8-9-2022

Quantifying numerical weather and surface model sensitivity to land use and land cover changes

Hossein Lotfi
Mississippi State University, agri.meteo@gmail.com

Quantifying numerical weather and surface model sensitivity to land use and land cover changes

By

Hossein Lotfi

Approved by:

Jamie L. Dyer (Major Professor)
Padmanava Dash
Wenmeng Tian
Andrew E. Mercer (Committee Member/Graduate Coordinator)
Rick Travis (Dean, College of Arts & Sciences)

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Earth and Atmospheric Sciences
in the Department of Geosciences

Mississippi State, Mississippi

August 2022

Name: Hossein Lotfi

Date of Degree: August 9, 2022

Institution: Mississippi State University

Major Field: Earth and Atmospheric Sciences

Major Professor: Jamie L. Dyer

Title of Study: Quantifying numerical weather and surface model sensitivity to land use and land cover changes

Pages in Study: 105

Candidate for Degree of Doctor of Philosophy

Land surfaces have changed as a result of human and natural processes, such as deforestation, urbanization, desertification and natural disasters like wildfires. Land use and land cover change impacts local and regional climates through various bio geophysical processes across many time scales. More realistic representation of land surface parameters within the land surface models are essential to for climate models to accurately simulate the effects of past, current and future land surface processes. In this study, we evaluated the sensitivity and accuracy of the Weather Research and Forecasting (WRF) model though the default MODIS land cover data and annually updated land cover data over southeast of United States. Findings of this study indicated that the land surface fluxes, and moisture simulations are more sensitive to the surface characteristics over the southeast US. Consequently, we evaluated the WRF temperature and precipitation simulations with more accurate observations of land surface parameters over the study area. We evaluate the model performance for the default and updated land cover simulations against observational datasets. Results of the study showed that updating land cover resulted in substantial variations in surface heat fluxes and moisture balances. Despite updated land use and land cover data provided more representative land surface characteristics, the WRF simulated 2-

m temperature and precipitation did not improved due to use of updated land cover data. Further, we conducted machine learning experiments to post-process the Noah-MP land surface model simulations to determine if post processing the model outputs can improve the land surface parameters. The results indicate that the Noah-MP simulations using machine learning remarkably improved simulation accuracy and gradient boosting, and random forest model had smaller mean error bias values and larger coefficient of determination over the majority of stations. Moreover, the findings of the current study showed that the accuracy of surface heat flux simulations by Noah-MP are influenced by land cover and vegetation type.

ACKNOWLEDGEMENTS

The author would like to express his deepest appreciation to Dr. Dyer for providing the opportunity to learn and work under his guidance. Gratitude is extended to the supporting Dr. Mercer for guiding this project and also other committee members including Dr. Dash and Dr. Tian. Thank you all for your individual time and efforts invested in my education both inside and out of the classroom. I would also like to acknowledge my wife, family, and friends who encouraged me in my walk with the Lord and my pursuit of higher education.

The author would like to thank National Oceanic and Atmospheric Administration for supporting this project financially (Award #NA19OAR4590411, awarded to Mississippi State University). In addition to that, the author would like to thank Mississippi State University High Performance Computing Collaboratory for providing computational resources. Finally, the author would like to than NASA Goddard Earth Science Data and Information Service Center for providing MODIS land cover data.

The author would like to acknowledge AmeriFlux network data providers for providing this study with the surface flux and soil moisture data. In addition to that, the author would like to thank NASA Goddard Earth Science Data and Information Service Center for providing NLDAS-2 data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Physical characteristics of the land surface have various impacts on climate by altering land surface albedo, evapotranspiration, and surface roughness that in turn alter atmospheric circulations, energy budgets, and hydrologic cycles. Thus, in atmospheric modeling, land surface characteristics play a key role in partitioning the energy and moisture fluxes between the land and atmosphere. As a result, improved characterizations of land surface characteristics within numerical weather prediction (NWP) models should lead to increased accuracy of simulations. This is especially true during the growing season over agricultural areas, when increased vegetation coverage enhances evapotranspiration and associated processes such as convective rainfall. Results from global modeling studies indicate a change in surface temperatures due to deforestation, but the impacts of land use and land cover (LULC) vary by region and season. In addition, studies showed the impacts of land cover modification on the spatiotemporal changes of precipitation, air temperature, surface heat fluxes and moisture flux (Bin Li & Avissar, 1994; Jianduo Li et al., 2018; Qu, Weiqing, 1998; Z. Yang & Dominguez, 2019; W. Zhang et al., 2016). Although many investigations assessed the sensitivity of dominant land cover changes to numerical weather models such as the Weather and Forecasting (WRF) model (Bin Li & Avissar, 1994; J. Dyer, 2011; Paul et al., 2016), it is important to also consider the percent coverage of all land cover categories to provide a robust analysis of the impacts of real-world land cover and vegetation types on land surface processes. In comparison with the current land cover data in the

WRF model (Al, 2017; Kirthiga & Patel, 2018; Jiayang Li et al., 2018; Schneider et al., 2004), the near-real-time land cover data add more representative land surface characteristics to the model simulations. This study evaluates the impact of the near real-time LULC data derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) observations on the performance of the WRF model to clearly understand the influence of land use on surface processes with the aim of improving surface-related weather simulations over the southeast United States. The simulation experiments involve two independent model runs: (1) standard WRF land cover (default simulation), and (2) updated land cover from MODIS (updated simulation) for a dry year (2007), a normal year (2014), and a wet year (2018). Further, all simulations were performed over the growing season (April 15 – October 1), with data from May 1 – October 1 being used for analysis (15-day model spin-up) over a 4-km nested domain.

Numerical weather prediction models simulate atmospheric and surface parameters using dynamic mechanisms between the land and atmosphere (Yoon et al., 2021). However, course horizontal resolution of these models often introduces prediction error into the model simulations. In addition, errors and uncertainties in NWP models occur in response to errors in land surface model (LSM) outputs, which are used to simulate land surface processes within numerical weather and climate models. Although updated LULC can provide more representative land surface properties, there are other unknown factors (Liu et al., 2016; G. Niu et al., 2011) in land-atmosphere systems that affect the accuracy of the simulations in both magnitude and sign. The relationship between NWP simulation and observed historical data at weather stations is used to calibrate the weather simulations. In recent years, statistical learning model are used to leverage accuracy of numerical land surface model simulations. Machine learning models are used to post-process Noah-MP land surface model outputs for improved hourly surface heat fluxes, soil

moisture, and soil temperature at AmeriFlux (Workshop, 2015) towers across the continental United States. The Noah-MP (Levis et al., 2012; Liu et al., 2016; Tsvetsinskaya et al., 2001) is a physics-based land-atmosphere simulator which is used within the NWP and climate models. The Noah-MP model outputs were used as input variables, and observed sensible heat flux, latent heat flux, soil moisture and soil temperature data were used as target features to train the ML models. Noah-MP model outputs, Noah-MP bias corrected outputs, and machine learning estimations were compared against the AmeriFlux observations.

Overall, the key objectives of this study are to investigate the importance of the LULC change within the WRF simulations to better understand the surface processes and its interactions with the near surface atmosphere. Since the land surface processes are the most challenging parts of regional climate modeling, better understanding of sensitivity of the surface heat and moisture flux simulations to land surface characteristics can be used to improve the weather predictions. Therefore, the potential improvements in WRF-simulated temperature and precipitation skills as a result of updating LULC was investigated. In addition, machine learning models were used to improve Noah-MP surface heat fluxes, soil moisture and soil temperature simulations relative to available surface observations. Although the findings of this study showed strong sensitivity of WRF simulations towards the representation of surface characteristics, updating LULC did not improved the temperature and precipitation simulations over all stations. Results indicated that post-processing Noah-MP simulations using machine learning substantial improved simulation accuracy. The findings of this study can be used to improve the NWP and land surface model simulations to and enhance numerical weather predictions.

