of surface and root-zone soil moisture simulated by Noah-MP, in terms of higher correlations and lower RMSEs in soil moisture anomalies, and 2) the resolution of input precipitation has a critical role in improving time series of both surface and root zone soil moisture standard-normal deviates.

Although land surface models can provide spatial and temporal soil moisture variability information, they are well known to introduce biases, which is the reason why biases were not discussed in this work. Nevertheless, assimilation of satellite-based soil moisture products (e.g., SMAP, SMOS) can help to correct these first-order large-scale biases. Future research work should not only evaluate the proposed methodology with different land surface models and in different regions of the world, but also investigate the potential to combine hyper-resolution land surface model estimates with satellite-based products in a land data assimilation system.

3. THE EFFICIENCY OF ASSIMILATING SATELLITE-BASED OBSERVATIONS IN A LAND SURFACE MODEL

This chapter focuses on combining modeling estimates of surface soil moisture with satellite-based observations within a land data assimilation system. Specifically, model states simulated by the Noah-MP land surface model described in the previous chapter are updated using an Ensemble Kalman Filter with products from the NASA SMAP (Soil Moisture Active Passive) satellite mission at two different resolutions (36 km and 9 km). The land surface model is run on a 12.5 km regular grid and forced with the NLDAS-2 (North America Land Data Assimilation System) dataset (also at 12.5 km resolution) to produce surface and root zone soil moisture estimates across Oklahoma during 2015. The Oklahoma Mesonet ground observations are used for validation purposes. Ground observations are compared to surface and root zone soil moisture output simulated by three different Noah-MP model runs: i) an open loop simulation (in which no satellite data are assimilated); ii) assimilation of the 36 km SMAP product; and iii) assimilation of the 9 km SMAP product. Results show that SMAP soil moisture retrievals (at both resolutions) improve the model performance (i.e., in comparison with the open loop run), particularly at higher elevations. Although root zone soil moisture is not directly assimilated (since satellite observations are limited to the top 5 cm of the soil column), the assimilation of SMAP products at the surface is transferred to lower layers by the modeled physical processes and is shown to improve root zone soil moisture estimates as well.

3.1 **Introduction**

Estimating soil moisture at high spatial and temporal resolutions is critical for accurate prediction of weather, droughts, floods, and energy exchange between the land and the atmosphere. Traditional ground-based soil-moisture measurements are discrete and often sparse, thus, not adequate to obtain holistic information regarding the wetness of soil over large regions. Satellite observations using microwave brightness temperature and radar backscatter can provide an alternative to sample soil-moisture at large scales (Jackson 1993; Njoku et al. 2003).

Several satellite mission have been dedicated to the retrieval of soil moisture in the recent past, such as the European Space Agency (ESA) Soil Moisture and Ocean Salinity (SMOS) mission (Kerr et al. 2010) and the NASA Soil Moisture Active Passive (SMAP) mission (Entekhabi et al. 2010a). However, satellite measurements are only sensitive to the top few centimeters of the soil column, which is particularly limiting for most agricultural, hydrological, and meteorological applications that require root zone information. Satellite retrievals are also limited by sensor accuracy and sampling (i.e., the coverage is often not spatially and temporally continuous), heavily depend on the parameterization of the retrieval algorithm (i.e., difficulties in defining the physical processes that relate brightness temperature to soil moisture and uncertainty in the algorithm parameters), and are affected by land cover heterogeneity within the pixel. Furthermore, satellite products are limited in areas where the fraction of open water is significant and where the soil is frozen or densely vegetated.

Land surface modeling is a valid alternative to provide soil moisture estimates that are continuous, at high spatial resolution, and that varies with depth in the soil column. The key challenges in land surface modeling are the susceptibility to errors in the model forcing, structure, and parameterization. It has been long argued that the observational and modeling limitation can be partially resolved by Land Data Assimilation Systems (LDAS) that combine the soil moisture observations (either in situ or satellite-based) with the state of a land surface model to maximize the spatio-temporal coverage, consistency, and accuracy of such simulations (Maggioni and Houser 2017; Reichle 2008; Reichle and Koster 2005).

Several past studies have investigated the efficiency of LDASs for soil moisture estimation (Crow and Van Loon 2006). Reichle et al. (2005) have assimilated global soil moisture retrievals from the Scanning Multichannel Microwave Radiometer (SMMR) into the NASA Catchment land surface model and demonstrated that the assimilation of SMMR data yields improved soil moisture estimates over those obtained from the model or the satellite data alone (Reichle and Koster 2005). Maggioni et al. (2012) have compared the efficiency of assimilating near-surface soil moisture from Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) using two different satellite rainfall error models (a complex multidimensional satellite rainfall error model and the simpler model which is commonly used in the NASA Goddard Earth Observing System Model) and found that the LDAS soil moisture estimate improve over the satellite retrievals and the open loop (no assimilation) land surface model estimates (Maggioni et al. 2012).

Among the several techniques available, the Ensemble Kalman Filter (EnKF) is thought to be an ideal sequential data assimilation method in hydrology and has been successfully applied to meteorological and oceanographic problems of moderate complexity in small-to medium size domains (Evensen and van Leeuwen 1996; Houtekamer and Mitchell 1998; Keppenne 2000; Lermusiaux 1999; Madsen and Canizares 1999). The EnKF is based on Monte Carlo runs that represent the forecast uncertainty, which is obtained by perturbing the model forcing and state variables to obtain an ensemble of state fields. The EnKF is very flexible at treating errors in model equations and parameters and is particularly suitable for nonlinear problems, such as soil dynamics (Durand and Margulis 2008; S. Kumar et al. 2008; Pan and Wood 2006).

For instance, Evensen et al. (1996) have used an EnKF to assimilate Geosat altimeter data in a two-layer ocean model to estimate the ring-shedding process in the Agulhas current (Evensen and van Leeuwen, 1996). Madson et al. (1999) have compared the efficiency of an extended and an ensemble Kalman filter for data assimilation in coastal areas and proven the EnKF superiority in highly non-linear dynamic problems (Madsen and Canizares, 1999). The performance of EnKFs for soil moisture estimation has been assessed by assimilating the L-band (1.4 GHz) microwave radio-brightness observations into a land surface model and EnKF was proven to be a flexible and robust data assimilation technique even at moderate ensemble sizes (Reichle et al. 2002a).

In this work, we examine the efficiency of using an EnKF within the Noah-MP land surface model (presented extensively in the previous chapter) for assimilating satellite-based soil moisture products. The methodology section describes the

experimental setup and the satellite products adopted in this study. The efficiency of the data assimilation system is assessed in the Results section by comparisons with the corresponding open-loop (i.e., no data assimilation) simulations, using the Mesonet ground observations as reference. The main conclusions are drawn in the Conclusion section, which also summarizes the main limitations of the proposed experiment and future research directions.

3.2 **Methodology**

This work focuses on the same domain described in the two previous chapters, Oklahoma and in the same time period (the year of 2015). In this chapter, satellite soil moisture data are incorporated in the Noah-MP land surface model using an EnKF to simulate surface and root zone soil moisture. This section describes in detail the experimental setup, including a discussion of the SMAP products assimilated within Noah-MP and the data assimilation system.

3.2.1 **Soil Moisture Active Passive (SMAP) Products**

The NASA Soil Moisture Active Passive (SMAP) mission was launched on 31st January 2015, with the goal of measuring land surface brightness temperature and radar backscatter (but the radar failed in July 2015) and provide information on surface soil moisture (top 5 cm of the soil column; Entekhabi et al. 2010). This works considers two products of the SMAP suite: the 36 km SMAP Level 3 product (hereinafter SMAP-36km) and the Enhanced L3 Radiometer Global Daily 9 km (hereinafter SMAP-9km).

The SMAP Level 3 (SPL3SMP) products are based on daily passive radiometer estimates of global land surface soil moisture (nominally 5 cm) that are resampled to a global, cylindrical 36 km Equal-Area Scalable Earth Grid, Version 2.0 (EASE-Grid 2.0; O'Neill et al. 2016). The SPL3SMP brightness temperature dataset is adjusted for the presence of water bodies. The brightness temperature is calibrated at the surface level, i.e., sky radiation and atmosphere contributions are corrected with auxiliary near surface data (De Lannoy et al. 2015). For this study, the most recent version 5 of SPL3SMP is used. In this version of the product, regions with permanent snow and ice, frozen ground, excessive static or transient open water in the cell, excessive radio-frequency interference (RFI) in the sensor data, and heavy vegetation (vegetation water content > 4.5 kg m^{-2}) are masked out using a passive freeze–thaw retrieval based on the normalized polarization ratio (NPR). Given the 1000 km swath and 98.5 min orbit, the SPL3SMP retrievals are spatially and temporally discontinuous, with 2–3-day gaps depending on location.

The recently released SMAP Enhanced L3 Radiometer Global Daily 9 km EASE-Grid Soil Moisture, Version 2 (SPL3SMP_E, Version 2, ONeill 2018) is also selected in this study. This SMAP L3 product is a daily global product that provides volumetric surface SM (0–5 cm, m^3/m^3). Corrected brightness temperatures are used in passive soil moisture retrieval and observed by the SMAP L-band radiometer (1.41 GHz) on the 9 km global cylindrical Equal-Area Scalable Earth (EASE) Grid 2.0 (Brodzik et al. 2012). This enhanced L3 product is a daily composite of SMAP enhanced L2 half-orbit products, where the L3 ascending and descending products are derived separately by only

considering the enhanced L2 SM products acquired (Chan 2016; ONeill 2018). The SMAP enhanced L2 SM product is derived from the SMAP Enhanced L1 Gridded Brightness Temperature Product (L1CTB_E) (posted at the 9 km grid cell) based on the Backus-Gilbert optimal interpolation technique (Chan et al. 2018; Colliander et al. 2018; ONeill 2018). An improved depth correction for effective soil temperature is used in this version of soil moisture retrievals to reduce dry biases seen when comparing SMAP products to in situ data at the SMAP core validation sites.