CHAPTER II

EVALUATING THE SENSITIVITY OF SURFACE-ATMOSPHERE INTERACTIONS TO LAND USE AND LAND COVER CHANGES USING WRF SIMULATIONS

Introduction

The Weather Research and Forecasting (WRF) model (Janjic et al., 2010) is an open-source atmospheric modeling system that is widely used for simulating and understanding land-atmosphere interactions (Al, 2017; Kirthiga & Patel, 2018; Jiayang Li et al., 2018; Schneider et al., 2004). Despite its popularity in dynamic downscaling of regional climate, the model utilizes a long-term average of land cover data as a static input for surface boundary conditions. In addition, the land-atmosphere simulations by the land surface models within the WRF system mostly rely on the dominant land use category or the land use fraction, which are used as static inputs in the WRF simulations (Mallard et al., 2018). The details of the land surface models are discussed in Chen & Dudhia (2001a). Land cover characteristics play a key role in controlling surface energy partitioning in the earth-atmosphere interactions, hence, this input can have an impact on convective precipitation generation (Holt et al., 2006). Vegetation type, for example, is one of the parameters that defines the canopy resistance in the WRF Noah land surface model (Noilhan, J., 1989), and Kishtawal et al. (2010) showed that changes in surface roughness and stomatal resistance caused 2-m air temperature to increase by 0.3° – 0.4°C in the eastern United states. Hence, accurate and more realistic representation of the physical characteristics of the land surface is critical for simulating land-atmosphere exchanges of heat and moisture fluxes within numerical

weather simulation models. This is especially important for agricultural areas, such as the Lower Mississippi River alluvial valley, where land cover fractions may change in the region due to variations in the distribution of active agricultural land, natural grasslands, and forested area.

Several researchers have presented evidence of land cover modification on the spatiotemporal changes of precipitation, air temperature, surface heat fluxes and moisture flux (Bin Li & Avissar, 1994; Jianduo Li et al., 2018; Qu, Weiqing, 1998; Z. Yang & Dominguez, 2019; W. Zhang et al., 2016). Although many investigations assessed the sensitivity of the WRF model to dominant land cover changes (Bin Li & Avissar, 1994; J. Dyer, 2011; Paul et al., 2016), it is important to also consider the percent coverage of all land cover categories to provide a robust analysis of the impacts of real-world land cover and vegetation types on land surface processes.

The current land use data set used in the WRF model is based on a long-term climatological dataset (Kumar et al., 2014), which cannot well represent the annual changes of land use and land cover (LULC) and therefore can negatively impacts the accuracy of regional-scale simulations (H. Li et al., 2020). The Moderate Resolution Imaging Spectroradiometer (MODIS), a multispectral sensor mounted on the AQUA and TERRA satellites, provides detailed information of the land surface that enables the weather research community to address a much larger range of research questions associated with surface biochemical cycling and land cover change (Justice et al., 2002). High spatial and temporal resolution, as well as the variables included in the MODIS land products, such as land cover, vegetation, and albedo, make it a valuable sensor in land-atmosphere studies (A. Kumar, et al., 2014).

The impact of LULC changes on regional and global climate is one of the most challenging aspects of understanding anthropogenic climate change (Trail et al., 2013). The specific effects of land cover change also depend on the type of change, the spatial scale of assessment, the size of

the area under consideration, and the region where the change occurs (Pielke et al., 2011). Land use and land cover change involve changes in the biophysics of the terrestrial surface, and therefore can impact regional climate through direct alternations in surface radiation and boundary layer turbulence. These factors then result in changes in the fluxes of momentum, heat, water vapor, and carbon dioxide between the surface and the atmosphere. Even small-scale changes in land cover due to agricultural activities (J. L. Dyer & Rigby, 2020) and urbanization (X. Li et al., 2018) can influence energy and moisture balance at the surface and consequently atmospheric processes. As a result, to incorporate the effects of the interannual and seasonal changes in land cover within numerical weather prediction (NWP) models, utilizing near real-time land cover data within the weather simulations is critical to properly evaluate land surface-atmosphere interactions.

Previous studies have attempted to enhance the skill of WRF forecasts by modifying land use and vegetation data using improved satellite products (Kumar et al., 2014; J. Yang et al., 2018); however, it is important to recognize the specific influence of LULC changes on model simulations to better understand the role of LULC over regions sensitive to surface characteristics. As a result, the objective of this study is to quantify the sensitivity of surface heat fluxes and moisture fluxes using a high-resolution numerical weather simulation framework with updated land cover (dominant and fractions) over the Continental US (CONUS), with a focus on regional patterns over the Southeastern US. Spatial and temporal analysis will be caried out to compare the regional weather simulations with default and updated land cover to evaluate the impacts of integration of the MODIS real-time land cover on the WRF simulations. The results of the study will provide valuable information to quantify the impacts of using up-to-date satellite-derived information in weather simulations focused on surface-atmosphere interactions and processes.

Data and Methods

MODIS land cover data

The combined Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) land cover data (MCD12Q1; Version 6) product was used to update the WRF default land cover types at annual time steps. These data were derived from MODIS reflectance data at 500-m spatial resolution using the random forest supervised classification method (S. Cai et al., 2014; Friedl et al., 2010). In addition to the reflectance dataset, the MODIS land-water mask product (MCD44W), vegetation continuous fields product (MCD44B), global cover type map, MODIS global map of urban areas, and land surface temperature were included as the input features to produce the MODIS land cover product (Sulla-Menashe et al., 2019). Random forest classifier, like its name implies, contains many individual decision trees that function as an ensemble. It creates a set of decision trees from randomly selected subsets of the training set, then performs a vote for each predicted result and aggregates the votes to decide the final class of the test set. The class with the most votes become the final prediction. The MCD12Q1 product provides annually updated land cover type since 2001, which is immensely important for a variety of regional and global studies (Aylas et al., 2020; Gogoi et al., 2019; H. Li et al., 2020). To quantify the uncertainties in MODIS land cover data, a set of validation sites across the world have been used, which covers a range of climate zones and major biomes (Myneni et al., 2002; Running et al., 1999). Moreover, high-resolution imagery data such as Landsat, ETM+, and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) are used to validate MODIS land data (Morisette et al., 2002).

The MODIS land cover product provides five classification schemes including International Geosphere-Biosphere Programme (IGBP), University of Maryland (UMD), Leaf

area index (LAI), BIOME-Biogeochemical Cycles (BGC), Plant Functional Types (PFT), and Land Cover Classification System (LCCS) whereby land cover was mapped using different classification systems. Table 2.1 shows five different classifications of the MCD12Q1 product and associated land cover types for each classification along with the WRF default IGBP MODIS land use classes. More details can be found in the user guide to collection six of the MCD12Q1 product (Sulla-Menashe, et al., 2018). Compared with MODIS IGBP land use classification, there are four more land use classes in the WRF IGBP land use classification including wooded tundra, mixed tundra, barren tundra, and lakes (Table 2.1). The WRF simulations were divided into two groups; the first simulation, referred to as “default”, is conducted with the WRF default IGBP MODIS land cover data, and the latter, referred to as “updated”, is conducted with the MODIS-derived IGBP annual land cover data.

Table 2.1 MODIS MCD12Q1 product classifications compared with the WRF IGBP MODIS land use classifications.

	WRF	MODIS				
Class number	IGBP	IGBP	UMD	LAI/FPAR	BGC	PFT
1	Evergreen needleleaf forest	Evergreen needleleaf forest	Evergreen needleleaf forest	Evergreen needleleaf forest	Evergreen needleleaf forest	Evergreen needleleaf forest
2	Deciduous needleleaf forest	Deciduous needleleaf forest	Deciduous needleleaf forest	Deciduous needleleaf forest	Deciduous needleleaf forest	Deciduous needleleaf forest
3	Evergreen broadleaf forest	Evergreen broadleaf forest	Evergreen broadleaf forest	Evergreen broadleaf forest	Evergreen broadleaf forest	Evergreen broadleaf forest
4	Deciduous broadleaf forest	Deciduous broadleaf forest	Deciduous broadleaf forest	Deciduous broadleaf forest	Deciduous broadleaf forest	Deciduous broadleaf forest
5	Mixed forest	Mixed forest	Mixed forest	-	-	-
6	Closed shrublands	Closed shrublands	Closed shrublands	Shrublands	-	Shrub
7	Open shrublands	Open shrublands	Open shrublands	Shrublands	-	Shrub
8	Woody Savannas	Woody Savannas	Woody Savannas	Savannas	-	-
9	Savannas	Savannas	Savannas		-	-
10	Grasslands	Grasslands	Grasslands	Grasses/ cereal crop	Annual grass vegetation	Grass
11	Permanent wetlands	Permanent wetlands	-	-	-	-
12	Croplands	Croplands	Croplands	Broadleaf crops	Annual broadleaf vegetation	Cereal crop
13	Urban and build-up land	Urban and build-up land	Urban and build-up land	Urban	Urban	Urban
14	Cropland/natural vegetation mosaics	Cropland/natural vegetation mosaics	-	-	-	Broadleaf crop

Table 2.1 (continued)

15	Permanent snow and ice	Permanent snow and ice	-	-	-	Snow and ice
16	Barren or sparsely vegetated	Barren or sparsely vegetated	Barren or sparsely vegetated	Unvegetated	Unvegetated	Barren or sparsely vegetated
17	Water	Water	Water	Water	Water	Water
18	Wooded Tundra	-	-	-	-	-
19	Mixed Tundra	-	-	-	-	-
20	Barren Tundra	-	-	-	-	-
21	Lake	-	-	-	-	-

The MODIS data was used since one of the land cover data ingested by the WRF model is based on MODIS land cover classification scheme. The MCD12Q1 data was downloaded from the NASA Earth sciences data and information service center (GES DISC) website (https://earthdata.nasa.gov), which provides annual MODIS land cover data at 500-m on a sinusoidal grid. This product was then reprojected to the WGS84 geographic projection, and then exported in binary format so that they could be used by the WRF Preprocessing System (WPS) and transferred to the WRF grid. The majority resampling method (mode) was used because the land cover data are categorical data. Along with the other gridded geostatic data, such as soil type, soil temperature, terrain height, vegetation fraction, leaf area index, and albedo, the WPS uses dominant land cover category and land cover fractions to represent the land surface physical characteristics in the WRF simulation grid. After generating WPS geogrid files for updated and default data, all grids containing the wooded tundra, mixed tundra, barren tundra, and lake categories in the default data were modified to have the same number of categories in the updated data to be consistent between simulations. It should be noted that although there are three additional vegetation categories (wooded tundra, mixed tundra, and barren tundra) in the WRF default IGBP global land cover data, the WPS generated files did not contain these three land cover types within the southeast US study area; therefore, only the lake grids (category 21) within the default data were updated to category 17 (water) in the updated land cover data.

WRF setup and study area

To identify the impacts of land cover changes on surface-atmosphere interactions and near-surface atmospheric conditions, the Advanced Research WRF (ARW) model, v.4-3 (Skamarock et al., 2021) was used to simulate meteorological conditions over the southeast and continental United States using 4-km (380 × 380 grids) and 12-km (230 × 390 grids) spatial resolution for the

nested and parent domains, respectively (Fig. 2.1). A 4-km spatial resolution was chosen as it provides adequate detail of land use variability and changes in the default and updated simulations while allowing convection to be directly simulated. The options used were 45 vertical levels arranged along a logarithmic profile with the following parameterizations: Rapid Radiative Transfer Model for GCMs (RRTMG) for longwave and shortwave radiation (Iacono et al., 2008); Noah land surface model (F. Chen & Dudhia, 2001a, 2001b); New Thompson scheme for microphysics (Thompson et al., 2008); and Mellor-Yamada-Janjic scheme (Janjić, 1994) for planetary boundary layer. Further references for these parameterization schemes can be found in (Dyer & Rigby, 2020; Skamarock et al., 2021)

Focusing on the annual percent precipitation over the southeast, the simulations were carried out separately for a dry year (2007), a normal year (2014), and a wet year (2018) from 1 May through 30 September, with the model runs starting on 15 April to allow proper spin-up time for the WRF momentum variables. A year whose annual precipitation was greater than the percent of normal precipitation (defined as average precipitation since 2005; NOAA 2022) were identified as wet year, while a year that was less than the percent of normal precipitation was identified as dry year (Fig. 2.2). The wet and dry years were selected to examine how the WRF-simulated surface heat fluxes and soil moisture are influenced by variable hydrologic conditions. In addition, the simulation period was selected since it covers the majority of the growing season for warm-season crops, which is when surface conditions have the largest impact on lower atmospheric processes. The initial and boundary atmospheric and soil conditions were provided by the National Center for Environmental Prediction (NCEP) 0.5-degree Global Forecasting System (GFS) analysis fields, from which the daily 0000 UTC data were utilized. These data are commonly used

for regional weather studies (López-Bravo et al., 2018; Ngan et al., 2015) over different locations, and are therefore considered appropriate for this study.

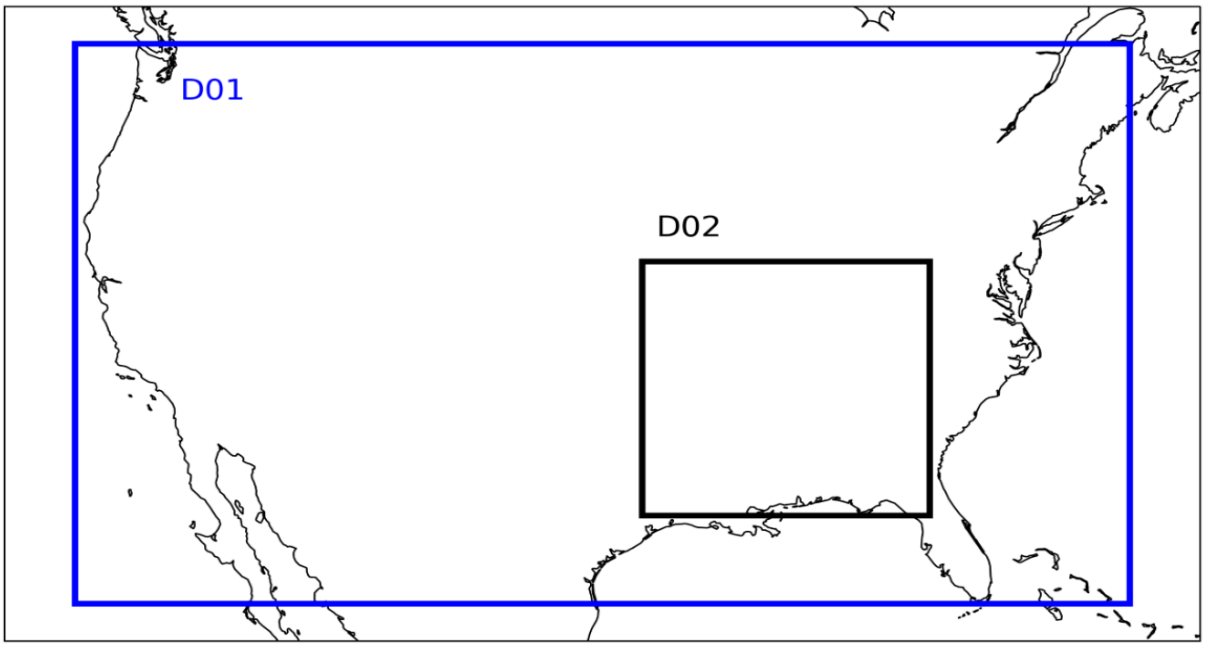

Figure 2.1 WRF model domains with 12-km spatial resolution and 230 × 390 grid points in the north-east and east-west, respectively, for the parent domain (blue), and 4-km resolution and 380 × 380 grid points for the nested domain (black).