3.2.2 **The Land Data Assimilation System**

The land surface model adopted in this study is the Community Noah Land Surface Model with Multi- Parameterization Options (Noah-MP version 3.6; Niu et al. 2011; Yang et al. 2011), extensively described in Chapter 2. Noah-MP is run within the NASA Land Information System (LIS; Peters-Lidard et al. 2007), which is particularly suitable for high performance land surface modeling and ensemble data assimilation. Noah-MP is forced with the NLDAS-2 atmospheric fields (i.e., near surface air temperature and specific humidity, downward longwave and shortwave radiation, eastward and northward wind, surface pressure, and total precipitation) at hourly temporal resolution and 12.5 km spatial resolution. The static input data for Noah-MP are obtained from the National Center for Atmospheric Research Application Laboratory website (https://ral.ucar.edu/solutions/products/noah-multiparameterization-land-surface-model-noah-mp-lsm), which are preprocessed onto the same model grid using the NASA Land surface Data Toolkit (LDT) public release of version 7.2 (Arsenault et al. 2018).

A multi-year spin-up loop approach is adopted to initialize the model (i.e., 20 years, looping over a 10-year period from 2005 to 2014). The model is run at a 15 min time step to produce hourly output during 2015 on a regular 12.5 km grid. Model parameters (e.g., land cover, land mask, soil texture, elevation, slope, aspect, greenness data, albedo, snow albedo, bottom temperature) are available at 1 km and are here averaged to the 12.5 km grid, similarly to the coarse-resolution experiment presented in Chapter 2.

Surface and root zone soil moisture are then simulated by three different Noah-MP model runs: an open loop simulation (hereinafter OL, in which no satellite data are assimilated) and two data assimilation (DA) simulations that adopt an EnKF to merge SMAP observations. To assess the impact of the resolution of the satellite retrievals in the data assimilation system, two DA experiments are conducted:

i. assimilation of the SMAP-36km product, hereinafter 36-km DA; and

ii. assimilation of the SMAP-9km product, hereinafter 9-km DA.

The assimilation of SMAP retrievals is implemented using the EnKF (Evensen 2003; Reichle et al. 2002b) assimilation system included in LIS (Kumar et al. 2008). The Kalman filter combines the Noah-MP forecast (background) with the SMAP soil moisture observations to generate an improved estimate of the modeled soil moisture. In an EnKF, an ensemble of model runs is used to represent the model state and its associated uncertainty. In the EnKF adopted in this study, model predictions of surface soil moisture are corrected with a stochastic filtering technique towards the SMAP observations, by accounting for the relative observation and model uncertainties. Based

on the previous studies (Kumar et al. 2008; Maggioni et al. 2012; Reichle et al. 2002a; Yin et al. 2015), an ensemble size of 24 members is chosen for all the DA simulations.

A major assumption of an EnKF is that the observations and model simulations have to be Gaussian distributed. In the DA experiments, the NLDAS-2 forcing inputs such as shortwave/longwave radiations and precipitation are perturbed hourly, similar to previous work (Kumar et al. 2019, 2018, 2014). Multiplicative perturbations are applied to shortwave radiation and precipitation with a mean of 1 and standard deviations of 0.3 and 0.5, respectively, whereas, longwave radiation is perturbed via an additive error model with a standard deviation of 30 W/m^2. Moreover, cross-correlations between the three meteorological forcing variables are imposed as follows: cross-correlation between shortwave radiation and precipitation is -0.8, cross-correlation between longwave radiation and precipitation is 0.5; and cross-correlation between shortwave and longwave radiations is -0.5. The SMAP soil moisture observations are perturbed via an additive model with a standard deviation of 0.01 cm^3/cm^3 (Kumar et al. 2012).

Another major assumption that implement an EnKF is that the observations and model estimates are not biased from each other. To ensure unbiasedness, two different scaling methods are initially applied to the SMAP products: a yearly cumulative distribution function (CDF) matching method and a monthly CDF matching method (following Reichle and Koster 2004). Recently published work has shown that monthly CDF matching procedures are preferred when applied to SMAP observations in an LDAS (Yin and Zhan 2018). Thus, a comparison between the two is assessed in the Results section.

At each assimilation step, the new state of each ensemble member is computed as a weighted average of the observations (satellite soil moisture retrieval) and the model background. Weighting is determined by the relative size of the background and observation errors, with the background error governed by the ensemble covariance. The observation operator is simply a rescaling of the top layer of the soil moisture state (0–10 cm in Noah). Thus, only the top layer is directly adjusted, but deeper layers of soil moisture can slightly change in the DA step due to background covariances. A more significant impact (over time) on the deeper layers can occur resulting from the modeled physical processes of drainage and diffusion (Sabater et al. 2007).

3.2.3 **Validation**

As in the previous study presented in Chapter 2, the Oklahoma Mesonet network is used as reference for validating the model soil moisture estimates (Figure 11). The network consists of 120 monitoring stations that measure a suite of atmospheric and hydrologic variables every 5 minutes. Soil moisture data are provided at four different depths (5, 25, 60, and 75 cm) and are quality controlled. Ground observations are compared to soil moisture output from the three simulations described in the methodology (OL, 36-km DA, and 9-km DA). Surface soil moisture refers to the Mesonet soil moisture measured at 5 cm depth and the 0–10 cm soil moisture output from Noah-MP. Root zone soil moisture is defined as the Mesonet observation at 60 cm depth and the 40–100 cm (the mid-point of the layer is 70 cm) model output.

To validate the efficiency of assimilating SMAP in Noah-MP, we focus on time series of standard-normal deviates of surface and root zone soil moisture, defined by subtracting the 2015 yearly mean and dividing by the corresponding standard deviation. As discussed in the previous chapter, this is due to several reasons: i) there are systematic differences between in situ observations, satellite products, and model estimates due to the point-scale character of the first versus the distributed nature of the other due; ii) there is a mismatch in the available ground and satellite measurement depths and the vertical resolution of the land surface model; iii) the nature of the three estimates is very different (in situ thermocouple measurements for the Mesonet data, brightness temperature retrieved by a passive microwave radiometer in space, and model simulations). In addition to that, one of the inherent assumptions of the EnKF is unbiasedness in the model and observations. Biases in the SMAP products are removed via CDF matching, as mentioned above and discussed in the next section, and therefore not tested in the validation phase.

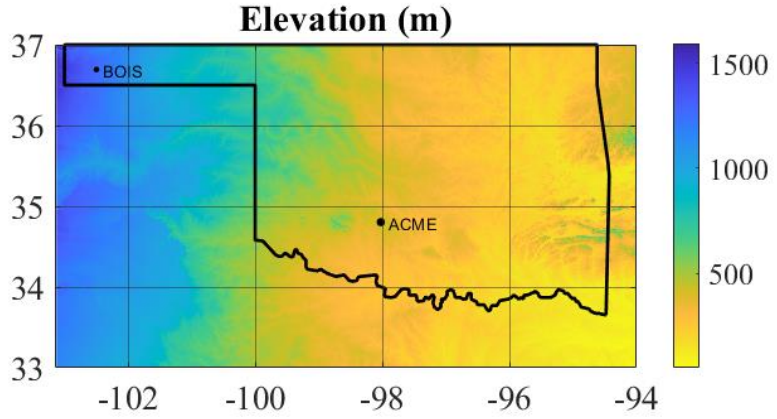

Figure 18. Domain area and location of two Mesonet stations (BOIS and ACME) used for validation purposes

Model estimates of soil moisture are evaluated against the Mesonet in situ observations, using time series of standard-normal deviates, in order to capture the phase correspondence between two estimates, regardless of potential mean biases or differences in dynamic range (Maggioni et al. 2011). Time series are presented both for domain averages and for two Mesonet stations of interest (and corresponding model pixels), one located in the panhandle, named BOIS (characterized by higher elevation) and one in the central region, named ACME (characterized by lower elevation), as shown in Figure 18. Correlation coefficients (R) and root mean square errors (RMSE) of simulated soil moisture standard deviates (for each of the three model runs) are computed against Mesonet observations and SMAP retrievals.

3.3 **Results**

First off, two different scaling methods applied to the SMAP observations prior to the assimilation are compared against each other: a yearly cumulative distribution function (CDF) matching method and a monthly CDF matching method. Figure 19 shows time series of domain averaged surface (top panels) and root zone (bottom panels) soil moisture values for the OL run and a set of DA runs that use the two different scaling methods (yearly in red and monthly in pink) applied to the SMAP-36km (left panels) and SMAP-9km (right panels) products.

A first thing to notice is that, as expected, the data assimilation of SMAP retrievals largely impacts the model free run (i.e., OL), which also confirms that SMAP adds information on the soil dynamics to the ones modeled by Noah-MP. Figure 19

demonstrates how the OL is pushed towards the SMAP products, when the latter are merged to the model forecast. As SMAP products are drier than the modeled soil moisture, the DA runs also result drier than the initial OL simulation.

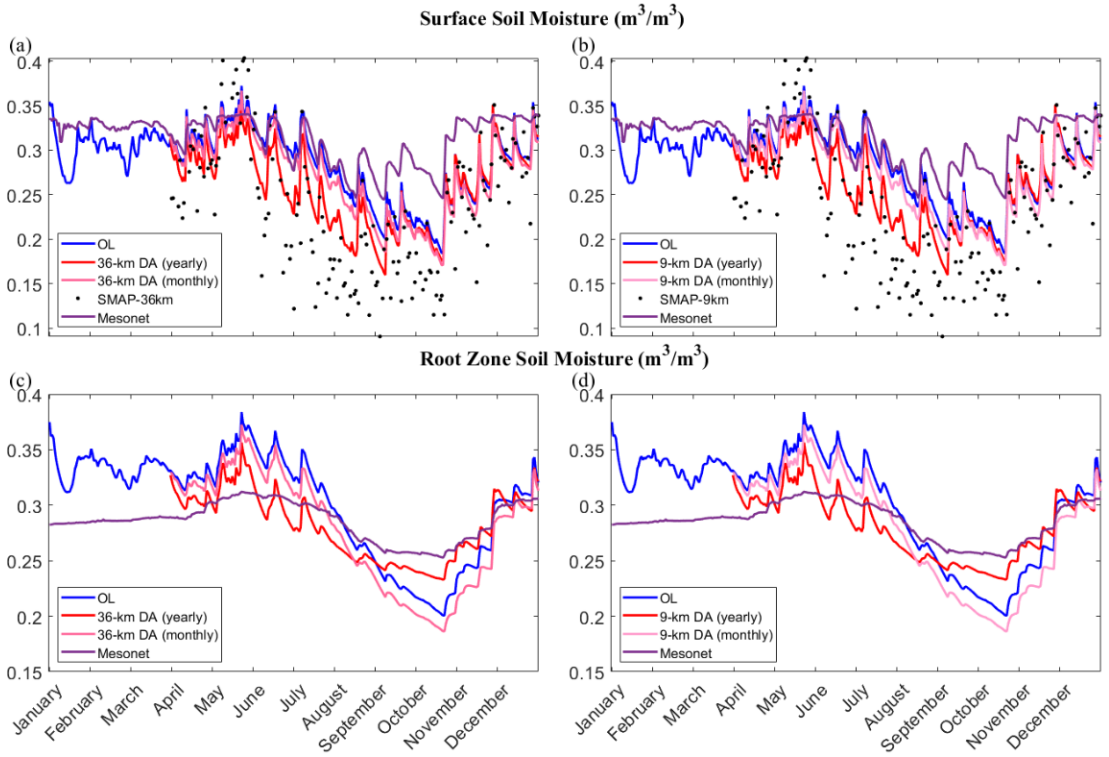

Figure 19. Daily domain averaged values of surface soil moisture (a, b) and root zone soil moisture (c, d) for i) OL, ii) Mesonet observations, iii) DA using a yearly CDF matching, and iv) DA using a monthly CDF matching. DA of both SMAP-36km (a, c) and SMAP-9km (b, d) is presented. Note that SMAP observations are only availbale after March 30th, 2015.

When comparing the performance of the yearly CDF to the monthly CDF matching approach, we observe how the first results closer to the SMAP products (and often further from the Mesonet reference observations) than the latter. This is due to the fact that biases are resolved at a temporal frequency (yearly) that is too coarse, while such

biases are seasonally dependent, and also confirms what observed in past literature (Yin and Zhan, 2018). Regardless of which SMAP product is assimilated (either SMAP-36km or SMAP-9km), the monthly CDF matching yields the best performance, pushing the OL run closer to the Mesonet ground observations, which, in this work, represent the benchmark soil moisture values.

Although root zone soil moisture is not directly assimilated, its estimate is impacted by the SMAP DA because of the physical processes in the soil column modeled by Noah-MP, as shown in the bottom plots of Figure19. Similarly, to surface soil moisture, the DA with monthly CDF matching is closer to the OL than the DA with yearly CDF matching, which in some cases brings the model also closer to the Mesonet observations, but in some instances deteriorates the model performance. This is evident in the spring and fall months, when the soil moisture is particularly wet or dry, respectively. Nevertheless, DA with monthly CDF matching appears to be superior during the summer, when the soil starts to dry up.

Table 5. Correlation coefficient, root mean square error, and mean relative error of modeled soil moisture vs Mesonet observations for a set of DA simulations that use either SMAP-36km or SMAP-9km, scaled by either monthly or yearly CDF matching

Simulation	Surface Soil Moisture			Root Zone Soil Moisture		
	R	RMSE	MRE	R	RMSE	MRE
36-km DA yearly CDF	0.60	0.10	0.28	0.70	0.05	0.15
36-km DA monthly CDF	0.66	0.07	0.21	0.71	0.06	0.20
9-km DA yearly CDF	0.60	0.10	0.28	0.70	0.05	0.15
9-km DA monthly CDF	0.66	0.07	0.21	0.71	0.06	0.20

What observed in the time series is corroborated by the statistics presented in **Error! Reference source not found.**. Specifically, correlation coefficients, RMSEs, and Mean Relative Errors (MREs) between the modeled surface soil moisture and ground-based measurements improve when a monthly CDF matching is applied compared to the yearly CDF DA simulation. For root zone soil moisture, R for both 36-km DA and 9-km DA monthly CDF marginally increases compared to the corresponding DA simulations that utilize a yearly CDF approach, whereas the overall RMSE and MRE slightly worsen. Moreover, there exists no difference between the performance of 36-km DA and 9-km DA. In conclusion, the monthly CDF matching method is chosen for all the DA simulations discussed next.

Next, time series of domain averaged standard normal deviates of surface soil moisture and root zone soil moistures are investigated (**Error! Reference source not found.**). Seasonality is evident in both surface and root zone soil moisture anomalies, with summers that are drier than the rest of the year. Summer and fall months also show larger variations with respect to winter and spring. In general, Mesonet time series present more abrupt changes, whereas model simulations (OL, 36-km DA, 9-km DA) are characterized by more gradual dry outs. Moreover, the variation in water content in the top few centimeters of the soil column (which more quickly responds to precipitation) is clearly higher than the one in the deeper soil layer.

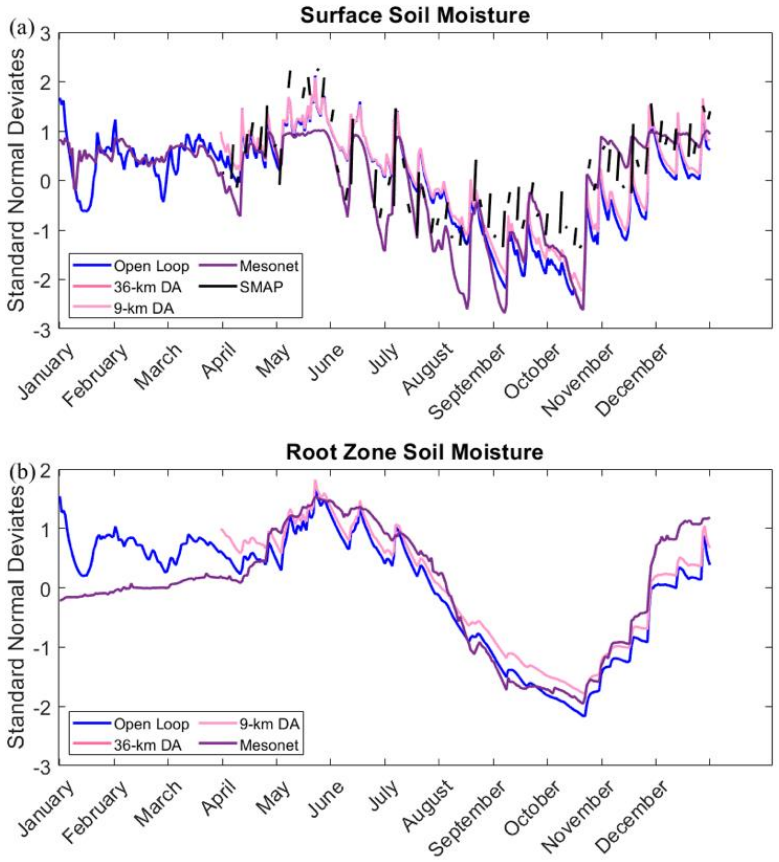

Figure 20. Timeseries of daily domain averaged standard normal deviates of surface soil moisture (a) and root zone soil moisture during 2015.

It is noteworthy how estimates of very different nature (ground measurements, satellite retrievals, and model simulations) all capture the same trend (i.e., seasonality) and variability. This is also demonstrated by the high correlation coefficients between model runs and Mesonet observations, presented in the boxplots in Figure 21 (median values are all greater than 0.6). For both surface and root zone soil moisture anomalies, the DA runs are closer to the reference Mesonet time series than the OL, showing that SMAP is able to improve the Noah-MP performance.

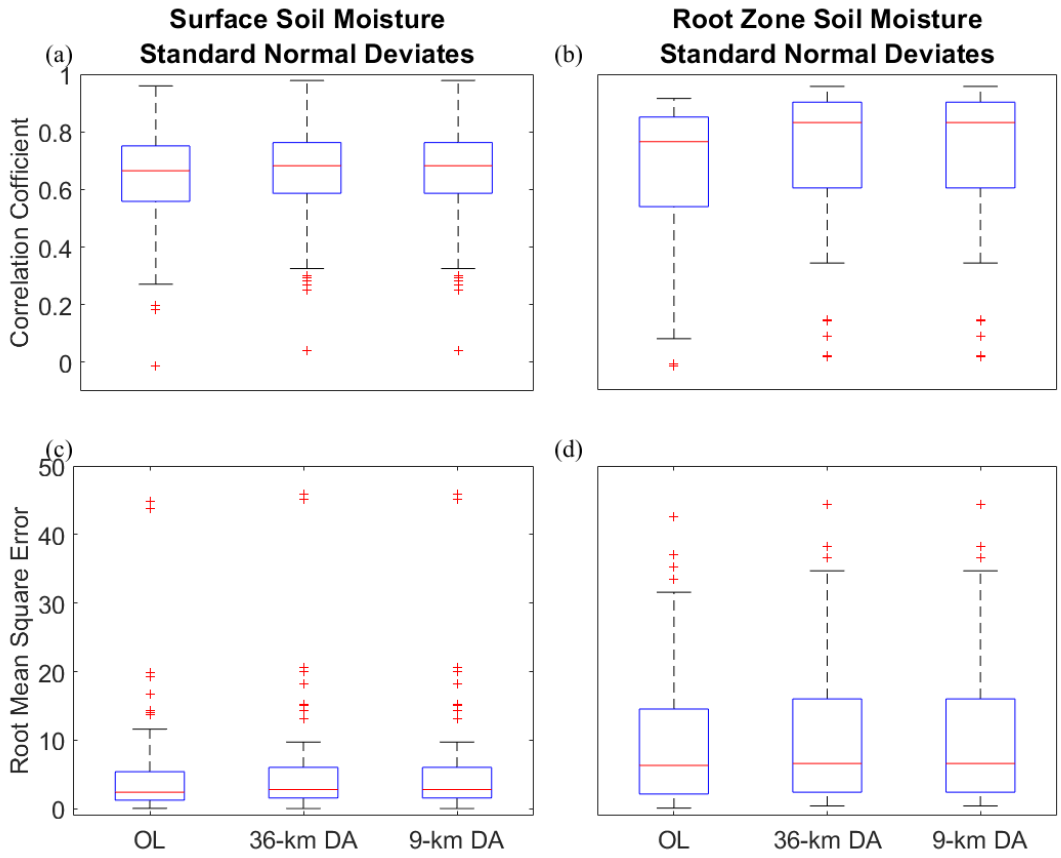

Figure 21 Boxplots of correlation coefficients (a, b) and RMSEs (c, d) of surface (a, c) and root zone (b, d) soil moisture standard-normal deviates computed at each Mesonet stations and corresponding model grid. In each box, the central mark indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively, the whiskers extend to the most extreme points not considered outliers, which are plotted individually as red crosses.