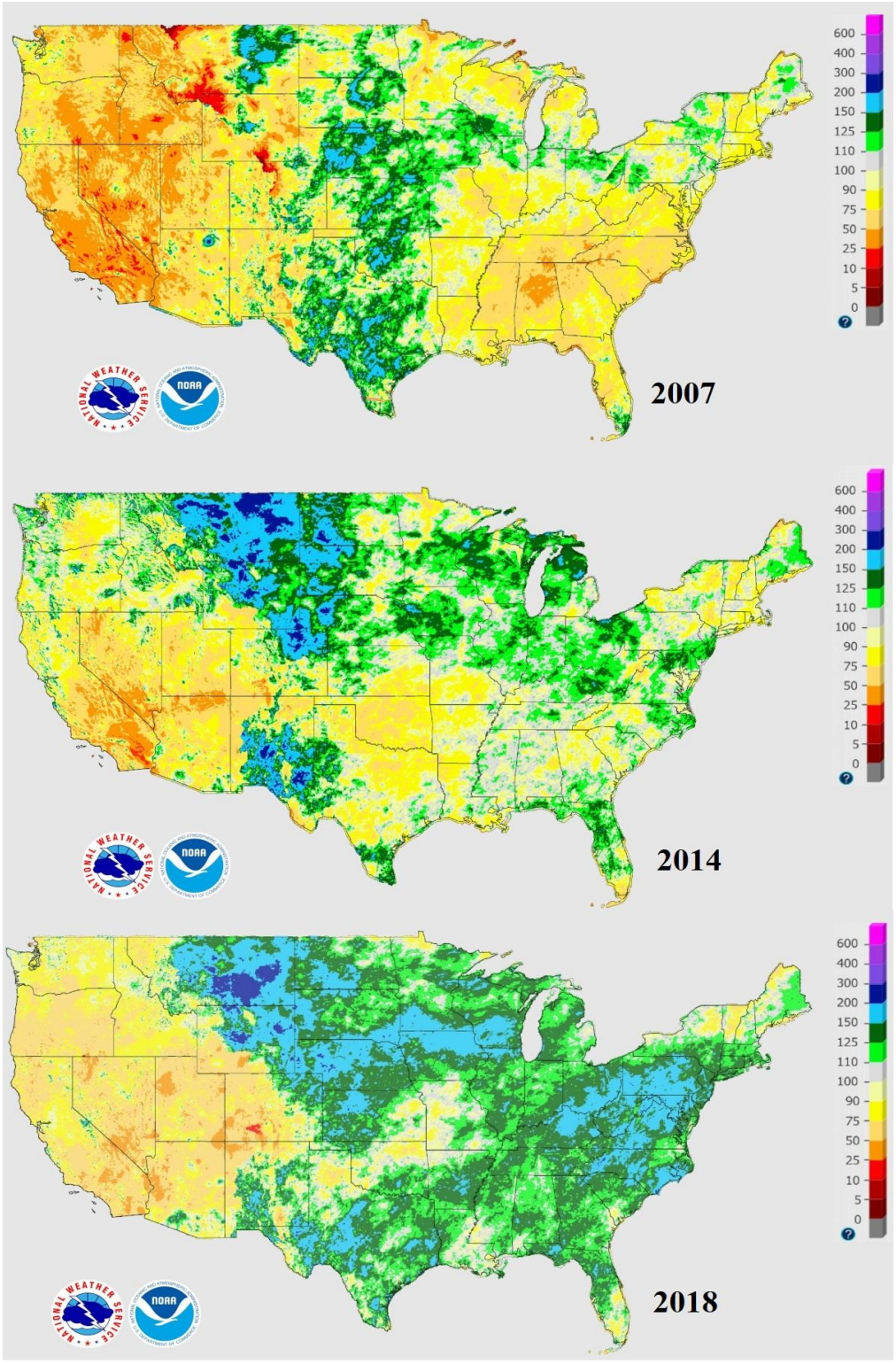

Figure 2.2 Percent of normal precipitation (defined as average precipitation since 2005; NOAA 2022) that occurred for the calendar year in each simulation year, departure from the climatological average.

For quantification of the impacts of updated LULC on 2-m average air temperature (T2), 2-m water vapor mixing ratio (Q2), and soil moisture at 30-cm, the default simulation was subtracted from the updated simulation of each variable over all grid points for all forecast hours. Along with the spatial pattern showing differences in the magnitude of the updated and default runs, the mean was used to evaluate the difference between the two simulations. Moreover, Kolmogorov-Smirnov (K-S) test was applied to determine if output from the updated WRF simulation was significantly different ($p < 0.05$) than output using the updated land cover. The K-S test was used as it does not assume a normal distribution, with the test statistic calculated under the hypothesis that the two samples being compared are from the same distribution. The two-sample K-S test is one of the most useful and general non-parametric methods for comparing two samples, as it is sensitive to differences in both the location and shape of the empirical cumulative distribution functions of the two samples. The analysis was performed over the 4-km nested domain and associated area of 12-km parent domain over the southeast to determine the impact of horizontal spatial resolution of LULC on the land surface simulations. To determine the difference in the actual values of the updated and default WRF model simulations, histograms were generated from the temporally accumulated precipitation, surface heat fluxes, and temporally averaged of air temperature, soil moisture and mixing ratio to represent the difference between the distribution of each variable over the study area. Along with comparing histograms of select variables between the default and updated WRF model simulations, the two-sample Kolmogorov-Smirnov (K-S) test was applied to determine if output from the updated WRF simulation framework was significantly different ($p < 0.05$) than output using the updated land cover. The K-S test was used as it does not assume a normal distribution, with the test statistic calculated under the hypothesis that the two samples being compared are from the same distribution. The two-sample K-S test is one of the

most useful and general non-parametric methods for comparing two samples, as it is sensitive to differences in both the location and shape of the empirical cumulative distribution functions of the two samples. The analysis was focused on the southeast because of the substantial changes in the dominant LULC observed over this region in the updated MODIS LULC data relative to the default data (Fig. 2.3-c, -d). In addition, the widespread agricultural activities in the area as well as the multiple transitions to (and variation in) forested and natural landscapes make LULC changes a potential driving factor in local-scale weather processes.

Results and discussions

Land cover changes

The area under different LULC categories showed noticeable changes between the updated and default land cover grids (Fig. 2.3-a, -b), where reductions were observed for Cropland/Natural Vegetation Mosaics (Southeast, Northeast, Great Lakes), Evergreen Needleleaf Forests (Florida), Mixed forests (Southeast, New England, and Great Lakes) and Open Shrublands (Southeast) in the updated LULC data compared with default WRF land cover data. In addition, the updated LULC data showed increases over some regions in Grasslands (Nevada, Utah, New Mexico), savannas (California), and woody savannas (California, East Texas, South Mississippi and Alabama). It should be noted that these patterns were consistent for all years. Areas where woody Savannas increased in the updated data (~1,300,000 km^2) were covered mostly by Cropland (360,000 km^2), Evergreen Needleleaf Forests (340,000 km^2), and Mixed Forests (310,000 km^2) in the default data. According to the attribute table of the WRF vegetation parameters, the magnitude of the maximum albedo values associated with these land cover categories are 0.23, 0.12, and 0.25, respectively, which are less than the value of the albedo associated with woody Savannas (0.3).

Furthermore, the maximum leaf area index (LAI) of woody Savannas (3.6) is less than the other three land cover categories (5.7, 5.5, and 6.4, respectively). Since albedo and LAI are correlated variables (Trail et al., 2013), the higher the LAI, the more green area to absorb sunlight. Investigating the combined effects of the LAI and stomatal resistance is beyond the scope of this study; however, it plays an important role in the near-surface atmosphere (Jefferson et al., 2017; Martin et al., 1999; Niyogi et al., 1998); therefore, changes in vegetation cover can act as direct controls on the other parameters at the Earth's surface such as evaporation and soil moisture content. The area covered by each land use category in the updated data compared with the default LULC over the nested domain (Fig. 2.3-c, -d) showed that Croplands/Natural vegetation Mosaics, Grasslands, and Mixed Forest decreased in the updated data relative to the default data.

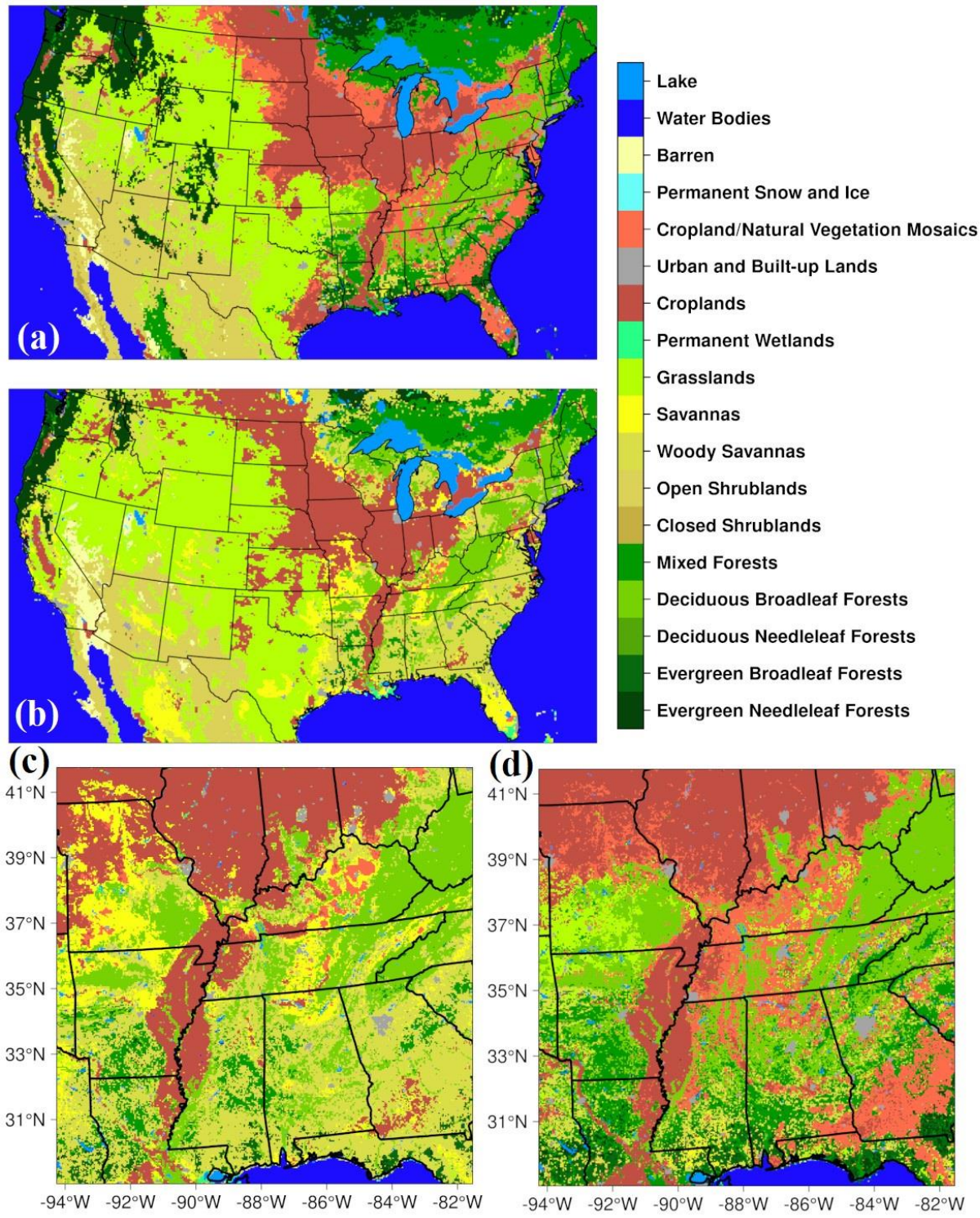

Figure 2.3 Maps of MODIS 2018 land use (a, c) and WRF default MODIS land use (b, d) data (right) for 12-km parent domain (a, b), and 4-km nested domain (c, d).

The study area is a region of approximately 1,617,984 km^2 covering the southeast US (29.0° - 42° N, 94° - 82°W) and characterized by rapid transitions from Cropland/Natural Vegetation Mosaics and mixed forests to woody savannas and savannas. In addition, LULC changes in this area are evident due to the widespread agricultural activities. Although the updated LULC grids from 2007, 2014, and 2018 did not show appreciable differences in dominant land cover between years (figures not shown), changes were detected for woody savannas (-36,000 km^2), and Mixed Forests (+29,000 km^2) from 2007 through 2018 in the updated data over the nested domain. Figure 2.4 shows the area of each LULC in the parent domain (a) and nested domain (b) in the default and updated data. It should be noted that these results are based on the 12-km and 4-km spatial resolution of the WRF domains; however, generating higher spatial resolution, for example 1km, of WPS geostatic data may represent detailed information about local-scale land cover changes due to agricultural activities, deforestation, construction, and wildfire events. Further comparison was performed on the area of dominant LULC over the nested domain and associated areas in the parent domain in the default data to show how the dominant land cover are represented at various spatial resolutions. As shown in Fig. 2.4-c, the area of the Cropland/Natural Vegetation Mosaics (16,752 km^2), Croplands (13,856 km^2), Deciduous Broadleaf Forests (11,680 km^2), Evergreen Broadleaf Forests (3,136 km^2), and Mixed forests (7,136 km^2) is higher in the 12-km domain relative to the 4-km domain; however, the areas of Evergreen Needleleaf Forests, Grasslands, and Lakes (~10,000 km^2), and woody savannas (17,600 km^2) is lower in the parent domain compared with the associated areas in the nested domain. Although the difference of areas of the various LULC categories at the different spatial resolutions is roughly 1% relative to the total area of the nested domain, 4-km land use provides much finer details for the Noah land surface model used in the WRF simulations.

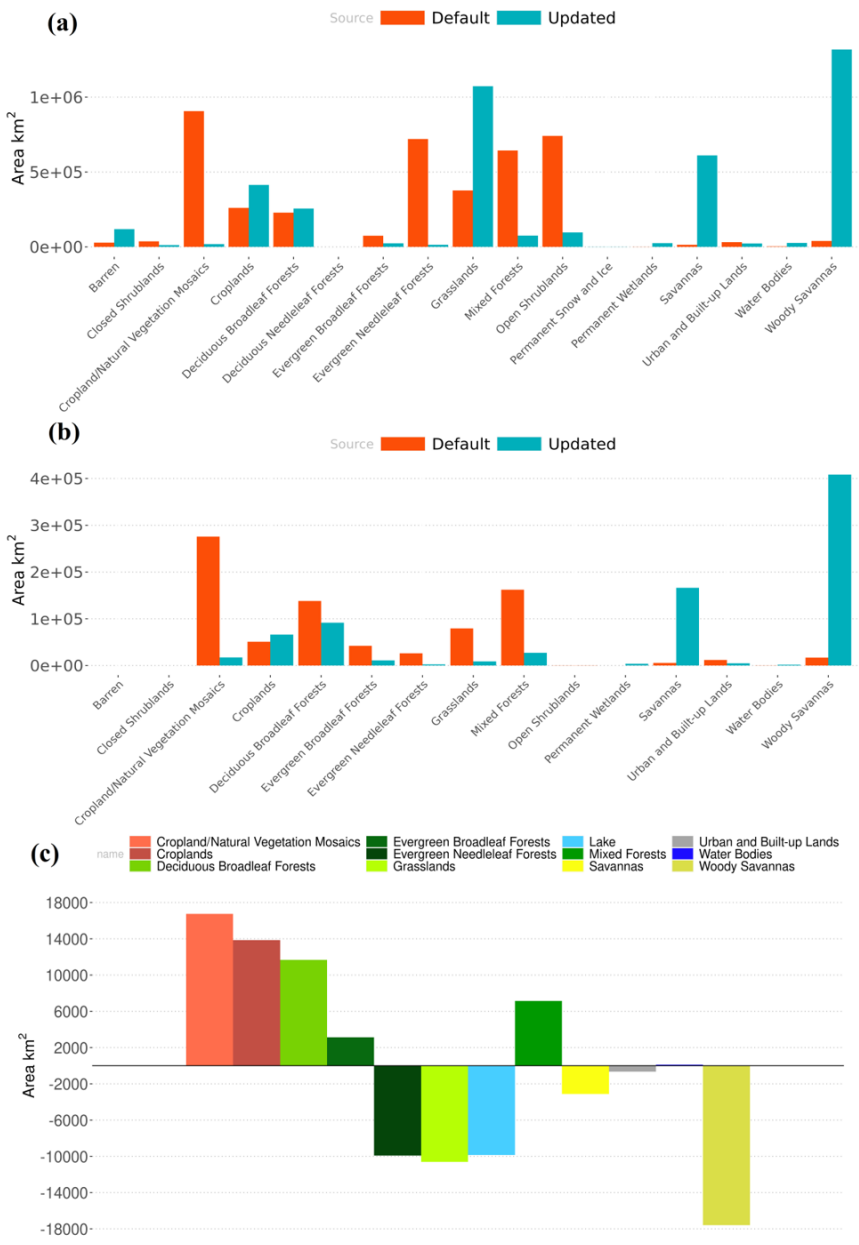

Figure 2.4 Differences between the WRF default MODIS land cover and 2018 MODIS land cover data in the parent domain (a) and nested domain (b). Panel (c) shows the difference (parent - nested) of area (km^2) of each land cover category over the nested domain relative to the associated area of the parent domain.

Surface heat fluxes

The average changes in sensible heat flux (HFX) between the updated and default (updated - default) LULC over the parent 12×12-km domain and nested 4×4-km domain (Fig. 2.5-a and 5-b, respectively) from May through September indicated that the updated simulation has a higher sensible heat flux relative to the default simulations over the southern portion of the area for the three study years. The spatial variability of sensible heat flux is high over the domains, approximately 30 W/m^2 increase was evident over the Ozarks in 2007, and Alabama, Georgia, and South Carolina at 12-km simulations during all three study years. These areas mostly correspond to woody savannas and cropland in the updated and default LULC respectively (Fig. 2.3-c, -d). The result indicated an area of maximum decrease (> 5 W/m^2) in southeast Missouri, south Indiana, and Ohio in 2014, and over the northern portions of the lower Mississippi River alluvial valley (LMRAV) in 2018 in the parent and nested domains. Time series of sensible heat flux were extracted from WRF updated and default simulations over the grid points and a two-sample K-S test was then applied to the extracted time series to determine if the sensible heat flux changes is statistically different. Results indicated that the decreased sensible heat flux simulations as a result of LULC change was statistically significant ($p < 0.05$) in many grids across the study area (Fig. 2.6-a). In contrast to the 12-km nested domain, sensible heat flux differences were smaller in 4-km simulation, and in the updated 4-km simulation relative to the default 4-km simulation the simulation was by approximately -20 W/m^2 smaller over central Kentucky in 2014; however, it was larger in the updated data in the 12-km simulation.