Correlation coefficients and RMSEs are computed for soil moisture standard normal deviates at each Mesonet station and corresponding model pixel. Boxplots in Figure 21 present their median, 25th and 75th percentiles, and minimum and maximum values for each model simulation (OL, 36-km DA, and 9-km DA) and for both surface and root zone soil moisture across the study area. In terms of R, not only slight improvements are observed when SMAP is assimilated, but the variability around the

median (i.e., difference between the 75th and 25th percentiles and difference between maximum and minimum values) also shrinks with respect to the OL simulation. The RMSE stays almost constant across the different simulations, with a slight decrease in the maximum value in the two DA runs with respect to the OL one for surface soil moisture.

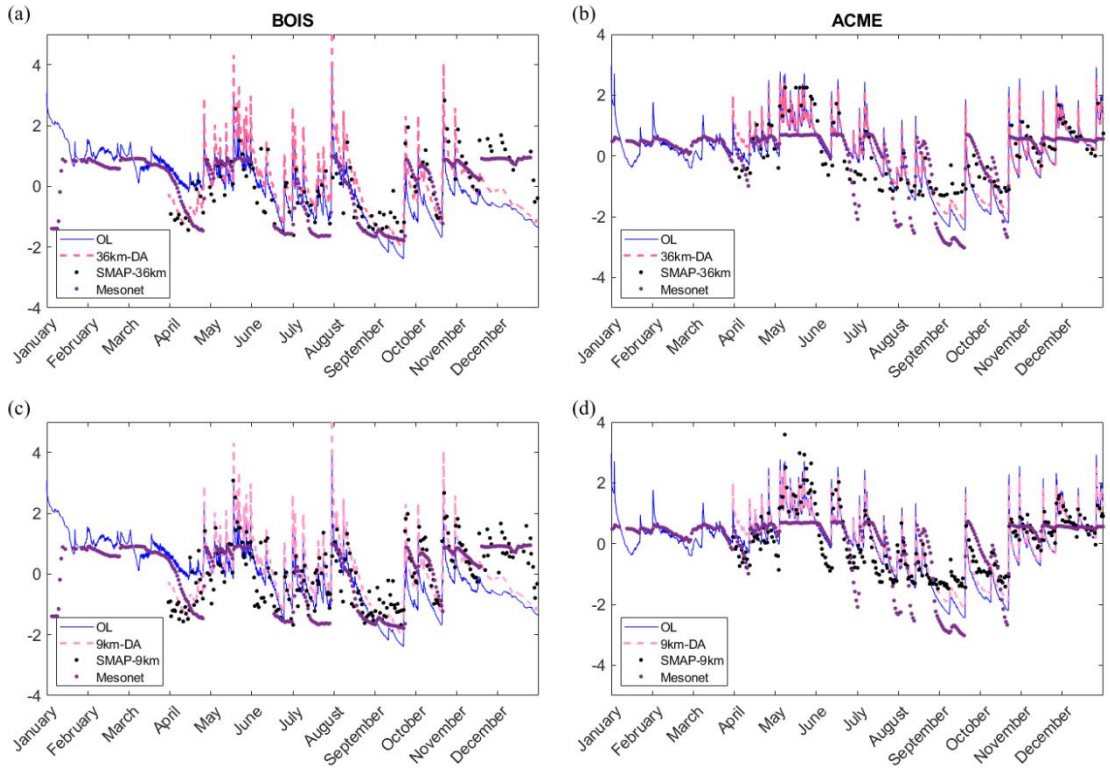

Figure 22 Timeseries of surface soil moisture standard normal deviates at two Mesonet stations, BOIS (a, c) and ACME (b, d).

In the case of root zone soil moisture, the median RMSE of the OL run is similar to the one of the two DA experiments that show an increase in the RMSE 75th percentile and maximum value. As mentioned in the comparison between the two CDF matching approaches, there exists no difference between the two DA simulations, showing that the

assimilation of the coarser SMAP products yields the same performance as the enhanced 9 km product.

In order to further investigate the performance of SMAP DA across the study area, **Error! Reference source not found.** focuses on time series of standard normal deviates of surface soil moisture at two Mesonet stations (and corresponding model pixels) of interest, one characterized by relatively high elevation (BOIS) and the other by lower elevation (ACME). A good agreement between the OL run and the Mesonet ground observations is observed at both locations. Nevertheless, SMAP data assimilation improves the correlation between model estimates and ground observations by 33% (from 0.51 in the OL to 0.68 in both 36-km DA and 9-km DA) at the BOIS station, as shown in Table 6. At the same location, the RMSE also decreases by 18% (from 1.04 in the OL to 0.85 in the DA runs). However, at the lower elevation location, DA does not seem to have an impact (1% deterioration in R and RMSE) on the Noah-MP performance.

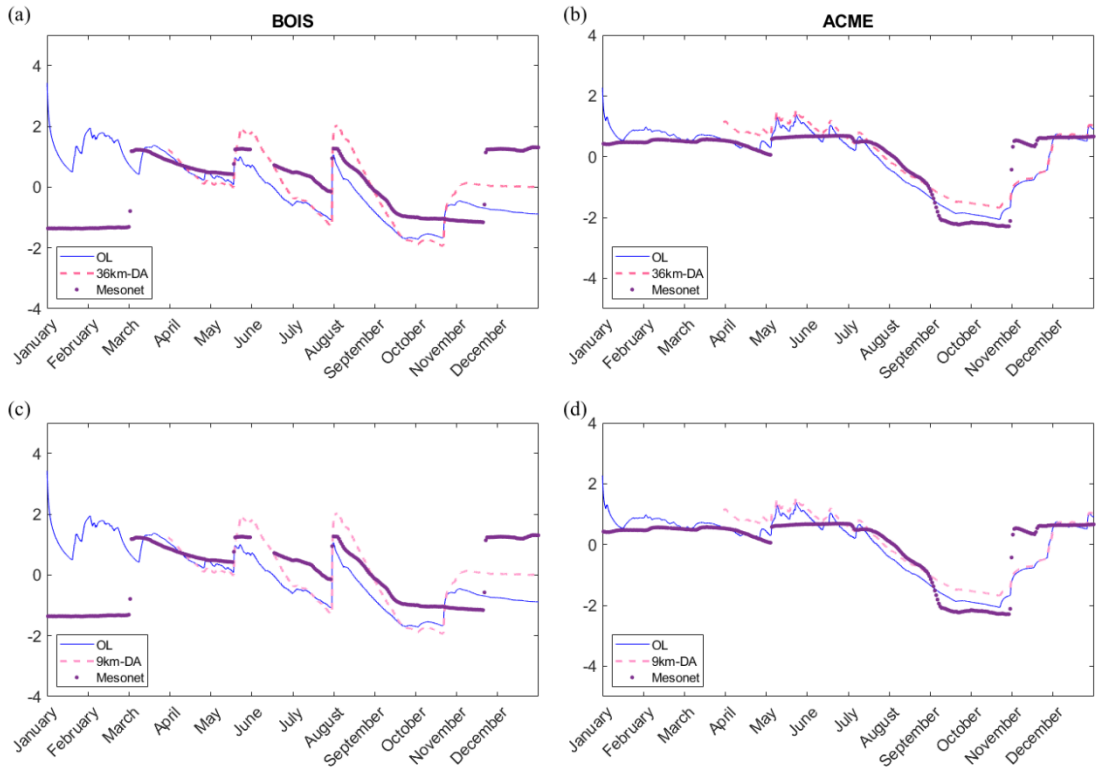

Figure 23 Same as in Figure 22 but for root zone soil moisture standard normal deviates

Figure 23 presents the time series of standard normal deviates of root zone soil moisture at the same two Mesonet stations (and corresponding model pixels). Similar to surface soil moisture, the OL run and the Mesonet ground observations show the same trends at both locations, with better agreement at the ACME station (characterized by relatively lower elevation). The impact of SMAP DA is then particularly evident at the BOIS station (relatively higher elevation) pushing the model run closer to the Mesonet observations. Specifically, DA improves R by 15% (from 0.58 in the OL to 0.67 in both 36-km DA and 9-km DA) at the BOIS station, as shown in Table 6. At the same location, the RMSE also decreases by 10% (from 0.89 in the OL to 0.80 in the DA runs). However,

at the lower elevation location, DA causes a 2% deterioration in R and 6% deterioration in RMSE, slightly worsening the Noah-MP performance. A pint to note is that correlations are still very high (>0.88) and RMSE low (<0.47) at the ACME location.

Table 6 Correlation and RMSE for two Mesonet locations of surface soil moisture standard normal deviates

Station	Simulation	Surface Soil Moisture		Root Zone Soil Moisture	
		R	RMSE	R	RMSE
BOIS	OL	0.51	1.04	0.58	0.89
	36-km DA	0.68	0.85	0.67	0.80
	9-km DA	0.68	0.85	0.67	0.80
ACME	OL	0.76	0.70	0.90	0.44
	36-km DA	0.75	0.71	0.88	0.47
	9-km DA	0.75	0.71	0.88	0.47

3.4 **Conclusions**

Land data assimilation systems that combine information from remotely-sensed retrievals and the spatially and temporally complete estimate from land surface models have the potential to provide a soil moisture product that is superior to the two parenting ones. Specifically, an LDAS incorporates observations (e.g., from satellites) that are limited in terms of temporal and spatial coverage (e.g., with soil depth) with estimates from a land surface model that captures the key land surface processes, such as the vertical transfer of water between the surface and root zone reservoirs.

The analysis of times series of soil moisture standard deviates shows how estimates of extremely different nature (ground observations, satellite products, and

model simulations) all capture the seasonality and variability of such anomalies and present high correlation between each other. Results show that the assimilation of SMAP soil moisture retrievals in a land surface model has the potential to improve the estimation of surface and root zone soil moisture, especially at higher elevations, increasing the correlation and reducing the random error between the model and the reference data. Although root zone soil moisture is not directly assimilated, it is impacted nonetheless by the assimilation of SMAP observations at the surface level, which brings root zone soil moisture estimates closer to the in situ observations. At the 12.5-km resolution adopted in this experiment, the two SMAP products (at 36 km and 9km resolution) behave very similarly, which may be attributed to the fact that they are both based on the same brightness temperature signal. Additionally, a monthly CDF matching method applied to the SMAP observations prior to the assimilation was found to be superior to a yearly CDF matching method. This proves that biases should be resolved at a finer temporal frequency than yearly to capture their seasonal variability.