The results show that the spatial extent of average HFX from the 4-km simulations was relatively lower in all years south of 35°N. For instance, there is a noticeable increase (> 25 W/m^2) in the HFX over central and southern Georgia within the 12-km model output in 2014; however,

an area of increased sensible heat flux greater than 25 W/m^2 is apparent only along central and into northeast Georgia within the 4-km simulation. Comparing the default and updated LULC showed that the extreme negative values of the HFX differences were associated with changes from open shrublands (default) to barren (updated). Figure 2.5-b shows more localized average changes in sensible heat flux between the updated and default LULC from May through September for the three study years. In addition, the sensible heat flux based on the WRF simulations for dry, normal, and wet years indicated that the difference between the default and updated simulations is maximized in both intensity and spatial extent for the dry year (2007) relative to the wet year (2018) over the Ozarks and Southeast United States. One possible explanation for the changes in the surface heat flux is that the changes in LULC lead to changes in soil moisture availability and surface heat capacity, which are two main factors that influence sensible heat flux (Esteve, 2015). For example, according to the WRF land use table, the magnitude of soil moisture percentage and surface heat capacity for open shrublands are 0.15% and 20.8×10^5 J m^{-3} K^{-1}, respectively, and associated values for the barren or sparsely vegetated land cover are 0.02% and 12.0×10^5 J m^{-3} K^{-1}. As a result, decreasing soil heat capacity and soil moisture availability would result in reducing evapotranspiration and increasing surface temperature.

The changed surface physical characteristics based on the updated MODIS land cover data within the WRF simulations leads to ~30W/m^2 decreases in the mean surface latent heat fluxes during the years studied (Fig. 2.5-c, -d). Similar to HFX, the areas with higher decreases in latent heat flux (>30 W/m^2) were mostly seen over the southern portion of the 4-km nested domain, which is mostly associated with transitions from croplands in the default data to woody savannas in the updated land use data. A two-sample K-S test results indicated that the changes in latent heat flux as a result of LULC change was statistically significant ($p < 0.05$), specially over the

Mississippi, Alabama, and Georgia (Fig. 2.6-b). Moreover, areas in south Illinois and Indiana showed an increase (>10 W/m^2) in latent heat flux in the 4-km simulation in 2014, and there was an additional increase over southeast Missouri in 2018. The areas with higher decreases in latent heat flux (<30 W/m^2) are associated primarily with transitions from cropland/natural vegetation mosaic and mixed forests land use categories in the default LULC to woody savannas in the updated LULC (Fig. 2.3). In general, modeled latent heat fluxes with updated LULC were lower relative to those with default LULC by approximately 14% through the whole nested domain. As a result, according to the WRF vegetation attributes table, decrease in the latent heat flux could be due to decrease in the leaf area index for woody savannas (min=0.5, max=3.6) compared with cropland/natural vegetation mosaic (min=2.3, max=4.3) and Mixed Forests (min=2.8, max=5.5). The stomatal resistance of woody savannas (300 s/m) is higher than the associated value for cropland/natural vegetation mosaic (40 s/m), which should act to decrease the magnitude of evapotranspiration; however, it should be noted that canopy resistance is affected by the other environmental parameters such as incoming solar radiation, air temperature, and vapor pressure (Chang et al., 2020). Another possibility for the increase in the latent heat flux could be due to a horizontal temperature gradient resulting from the sharp spatial transition in LULC (Dyer and Rigby, 2020). Further analysis of leaf area index, stomatal resistance, and green vegetation fraction is needed to provide additional interpretation of this process.

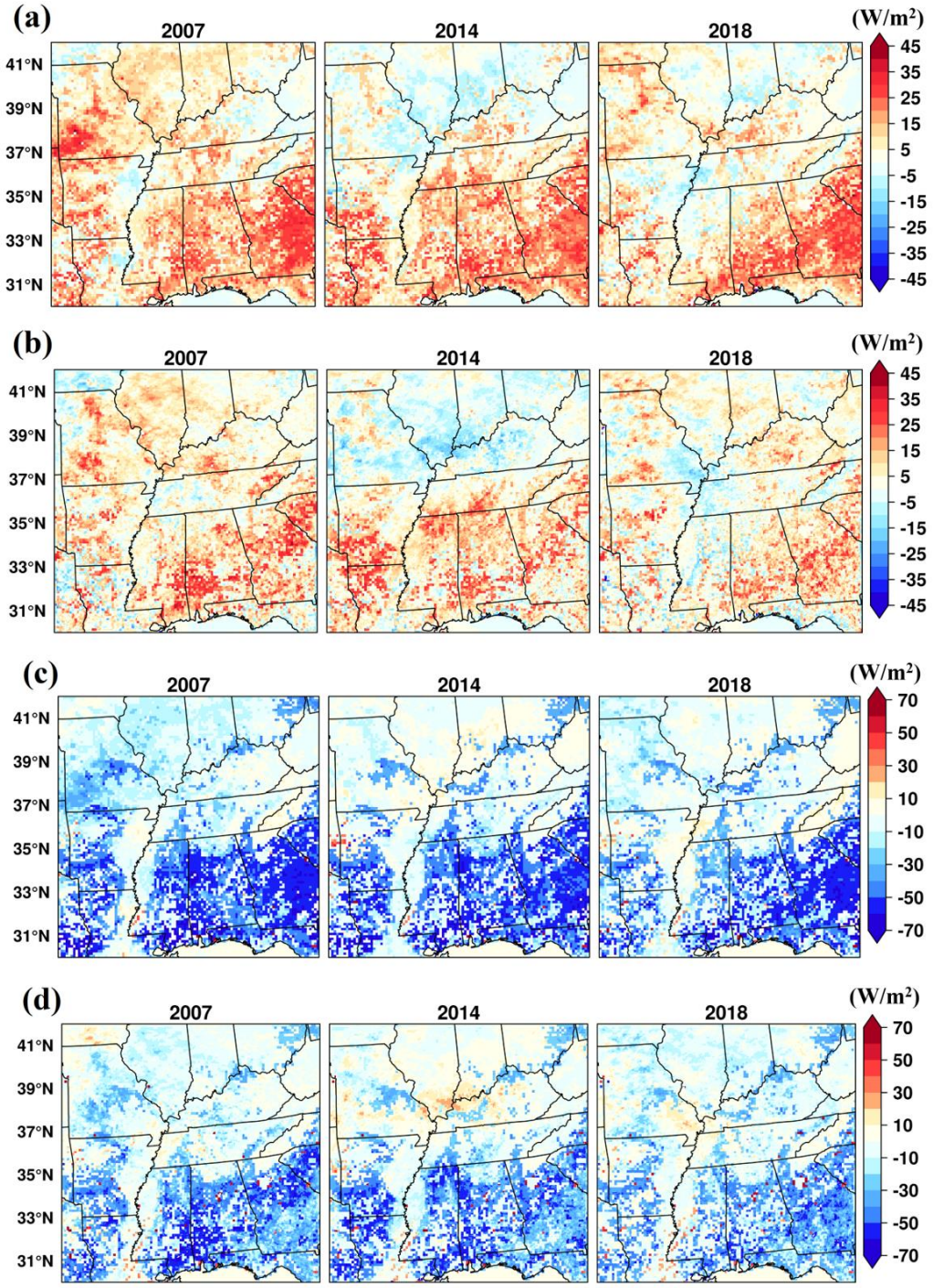

Figure 2.5 Average of sensible heat flux (a-b) and latent heat flux (c-d) differences across the three simulation years analyzed over 12-km nested domain (a and c) and 4-km parent domain (b and d). Differences were calculated as time series of the updated simulations minus the default simulations over the period from May – August.