Future work should investigate the assimilation of different satellite-based soil moisture products (e.g., SMOS) using different land surface models, focus on different regions of the world (including areas characterized by complex terrain and by denser vegetation), and extend the time series to a multi-year analysis. Future research should also apply the methodology proposed in Chapters 1 and 2 to downscale a set of atmospheric variables that will force a land surface model merged with satellite-based observations within land data assimilation system. This has the potential to improve at once the resolution and quality of surface and root zone soil moisture estimates.

CONCLUDING REMARKS

This work developed a method to improve the estimation of surface and root zone soil moisture at high resolution. First, an innovative physically-based scheme to downscale a suite of atmospheric variables (air temperature, pressure, humidity, wind speed, incident longwave and shortwave radiation) is proposed. The proposed downscaling framework, which comprises a statistical interpolation and a set of deterministic physical rules, is tested across Oklahoma during 2015 and applied to the NLDAS-2 dataset (from 12.5 km to 500 m). Second, the downscaled NLDAS-2 product is used to force a land surface model to simulate 500 m resolution surface and root zone soil moisture. Third, a land data assimilation system is adopted to merge land surface model simulations with products from the NASA SMAP (Soil Moisture Active Passive) satellite mission at two different resolutions (36 km and 9 km).

The science questions posed in the introductory section were addressed as follows:

1. *Is physically-based downscaling a viable approach to produce hyper-resolution atmospheric forcings?*

   The physically-based approach developed in this work was demonstrated to be a viable approach to downscale atmospheric forcings from a resolution of 12.5 km to 500 m. The downscaled product is evaluated with respect to high-quality, high-resolution ground-based observations collected by the Oklahoma Mesonet network. For most atmospheric variables, an improvement was observed

in terms of several performance metrics (bias, correlation coefficient, and RMSE). Although the proposed downscaled precipitation product was shown to only marginally improve the original resolution NLDAS-2 dataset, the level of detail that the 500 m atmospheric forcings presented compared to the 12.5 km resolution was noteworthy.

2. *What is the role of forcing resolution in land surface modeling? And in particular, what is the role of precipitation resolution relative to the other atmospheric forcings on simulated soil moisture?*

Finer resolution forcing data have the potential to improve estimates of surface and root zone soil moisture simulated by a land surface model, producing higher correlations and lower RMSEs with respect to Mesonet in situ observations. The resolution of input precipitation plays a critical role in improving the estimation of both surface and root zone soil moisture. The model forced with all the hyper-resolution atmospheric variables (including precipitation) is proven to be the most successful at pushing the model closer to the Mesonet surface and root zone soil moisture observations.

3. *How useful is SMAP to land surface modeling? What is the efficiency of a data assimilation system to estimate surface and root-zone soil moisture?*

Land data assimilation systems that combine information from remotely-sensed retrievals and the spatially and temporally complete estimate from land surface models have the potential to provide a soil moisture product that is superior to the two parenting ones. SMAP data assimilation was shown to improve estimates of surface and root zone

soil moisture anomalies, especially at higher elevation, by increasing the correlation with ground observations and reducing random errors.

The methodologies developed in this research have the potential to be generalized to different datasets, models, longer time series, different resolutions, different products, and other regions of the world. Future studies should investigate the efficiency of assimilating SMAP products in a land surface model forced with a hyper-resolution atmospheric variable dataset. As more SMAP products/versions become available, the framework developed here should be expanded to a larger scale and multiple spatio/temporal resolutions.

REFERENCES

Allen Richard G., Tasumi Masahiro, Trezza Ricardo, 2007. Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC)—Model. Journal of Irrigation and Drainage Engineering 133, 380–394. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:4(380)

Arsenault, K.R., Kumar, S.V., Geiger, J.V., Wang, S., Kemp, E., Mocko, D.M., Beaudoing, H.K., Getirana, A., Navari, M., Li, B., Jacob, J., Wegiel, J., Peters-Lidard, C.D., 2018. The Land surface Data Toolkit (LDT v7.2) – a data fusion environment for land data assimilation systems. Geosci. Model Dev. 11, 3605–3621. https://doi.org/10.5194/gmd-11-3605-2018

Arya, L.M., Paris, J.F., 1981. A Physicoempirical Model to Predict the Soil Moisture Characteristic from Particle-Size Distribution and Bulk Density Data 1. Soil Science Society of America Journal 45, 1023–1030. https://doi.org/10.2136/sssaj1981.03615995004500060004x

Badas, M.G., Deidda, R., Piga, E., 2006. Modulation of homogeneous space-time rainfall cascades to account for orographic influences. Nat. Hazards Earth Syst. Sci. 6, 427–437. https://doi.org/10.5194/nhess-6-427-2006

Baker, I.T., Sellers, P.J., Denning, A.S., Medina, I., Kraus, P., Haynes, K.D., Biraud, S.C., 2017. Closing the scale gap between land surface parameterizations and GCMs with a new scheme, SiB3-Bins. Journal of Advances in Modeling Earth Systems 9, 691–711. https://doi.org/10.1002/2016MS000764

Ball, J.T., Woodrow, I.E., Berry, J.A., 1987. A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions, in: Progress in Photosynthesis Research. Springer, pp. 221–224.

Bastiaanssen, W.G.M., 2000. SEBAL-based sensible and latent heat fluxes in the irrigated Gediz Basin, Turkey. Journal of Hydrology 229, 87–100. https://doi.org/10.1016/S0022-1694(99)00202-4

Beljaars, A.C.M., Viterbo, P., Miller, M.J., Betts, A.K., 1996. The Anomalous Rainfall over the United States during July 1993: Sensitivity to Land Surface Parameterization and Soil Moisture Anomalies. Mon. Wea. Rev. 124, 362–383. https://doi.org/10.1175/1520-0493(1996)124<0362:TAROTU>2.0.CO;2

Beven, K., Cloke, H., Pappenberger, F., Lamb, R., Hunter, N., 2015. Hyperresolution information and hyperresolution ignorance in modelling the hydrology of the land

surface. Sci. China Earth Sci. 58, 25–35. https://doi.org/10.1007/s11430-014-5003-4

Brock, F.V., Crawford, K.C., Elliott, R.L., Cuperus, G.W., Stadler, S.J., Johnson, H.L., Eilts, M.D., 1994. The Oklahoma Mesonet: A Technical Overview. J. Atmos. Oceanic Technol. 12, 5–19. https://doi.org/10.1175/1520-0426(1995)012<0005:TOMATO>2.0.CO;2

Brodzik, M.J., Billingsley, B., Haran, T., Raup, B., Savoie, M.H., 2012. EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets. ISPRS International Journal of Geo-Information 1, 32–45. https://doi.org/10.3390/ijgi1010032

Brutsaert, W., 1982. Evaporation into the atmosphere: Theory. History, and Applications 1.

Cai, X. (orcid:0000000247984954), Yang, Z.-L., Fisher, J.B., Zhang, X. (orcid:0000000347117751), Barlage, M., Chen, F., 2016. Integration of nitrogen dynamics into the Noah-MP land surface model v1.1 for climate and environmental predictions. Geoscientific Model Development (Online) 9. https://doi.org/10.5194/gmd-9-1-2016

Cai, X., Yang, Z.-L., David, C.H., Niu, G.-Y., Rodell, M., 2014. Hydrological evaluation of the Noah-MP land surface model for the Mississippi River Basin. Journal of Geophysical Research: Atmospheres 119, 23–38. https://doi.org/10.1002/2013JD020792

Chan, S.K., Bindlish, R., O'Neill, P., Jackson, T., Njoku, E., Dunbar, S., Chaubell, J., Piepmeier, J., Yueh, S., Entekhabi, D., Colliander, A., Chen, F., Cosh, M.H., Caldwell, T., Walker, J., Berg, A., McNairn, H., Thibeault, M., Martínez-Fernández, J., Uldall, F., Seyfried, M., Bosch, D., Starks, P., Holifield Collins, C., Prueger, J., van der Velde, R., Asanuma, J., Palecki, M., Small, E.E., Zreda, M., Calvet, J., Crow, W.T., Kerr, Y., 2018. Development and assessment of the SMAP enhanced passive soil moisture product. Remote Sensing of Environment 204, 931–941. https://doi.org/10.1016/j.rse.2017.08.025

Chan, Steven, 2016. Enhanced Level 3 Passive Soil Moisture Product Specification Document [WWW Document]. URL https://scholar.google.com/scholar_lookup?title=Enhanced%20Level%203%20Passive%20Soil%20Moisture%20Product%20Specification%20Document&publication_year=2016&author=S.%20Chan (accessed 4.2.20).

Chen, F., Barlage, M., Tewari, M., Rasmussen, R., Jin, J., Lettenmaier, D., Livneh, B., Lin, C., Miguez-Macho, G., Niu, G.-Y., Wen, L., Yang, Z.-L., 2014. Modeling seasonal snowpack evolution in the complex terrain and forested Colorado Headwaters region: A model intercomparison study. Journal of Geophysical Research: Atmospheres 119, 13,795-13,819. https://doi.org/10.1002/2014JD022167

Chen, F., Dudhia, J., 2001. Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. Mon. Wea. Rev. 129, 569–585. https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2

Colliander, A., Jackson, T.J., Chan, S.K., O'Neill, P., Bindlish, R., Cosh, M.H., Caldwell, T., Walker, J.P., Berg, A., McNairn, H., Thibeault, M., Martínez-Fernández, J., Jensen, K.H., Asanuma, J., Seyfried, M.S., Bosch, D.D., Starks, P.J., Holifield Collins, C., Prueger, J.H., Su, Z., Lopez-Baeza, E., Yueh, S.H., 2018. An assessment of the differences between spatial resolution and grid size for the SMAP enhanced soil moisture product over homogeneous sites. Remote Sensing of Environment 207, 65–70. https://doi.org/10.1016/j.rse.2018.02.006

Cosgrove, B.A., Lohmann, D., Mitchell, K.E., Houser, P.R., Wood, E.F., Schaake, J.C., Robock, A., Marshall, C., Sheffield, J., Duan, Q., Luo, L., Higgins, R.W., Pinker, R.T., Tarpley, J.D., Meng, J., 2003. Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. Journal of Geophysical Research: Atmospheres 108. https://doi.org/10.1029/2002JD003118

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

De Lannoy, G.J.M., Reichle, R.H., Peng, J., Kerr, Y., Castro, R., Kim, E.J., Liu, Q., 2015. Converting Between SMOS and SMAP Level-1 Brightness Temperature Observations Over Nonfrozen Land. IEEE Geoscience and Remote Sensing Letters 12, 1908–1912. https://doi.org/10.1109/LGRS.2015.2437612

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., Berg, L. van de, Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., Rosnay, P. de, Tavolato, C., Thépaut, J.-N., Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society 137, 553–597. https://doi.org/10.1002/qj.828

Dingman, S.L., 2015. Physical Hydrology: Third Edition. Waveland Press.

Dirmeyer, P.A., 2000. Using a Global Soil Wetness Dataset to Improve Seasonal Climate Simulation. J. Climate 13, 2900–2922. https://doi.org/10.1175/1520-0442(2000)013<2900:UAGSWD>2.0.CO;2

Dong, Z., Gao, S., Fryrear, D.W., 2001. Drag coefficients, roughness length and zero-plane displacement height as disturbed by artificial standing vegetation. Journal of Arid Environments 49, 485–505. https://doi.org/10.1006/jare.2001.0807

Durand, M., Margulis, S.A., 2008. Effects of uncertainty magnitude and accuracy on assimilation of multiscale measurements for snowpack characterization. Journal of Geophysical Research: Atmospheres 113. https://doi.org/10.1029/2007JD008662

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, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Van Zyl, J., 2010a. The Soil Moisture Active Passive (SMAP) Mission. Proceedings of the IEEE 98, 704–716. https://doi.org/10.1109/JPROC.2010.2043918

Entekhabi, D., Reichle, R.H., Koster, R.D., Crow, W.T., 2010b. Performance Metrics for Soil Moisture Retrievals and Application Requirements. J. Hydrometeor. 11, 832–840. https://doi.org/10.1175/2010JHM1223.1

Evensen, G., 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dynamics 53, 343–367. https://doi.org/10.1007/s10236-003-0036-9

Evensen, G., van Leeuwen, P.J., 1996. Assimilation of Geosat Altimeter Data for the Agulhas Current Using the Ensemble Kalman Filter with a Quasigeostrophic Model. Mon. Wea. Rev. 124, 85–96. https://doi.org/10.1175/1520-0493(1996)124<0085:AOGADF>2.0.CO;2

Fiddes, J., Gruber, S., 2014. TopoSCALE v.1.0: Downscaling gridded climate data in complex terrain. Geoscientific Model Development 7, 387–405.

Fischer, E.M., Seneviratne, S.I., Vidale, P.L., Lüthi, D., Schär, C., 2007. Soil Moisture–Atmosphere Interactions during the 2003 European Summer Heat Wave. J. Climate 20, 5081–5099. https://doi.org/10.1175/JCLI4288.1

Franz, T.E., Loecke, T.D., Burgin, A.J., Zhou, Y., Le, T., Moscicki, D., 2017. Spatiotemporal predictions of soil properties and states in variably saturated landscapes. Journal of Geophysical Research: Biogeosciences 122, 1576–1596. https://doi.org/10.1002/2017JG003837

Garnaud, C., Bélair, S., Berg, A., Rowlandson, T., 2015. Hyperresolution Land Surface Modeling in the Context of SMAP Cal–Val. J. Hydrometeor. 17, 345–352. https://doi.org/10.1175/JHM-D-15-0070.1

Gaur, A., Simonovic, S.P., 2017. Accessing vulnerability of land-cover types to climate change using physical scaling downscaling model. International Journal of Climatology 37, 2901–2912. https://doi.org/10.1002/joc.4887

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao,

B., 2017. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). J. Climate 30, 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Gibson, J., Franz, T.E., Wang, T., Gates, J., Grassini, P., Yang, H., Eisenhauer, D., 2017. A case study of field-scale maize irrigation patterns in western Nebraska: implications for water managers and recommendations for hyper-resolution land surface modeling. Hydrology and Earth System Sciences 21, 1051–1062. https://doi.org/10.5194/hess-21-1051-2017

Girotto, M., Margulis, S.A., Durand, M., 2014. Probabilistic SWE reanalysis as a generalization of deterministic SWE reconstruction techniques. Hydrological Processes 28, 3875–3895. https://doi.org/10.1002/hyp.9887

Gottschalck, J., Meng, J., Rodell, M., Houser, P., 2005. Analysis of Multiple Precipitation Products and Preliminary Assessment of Their Impact on Global Land Data Assimilation System Land Surface States. J. Hydrometeor. 6, 573–598. https://doi.org/10.1175/JHM437.1

Gupta, H.V., Nearing, G.S., 2014. Debates—the future of hydrological sciences: A (common) path forward? Using models and data to learn: A systems theoretic perspective on the future of hydrological science. Water Resources Research 50, 5351–5359. https://doi.org/10.1002/2013WR015096

Häntzschel, J., Goldberg, V., Bernhofer, C., 2005. GIS-based regionalisation of radiation, temperature and coupling measures in complex terrain for low mountain ranges. Meteorological Applications 12, 33–42. https://doi.org/10.1017/S1350482705001489

Haylock, M.R., Cawley, G.C., Harpham, C., Wilby, R.L., Goodess, C.M., 2006. Downscaling heavy precipitation over the United Kingdom: a comparison of dynamical and statistical methods and their future scenarios. International Journal of Climatology 26, 1397–1415. https://doi.org/10.1002/joc.1318

Hazra, A., Maggioni, V., Houser, P., Antil, H., Noonan, M., 2019. A Monte Carlo-based multi-objective optimization approach to merge different precipitation estimates for land surface modeling. Journal of Hydrology 570, 454–462. https://doi.org/10.1016/j.jhydrol.2018.12.039

Hossain, F., Anagnostou, E.N., 2005. Numerical investigation of the impact of uncertainties in satellite rainfall estimation and land surface model parameters on simulation of soil moisture. Advances in Water Resources 28, 1336–1350. https://doi.org/10.1016/j.advwatres.2005.03.013

Houtekamer, P.L., Mitchell, H.L., 1998. Data Assimilation Using an Ensemble Kalman Filter Technique. Mon. Wea. Rev. 126, 796–811. https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2

Illston, B.G., Basara, J.B., Fiebrich, C.A., Crawford, K.C., Hunt, E., Fisher, D.K., Elliott, R., Humes, K., 2008. Mesoscale Monitoring of Soil Moisture across a Statewide

Network. J. Atmos. Oceanic Technol. 25, 167–182. https://doi.org/10.1175/2007JTECHA993.1

Jackson, T.J., 1993. III. Measuring surface soil moisture using passive microwave remote sensing. Hydrological Processes 7, 139–152. https://doi.org/10.1002/hyp.3360070205

Ji, P., Yuan, X., Liang, X.-Z., 2017. Do Lateral Flows Matter for the Hyperresolution Land Surface Modeling? Journal of Geophysical Research: Atmospheres 122, 12,077-12,092. https://doi.org/10.1002/2017JD027366

Jordan, R., 1991. A one-dimensional temperature model for a snow cover, Spec. Rep. 91–16, Cold Reg. Res. and Eng. Lab., US Army Corps of Eng., Hanover, NH.

Keppenne, C.L., 2000. Data assimilation into a primitive-equation model with a parallel ensemble Kalman filter. Monthly Weather Review 128, 1971–1981.

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., Martín-Neira, M., Mecklenburg, S., 2010. The SMOS Mission: New Tool for Monitoring Key Elements ofthe Global Water Cycle. Proceedings of the IEEE 98, 666–687. https://doi.org/10.1109/JPROC.2010.2043032

Kerr, Y.H., Waldteufel, P., Wigneron, J.-P., Martinuzzi, J., Font, J., Berger, M., 2001. Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission. IEEE Transactions on Geoscience and Remote Sensing 39, 1729–1735. https://doi.org/10.1109/36.942551

Ko, A., Mascaro, G., Vivoni, E.R., 2019. Strategies to Improve and Evaluate Physics-Based Hyperresolution Hydrologic Simulations at Regional Basin Scales. Water Resources Research 55, 1129–1152. https://doi.org/10.1029/2018WR023521

Koster, R.D., Dirmeyer, P.A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C.T., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C.-H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Sud, Y.C., Taylor, C.M., Verseghy, D., Vasic, R., Xue, Y., Yamada, T., 2004. Regions of Strong Coupling Between Soil Moisture and Precipitation. Science 305, 1138–1140. https://doi.org/10.1126/science.1100217

Koster, R.D., Suarez, M.J., 2003. Impact of Land Surface Initialization on Seasonal Precipitation and Temperature Prediction. J. Hydrometeor. 4, 408–423. https://doi.org/10.1175/1525-7541(2003)4<408:IOLSIO>2.0.CO;2

Kumar, S., Peters-Lidard, C., Tian, Y., Reichle, R., Geiger, J., Alonge, C., Eylander, J., Houser, P., 2008. An integrated hydrologic modeling and data assimilation framework. Computer 41, 52–59. https://doi.org/10.1109/MC.2008.475

Kumar, S.V., Jasinski, M., Mocko, D.M., Rodell, M., Borak, J., Li, B., Beaudoing, H.K., Peters-Lidard, C.D., 2018. NCA-LDAS Land Analysis: Development and Performance of a Multisensor, Multivariate Land Data Assimilation System for

the National Climate Assessment. J. Hydrometeor. 20, 1571–1593. https://doi.org/10.1175/JHM-D-17-0125.1

Kumar, S.V., M. Mocko, D., Wang, S., Peters-Lidard, C.D., Borak, J., 2019. Assimilation of Remotely Sensed Leaf Area Index into the Noah-MP Land Surface Model: Impacts on Water and Carbon Fluxes and States over the Continental United States. J. Hydrometeor. 20, 1359–1377. https://doi.org/10.1175/JHM-D-18-0237.1