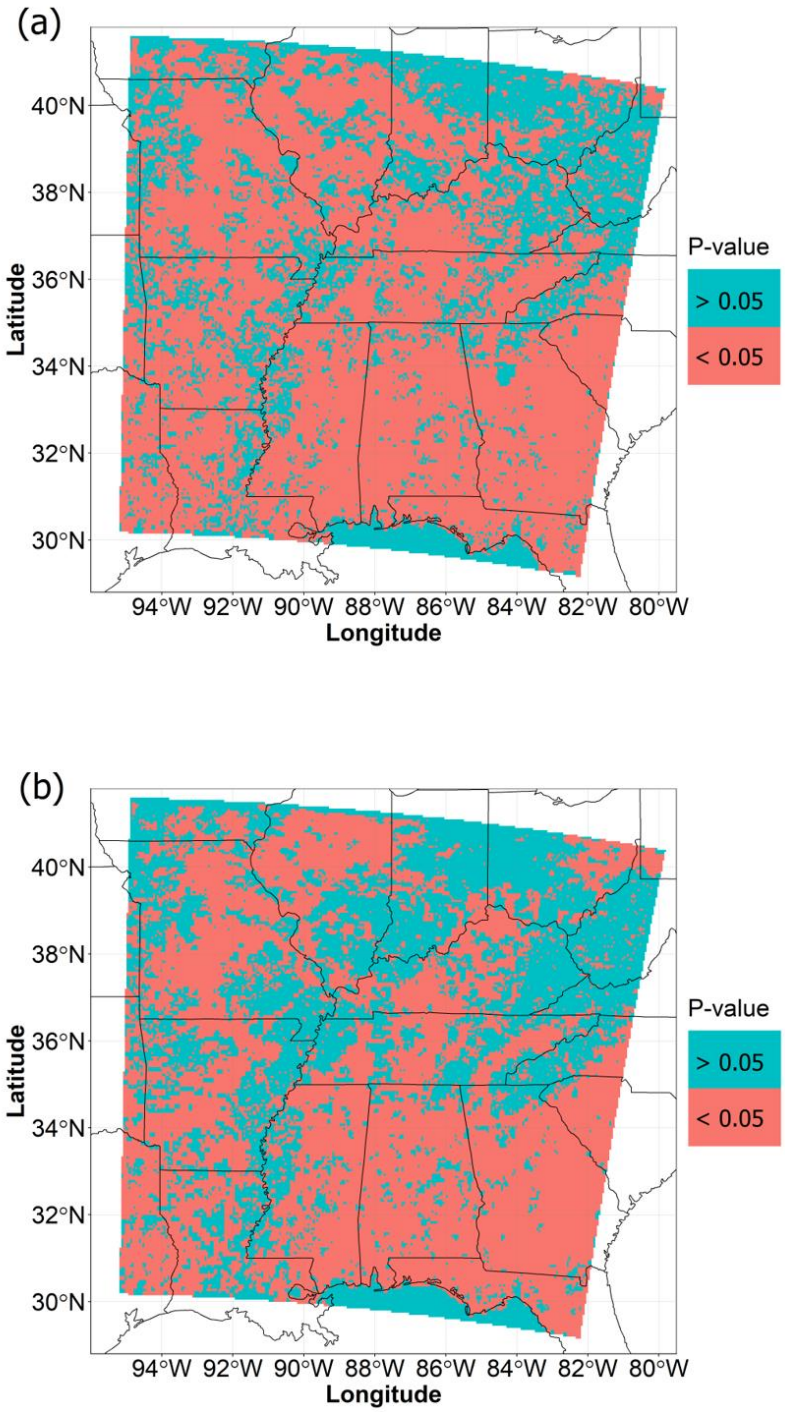

Figure 2.6 K-S two-sample test results of sensible heat flux (a) and latent heat flux (b) in 2007 analyzed over 4-km nested domain. This test applied on the timeseries of each grid individually.

Given the variability of surface heat fluxes between the default and updated simulations over the southeast, additional comparisons were performed on the daily maximum values of the surface heat fluxes for the 2014 simulation to highlight the impact of LULC changes on extreme values. Results indicated that the average of maximum sensible heat fluxes was 306 W/m^2 and 330 W/m^2 in the default and updated simulations, respectively. Figure 2.6a shows the difference of spatially averaged daily maximum sensible heat flux over the southeast in 2014, where from May through July, daily maximum sensible heat flux with the updated LULC was approximately 20 W/m^2 (± 20 W/m^2) higher relative to the default case. From May through July, the standard deviation of differences was lower relative to the period from July through September; however, from July through September the differences had an average of 0 W/m^2. In addition, the simulated latent heat flux (updated) was approximately -50 W/m^2 lower than the default simulation during May and July (Fig. 2.6b) which was statistically significant ($p < 0.05$), and the standard deviation of differences between the simulations was relatively high. After this period, the difference between the simulations was lower than the period of May through July. During June and July when insolation is maximized and potential evapotranspiration is high, leaf area index and the availability of soil moisture are primary factors in defining the intensity of the latent heat flux. During September, as sun angle decreases and available surface energy goes down, and leaf area index also decrease due to seasonal vegetation patterns, the potential evapotranspiration is lower relative to June and July. Therefore, seasonal variability of vegetation and land surface representation can modulate the partitioning of available energy between the land surface and atmosphere (Matsui et al., 2005).

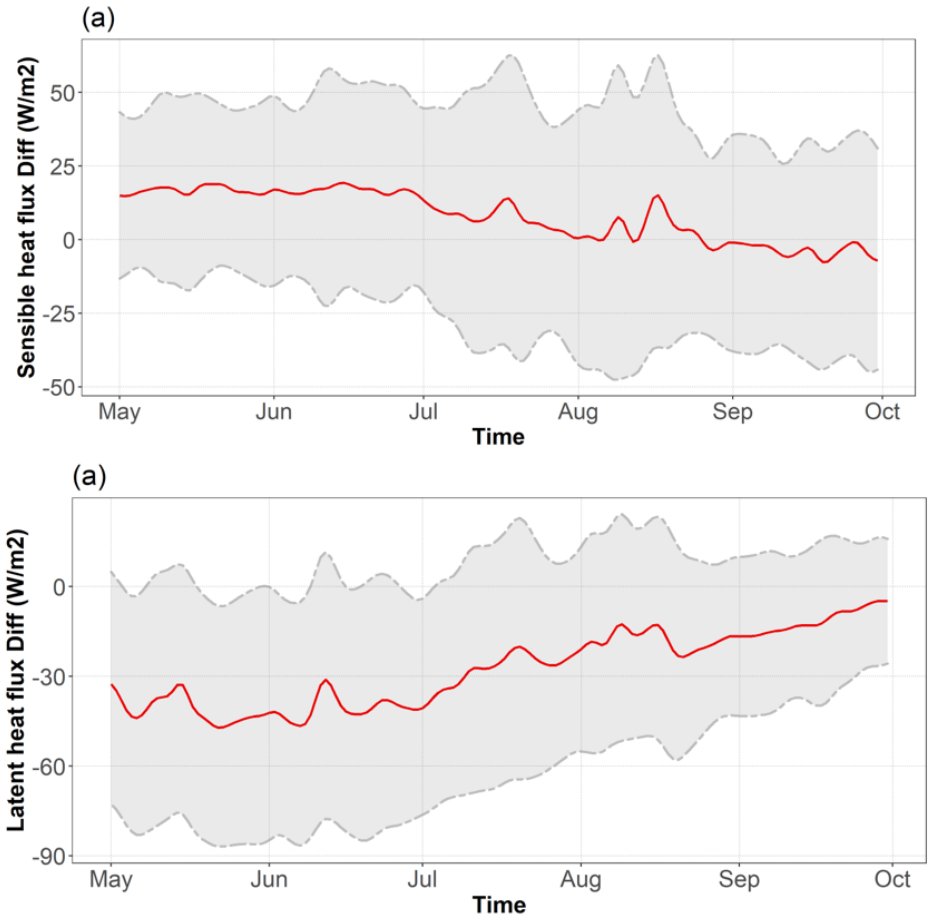

Figure 2.7 Time series of spatially averaged of daily maximum sensible heat flux (a) and latent heat flux (b) for updated and default simulations over the 4-km resolution simulation. The red line represents the difference of the simulations which were calculated as time series of the updated simulation minus the default simulation over the study period (1-May-30-September 2014). The gray areas represent ±1 standard deviation of the differences.

Air temperature and mixing ratio

The change of land cover from the default dataset to the updated data caused increased 2-m temperature (T2) in the southeastern United States by approximately 0.65 °C and 0.4 °C in the 12-km and 4-km simulations, respectively (Fig. 2.7-a, -b). A widespread increase (1 °C – 2 °C) in T2 was observed through the southern Midwest in 2007 as well as across central Georgia in the

12-km simulations across the three study years. These areas correspond to locations where dominant land use is changed from Croplands in the default data to woody Savannas in the updated LULC data (Fig. 2.3). Additionally, the results showed that T2 simulation at 12-km horizontal resolution was approximately 1°C higher than the associated simulation at 4-km resolution. This result shows how the weather simulations could be influenced by the horizontal spatial resolution of the LULC grids, which can be critical for downscaling weather predictions and climate model simulations such as the Coupled Model Intercomparison Project phase 6 (CMIP6) projections (M. Chen et al., 2020; Jianduo Li et al., 2021). Additionally, the results showed that the spatial extent of the areas with an increased air temperature was higher in 2007 relative to the 2014 and 2018 simulations. The greatest spatial extent of the area with decreased (0.5-1°C) air temperature was seen over south Illinois in 2014 at 12-km and 4-km resolutions, although updated WRF simulations indicated an increased temperature over this area in 2007. Additional decrease of T2 by roughly 0.5 °C was evident over southwest Texas in 2007 and 2018. The spatial extent and the intensity of the decreased T2 over south Illinois was higher in the 4-km simulation relative to the 12-km simulation (Fig. 2.7b). Focusing on temporal variation of air temperature simulations, a two-sample K-S test indicates that there are significant differences ($p < 0.05$) between the updated and default simulations. This result suggests that if more recent MODIS land cover data are used, the WRF simulations show slightly higher daily maximum temperature (+0.8 °C). The 0.5-1°C increase in the magnitude of 2-m air temperature over the southeast as a result of land cover change indicates the considerable impact of the land surface characteristics on regional climate conditions and associated model simulations.