Kumar, S.V., Peters-Lidard, C.D., Mocko, D., Reichle, R., Liu, Y., Arsenault, K.R., Xia, Y., Ek, M., Riggs, G., Livneh, B., Cosh, M., 2014. Assimilation of Remotely Sensed Soil Moisture and Snow Depth Retrievals for Drought Estimation. J. Hydrometeor. 15, 2446–2469. https://doi.org/10.1175/JHM-D-13-0132.1

Kumar, S.V., Reichle, R.H., Harrison, K.W., Peters-Lidard, C.D., Yatheendradas, S., Santanello, J.A., 2012. A comparison of methods for a priori bias correction in soil moisture data assimilation. Water Resources Research 48. https://doi.org/10.1029/2010WR010261

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. Advances in Water Resources, Hydrologic Remote Sensing 31, 1419–1432. https://doi.org/10.1016/j.advwatres.2008.01.013

Lawrence, M.G., 2005. The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications. Bull. Amer. Meteor. Soc. 86, 225–234. https://doi.org/10.1175/BAMS-86-2-225

Lermusiaux, P.F.J., 1999. Data Assimilation via Error Subspace Statistical Estimation. Mon. Wea. Rev. 127, 1408–1432. https://doi.org/10.1175/1520-0493(1999)127<1408:DAVESS>2.0.CO;2

Liston, G.E., Elder, K., 2006. A Meteorological Distribution System for High-Resolution Terrestrial Modeling (MicroMet). J. Hydrometeor. 7, 217–234. https://doi.org/10.1175/JHM486.1

López López, P., Immerzeel, W.W., Rodríguez Sandoval, E.A., Sterk, G., Schellekens, J., 2018. Spatial Downscaling of Satellite-Based Precipitation and Its Impact on Discharge Simulations in the Magdalena River Basin in Colombia. Front. Earth Sci. 6. https://doi.org/10.3389/feart.2018.00068

Ma, N., Niu, G.-Y., Xia, Y., Cai, X., Zhang, Y., Ma, Y., Fang, Y., 2017. A Systematic Evaluation of Noah-MP in Simulating Land-Atmosphere Energy, Water, and Carbon Exchanges Over the Continental United States. Journal of Geophysical Research: Atmospheres 122, 12,245-12,268. https://doi.org/10.1002/2017JD027597

Madsen, H., Canizares, R., 1999. Comparison of extended and ensemble Kalman filters for data assimilation in coastal area modelling. International Journal for Numerical Methods in Fluids 31, 961–981.

Maggioni, V., Anagnostou, E.N., Reichle, R.H., 2012. The impact of model and rainfall forcing errors on characterizing soil moisture uncertainty in land surface modeling. Hydrology and Earth System Sciences 16, 3499–3515. https://doi.org/10.5194/hess-16-3499-2012

Maggioni, V., Houser, P.R., 2017. Soil Moisture Data Assimilation, in: Park, S.K., Xu, L. (Eds.), Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. III). Springer International Publishing, Cham, pp. 195–217. https://doi.org/10.1007/978-3-319-43415-5_9

Maggioni, Viviana, Reichle, R.H., Anagnostou, E.N., 2012. The Efficiency of Assimilating Satellite Soil Moisture Retrievals in a Land Data Assimilation System Using Different Rainfall Error Models. J. Hydrometeor. 14, 368–374. https://doi.org/10.1175/JHM-D-12-0105.1

Maggioni, V., Reichle, R.H., Anagnostou, E.N., 2011. The Effect of Satellite Rainfall Error Modeling on Soil Moisture Prediction Uncertainty. J. Hydrometeor. 12, 413–428. https://doi.org/10.1175/2011JHM1355.1

Maidment, D.R., 2017. Conceptual Framework for the National Flood Interoperability Experiment. JAWRA Journal of the American Water Resources Association 53, 245–257. https://doi.org/10.1111/1752-1688.12474

Martens, B.V. der V., Illston, B.G., Fiebrich, C.A., 2017. The Oklahoma Mesonet: A Pilot Study of Environmental Sensor Data Citations. Data Science Journal 16, 47. https://doi.org/10.5334/dsj-2017-047

McPherson, R.A., Fiebrich, C.A., Crawford, K.C., Kilby, J.R., Grimsley, D.L., Martinez, J.E., Basara, J.B., Illston, B.G., Morris, D.A., Kloesel, K.A., Melvin, A.D., Shrivastava, H., Wolfinbarger, J.M., Bostic, J.P., Demko, D.B., Elliott, R.L., Stadler, S.J., Carlson, J.D., Sutherland, A.J., 2007. Statewide Monitoring of the Mesoscale Environment: A Technical Update on the Oklahoma Mesonet. J. Atmos. Oceanic Technol. 24, 301–321. https://doi.org/10.1175/JTECH1976.1

Mei, Y., Maggioni, V., Houser, P., Xue, Y., Rouf, T., 2020. A Nonparametric Statistical Technique for Spatial Downscaling of Precipitation over High Mountain Asia. Water Resources Research.

Mei, Y., Maggioni, V., Houser, P., Xue, Y., Rouf, T., 2018. A Nonparametric Statistical Technique for Spatial Downscaling of Precipitation over High Mountain Asia. AGU Fall Meeting Abstracts 21.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E.H., Ek, M.B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., Shi, W., 2006. North

American Regional Reanalysis. Bull. Amer. Meteor. Soc. 87, 343–360. https://doi.org/10.1175/BAMS-87-3-343

Mesonet | Home Page [WWW Document], n.d. URL https://www.mesonet.org/index.php (accessed 1.16.19).

Michalsky, J., Lantz, K., National Oceanic and Atmospheric Administration, University of Colorado, Boulder, 2016. Two-Column Aerosol Project (TCAP): Ground-Based Radiation and Aerosol Validation Using the NOAA Mobile SURFRAD Station Field Campaign Report (No. DOE/SC--ARM-14-043, 1254181). https://doi.org/10.2172/1254181

Mitchell, K.E., 2004. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. Journal of Geophysical Research 109. https://doi.org/10.1029/2003JD003823

Nicholas, F.W., Lewis, J.E., 1980. Relationships between aerodynamic roughness and land use and land cover in Baltimore, Maryland. U.S. G.P.O.

Niu, G.-Y., Yang, Z.-L., 2006. Effects of Frozen Soil on Snowmelt Runoff and Soil Water Storage at a Continental Scale. J. Hydrometeor. 7, 937–952. https://doi.org/10.1175/JHM538.1

Niu, G.-Y., Yang, Z.-L., Dickinson, R.E., Gulden, L.E., Su, H., 2007. Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data. Journal of Geophysical Research: Atmospheres 112. https://doi.org/10.1029/2006JD007522

Niu, G.-Y., Yang, Z.-L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., Xia, Y., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. Journal of Geophysical Research: Atmospheres 116. https://doi.org/10.1029/2010JD015139

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. https://doi.org/10.1109/TGRS.2002.808243

O'Neill, P., Chan, S., Colliander, A., Dunbar, S., Njoku, E., Bindlish, R., Chen, F., Jackson, T., Burgin, M., Piepmeier, J., Yueh, S., Entekhabi, D., Cosh, M., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., McNairn, H., Thibeault, M., Martínez-Fernández, J., González-Zamora, Á., Seyfried, M., Bosch, D., Starks, P., Goodrich, D., Prueger, J., Palecki, M., Small, E., Zreda, M., Calvet, J.-C., Crow, W., Kerr, Y., 2016. Evaluation of the validated Soil Moisture product from the SMAP radiometer, in: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Presented at the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 125–128. https://doi.org/10.1109/IGARSS.2016.7729023

ONeill, P.E., 2018. SMAP Enhanced L3 Radiometer Global Daily 9 km EASE-Grid Soil Moisture, Version 2. https://doi.org/10.5067/RFKIZ5QY5ABN

Oral history interview with Fred V. Brock [WWW Document], n.d. URL https://dc.library.okstate.edu/digital/collection/mesonet/id/266 (accessed 1.16.19).

Oral history interview with Ken Crawford [WWW Document], n.d. URL https://dc.library.okstate.edu/digital/collection/mesonet/id/293 (accessed 1.16.19).

Pan, M., Wood, E.F., 2006. Data Assimilation for Estimating the Terrestrial Water Budget Using a Constrained Ensemble Kalman Filter. J. Hydrometeor. 7, 534–547. https://doi.org/10.1175/JHM495.1

Peters-Lidard, C.D., Houser, P.R., Tian, Y., Kumar, S.V., Geiger, J., Olden, S., Lighty, L., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E.F., Sheffield, J., 2007a. High-performance Earth system modeling with NASA/GSFC's Land Information System. Innovations Syst Softw Eng 3, 157–165. https://doi.org/10.1007/s11334-007-0028-x

Peters-Lidard, C.D., Houser, P.R., Tian, Y., Kumar, S.V., Geiger, J., Olden, S., Lighty, L., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E.F., Sheffield, J., 2007b. High-performance Earth system modeling with NASA/GSFC's Land Information System. Innovations Syst Softw Eng 3, 157–165. https://doi.org/10.1007/s11334-007-0028-x

Pinker, R.T., Tarpley, J.D., Laszlo, I., Mitchell, K.E., Houser, P.R., Wood, E.F., Schaake, J.C., Robock, A., Lohmann, D., Cosgrove, B.A., Sheffield, J., Duan, Q., Luo, L., Higgins, R.W., 2003. Surface radiation budgets in support of the GEWEX Continental-Scale International Project (GCIP) and the GEWEX Americas Prediction Project (GAPP), including the North American Land Data Assimilation System (NLDAS) project. Journal of Geophysical Research: Atmospheres 108. https://doi.org/10.1029/2002JD003301

Reichle, R.H., 2008. Data assimilation methods in the Earth sciences. Advances in Water Resources, Hydrologic Remote Sensing 31, 1411–1418. https://doi.org/10.1016/j.advwatres.2008.01.001

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. https://doi.org/10.1029/2004GL021700

Reichle, R.H., Koster, R.D., 2004. Bias reduction in short records of satellite soil moisture. Geophysical Research Letters 31. https://doi.org/10.1029/2004GL020938

Reichle, R.H., McLaughlin, D.B., Entekhabi, D., 2002a. Hydrologic Data Assimilation with the Ensemble Kalman Filter. Mon. Wea. Rev. 130, 103–114. https://doi.org/10.1175/1520-0493(2002)130<0103:HDAWTE>2.0.CO;2

Reichle, R.H., Walker, J.P., Koster, R.D., Houser, P.R., 2002b. Extended versus Ensemble Kalman Filtering for Land Data Assimilation. J. Hydrometeor. 3, 728–740. https://doi.org/10.1175/1525-7541(2002)003<0728:EVEKFF>2.0.CO;2

Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R.D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C.R., Reichle, R., Robertson, F.R., Ruddick, A.G., Sienkiewicz, M., Woollen, J., 2011. MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. J. Climate 24, 3624–3648. https://doi.org/10.1175/JCLI-D-11-00015.1

Robock, A., Vinnikov, K.Y., Srinivasan, G., Entin, J.K., Hollinger, S.E., Speranskaya, N.A., Liu, S., Namkhai, A., 2000. The Global Soil Moisture Data Bank. Bull. Amer. Meteor. Soc. 81, 1281–1300. https://doi.org/10.1175/1520-0477(2000)081<1281:TGSMDB>2.3.CO;2

Rodell, M., Houser, P.R., Berg, A.A., Famiglietti, J.S., 2005. Evaluation of 10 Methods for Initializing a Land Surface Model. J. Hydrometeor. 6, 146–155. https://doi.org/10.1175/JHM414.1

Rogers, E., Black, T.L., Deaven, D.G., DiMego, G.J., Zhao, Q., Baldwin, M., Junker, N.W., Lin, Y., 1996. Changes to the operational "early" Eta analysis/forecast system at the National Centers for Environmental Prediction. Weather and Forecasting 11, 391–413.

Rouf, T., Mei, Y., Maggioni, V., Houser, P., Noonan, M., 2019. A Physically Based Atmospheric Variables Downscaling Technique. J. Hydrometeor. 21, 93–108. https://doi.org/10.1175/JHM-D-19-0109.1

Ruiz-Arias, J.A., Alsamamra, H., Tovar-Pescador, J., Pozo-Vázquez, D., 2010. Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions. Energy Conversion and Management 51, 881–893. https://doi.org/10.1016/j.enconman.2009.11.024

Sabater, J.M., Jarlan, L., Calvet, J.-C., Bouyssel, F., De Rosnay, P., 2007. From Near-Surface to Root-Zone Soil Moisture Using Different Assimilation Techniques. J. Hydrometeor. 8, 194–206. https://doi.org/10.1175/JHM571.1

Sen Gupta, A., Tarboton, D.G., 2016. A tool for downscaling weather data from large-grid reanalysis products to finer spatial scales for distributed hydrological applications. Environmental Modelling & Software 84, 50–69. https://doi.org/10.1016/j.envsoft.2016.06.014

Senatore, A., Mendicino, G., Gochis, D.J., Yu, W., Yates, D.N., Kunstmann, H., 2015. Fully coupled atmosphere-hydrology simulations for the central Mediterranean: Impact of enhanced hydrological parameterization for short and long time scales. Journal of Advances in Modeling Earth Systems 7, 1693–1715. https://doi.org/10.1002/2015MS000510

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

Serpetzoglou, E., Anagnostou, E.N., Papadopoulos, A., Nikolopoulos, E.I., Maggioni, V., 2010. Error Propagation of Remote Sensing Rainfall Estimates in Soil Moisture Prediction from a Land Surface Model. J. Hydrometeor. 11, 705–720. https://doi.org/10.1175/2009JHM1166.1

Shukla, J., Mintz, Y., 1982. Influence of Land-Surface Evapotranspiration on the Earth's Climate. Science 215, 1498–1501. https://doi.org/10.1126/science.215.4539.1498

Singh, R.S., Reager, J.T., Miller, N.L., Famiglietti, J.S., 2015. Toward hyper-resolution land-surface modeling: The effects of fine-scale topography and soil texture on CLM4.0 simulations over the Southwestern U.S. Water Resources Research 51, 2648–2667. https://doi.org/10.1002/2014WR015686

SRTM Data – CGIAR-CSI SRTM [WWW Document], n.d. URL http://srtm.csi.cgiar.org/srtmdata/ (accessed 1.16.19).

Stokes, G.M., Schwartz, S.E., 1994. The Atmospheric Radiation Measurement (ARM) Program: Programmatic Background and Design of the Cloud and Radiation Test Bed. Bulletin of the American Meteorological Society 75, 1201–1221. https://doi.org/10.1175/1520-0477(1994)075<1201:TARMPP>2.0.CO;2

Sud, Y.C., Shukla, J., Mintz, Y., 1985. Influence of land-surface roughness on atmospheric circulation and rainfall: A sensitivity study with a GCM. Presented at the its Res. Rev., 1983 p 129-130 (SEE N85-29433 18-47).

Sunyer, M.A., Hundecha, Y., Lawrence, D., Madsen, H., Willems, P., Martinkova, M., Vormoor, K., Bürger, G., Hanel, M., Kriaučiūnienė, J., Loukas, A., Osuch, M., Yücel, I., 2015. Inter-comparison of statistical downscaling methods for projection of extreme precipitation in Europe. Hydrology and Earth System Sciences 19, 1827–1847. https://doi.org/10.5194/hess-19-1827-2015

Survey, U.G., 2006. "Finished" 3-arc second SRTM Format Documentation.

Tao, J., Barros, A.P., 2018. Multi-year atmospheric forcing datasets for hydrologic modeling in regions of complex terrain – Methodology and evaluation over the Integrated Precipitation and Hydrology Experiment 2014 domain. Journal of Hydrology 567, 824–842. https://doi.org/10.1016/j.jhydrol.2016.12.058

Van Genuchten, M.Th., 1980. A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils 1. Soil Science Society of America Journal 44, 892–898. https://doi.org/10.2136/sssaj1980.03615995004400050002x

Venugopal, V., Foufoula-Georgiou, E., Sapozhnikov, V., 1999. A space-time downscaling model for rainfall. J. Geophys. Res. 104, 19705–19721. https://doi.org/10.1029/1999JD900338

Verseghy, D.L., 1991. Class—A Canadian land surface scheme for GCMS. I. Soil model. International Journal of Climatology 11, 111–133. https://doi.org/10.1002/joc.3370110202

Wood, E.F., Roundy, J.K., Troy, T.J., Beek, L.P.H. van, Bierkens, M.F.P., Blyth, E., Roo, A. de, Döll, P., Ek, M., Famiglietti, J., Gochis, D., Giesen, N. van de, 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 Resources Research 47. https://doi.org/10.1029/2010WR010090

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y., Mocko, D., 2012. Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. Journal of Geophysical Research: Atmospheres 117. https://doi.org/10.1029/2011JD016048

Xue, Y., Houser, P.R., Maggioni, V., Mei, Y., Kumar, S.V., Yoon, Y., 2019. Assimilation of Satellite-Based Snow Cover and Freeze/Thaw Observations Over High Mountain Asia. Front. Earth Sci. 7, 115. https://doi.org/10.3389/feart.2019.00115

Yang, Z.-L., Niu, G.-Y., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M., Xia, Y., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. Journal of Geophysical Research: Atmospheres 116. https://doi.org/10.1029/2010JD015140

Yin, J., Zhan, X., 2018. Impact of Bias-Correction Methods on Effectiveness of Assimilating SMAP Soil Moisture Data into NCEP Global Forecast System Using the Ensemble Kalman Filter. IEEE Geoscience and Remote Sensing Letters 15, 659–663. https://doi.org/10.1109/LGRS.2018.2806092

Yin, J., Zhan, X., Zheng, Y., Hain, C.R., Liu, J., Fang, L., 2015. Optimal ensemble size of ensemble Kalman filter in sequential soil moisture data assimilation. Geophysical Research Letters 42, 6710–6715. https://doi.org/10.1002/2015GL063366

Zhou, Q., Yang, S., Zhao, C., Cai, M., Lou, H., Luo, Y., Hou, L., 2016. Development and implementation of a spatial unit non-overlapping water stress index for water scarcity evaluation with a moderate spatial resolution. Ecological Indicators 69, 422–433. https://doi.org/10.1016/j.ecolind.2016.05.006

Zhou, X., Zhang, Y., Wang, Y., Zhang, H., Vaze, J., Zhang, L., Yang, Y., Zhou, Y., 2012. Benchmarking global land surface models against the observed mean

annual runoff from 150 large basins. Journal of Hydrology 470–471, 269–279. https://doi.org/10.1016/j.jhydrol.2012.09.002

Zorzetto, E., Marani, M., 2019. Downscaling of Rainfall Extremes From Satellite Observations. Water Resources Research 55, 156–174. https://doi.org/10.1029/2018WR022950

BIOGRAPHY

My name is Tasnuva Rouf. I received my Bachelor of Science degree in Water Resources Engineering in 2012 and Master of Science Degree in Water Resources Engineering in 2015 from the Bangladesh University of Engineering and Technology (BUET). After graduation, I have joined BUET as a Lecturer and taught several courses in civil engineering. I joined the Ph.D. program in Civil, Environmental, and Infrastructure Engineering at George Mason University (Mason) in Fall 2016 as a Graduate Research Assistant. I was the recipient of the Balfour Beatty Distinguished Graduate Fellowship for the academic year 2016/2017, of two Graduate Student Travel Fund scholarships for the academic years 2017/2018, 2018/2019, and 2019/2020 to cover travel expenses to national conferences. I have attended the Smithsonian-Mason School of Conservation course on 'Statistical Downscaling of Global Climate Models using SDSM 5.2' thanks to the ConocoPhillips Water and Biodiversity Stewardship Scholarship. I spent the summer of 2018 at the National Water Center Innovators Program Summer Institute at the University of Alabama in Tuscaloosa, thanks to the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUASHI) Fellowship. In Fall 2019, I served as a Graduate Teaching Assistant in the Civil, Environmental, and Infrastructure Engineering Department. In September 2019, I joined a field campaign to collect soil moisture data in Yanco, in Southwestern Australia. I was also awarded the Spring 2020 Provost Dissertation Completion Grant, which facilitated the timely completion of my doctoral thesis.