Similar results were noted for 2-m water vapor mixing ratio (Q2) as for T2 variation patterns, with decreases of approximately 1-2 g/kg over the southern part of the nested domain

within the 12-km simulations (Fig. 2.7-a, -c). The difference of the two model runs showed an approximately 1-2 g/kg decrease in Q2 across the three simulation years, although in 2007 Q2 decreased by 1-1.5 g/kg across the Ozarks and into the northeast US. Results also showed that the Q2 increased by 0.5 g/kg over south Illinois and Indiana in 2014, indicating that transitioning from evergreen needleleaf forests, mixed forests, and open shrublands (default data), to woody savannas (updated data) had a noticeable impact on the amount of water vapor released to the atmosphere. Woody savannas (updated case) have a lower LAI (3.6) than the evergreen needleleaf forests and mixed forests (6.4, and 5.16 respectively); therefore, more surface radiation is partitioned into sensible heat and less into latent heat resulting in higher T2 and lower Q2 (Ran et al., 2015). As a result, decreasing LAI due to changes in LULC results in higher T2 and lower Q2.

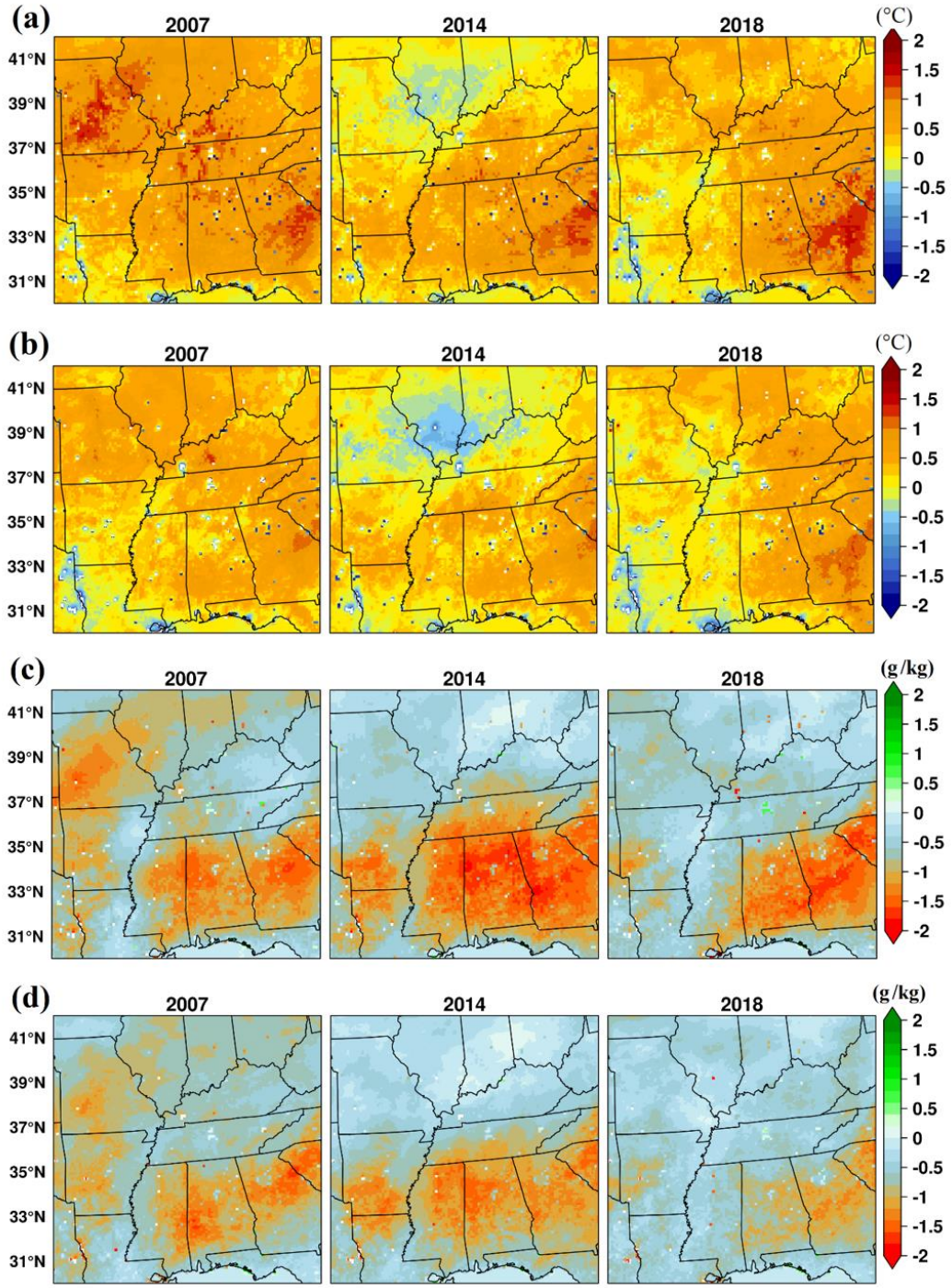

Figure 2.8 Average of 2-m air temperature (a-b) and water vapor mixing ratio (c-d) differences across the three simulation years analyzed over the nested domain at 12-km (a and c) and 4-km (b and d) spatial resolutions. Differences were calculated as time series of the updated simulations minus the default simulations over the period from May – August.

Soil moisture and precipitation

Analysis of the difference of WRF-simulated soil moisture with updated and default land cover indicated that modification of land surface physical characteristics can influence the root zone soil moisture. Figure 2.8-a shows the 30-cm volumetric soil moisture differences of updated and default simulations over the 12-km simulation across the three simulation years. As illustrated, the highest decreases (>10%) of 30-cm soil moisture occurred along the Ozarks and into the northeast (Illinois and Indiana) in 2007. The intensity of decreased 30-cm soil moisture in these areas was lower over the nested 4-km domain. Additional decreases (5-10%) in 30-cm soil moisture captured were seen over the center of the LMRAV, north Alabama, and eastern Tennessee in 2014, and north Georgia in 2018 in the parent 12-km domain; however, decreased 30-cm soil moisture content due to updated 4-km LULC was lower than associated areas in the 12-km simulation. Moreover, results showed the sensitivity of 30-cm soil moisture content to land cover change led to a roughly 10% increase in the over the northeast LMRAV, Kentucky in 2014, southeast Alabama, and Georgia in 2014 over the nested domain (Fig. 2.8-b).

These results indicated that soil moisture simulations are sensitive to LULC, because since within the Noah land surface model LSM used in the WRF simulations there are two primary variables - including vegetation type and the soil texture - upon which other secondary parameters such as minimal canopy resistance and soil hydraulic properties are determined (F. Chen & Dudhia, 2001b). As a result, changing the land cover type modifies the aerodynamic roughness length (m), thereby affecting the exchange coefficient (Ch) within the land surface model (Eq. 1).

$$C_h = \frac{K^2}{\left[ln(\frac{z}{z_{0m}}) - \psi_m \frac{z}{L}\right]\left[ln(\frac{z}{z_{0m}}) - \psi_h \frac{z}{L}\right]} \quad (2.1)$$

where ψ_m and ψ_h are the integrated similarity functions for momentum and heat, respectively, L is the Monin Obukhov length (m), k is the von Kármán constant (equal to 0.4), z_{0m} is the aerodynamic roughness length (m), and z is the measurement height (m). In this equation, the parameter z_{0m} is controlled by the LULC category. The areas of increased 100 cm soil moisture are associated primarily with mixed forests (z_{0m} = 50), deciduous broadleaf forests (z_{0m} = 50), and cropland (z_{0m} = 14) in the default LULC data, which were changed to woody savannas (z_{0m} = 5) in the 2018 LULCs. The C_h parameter directly affects the latent heat flux (Eq. 2), as a result, reduction in z_{0m} due to changes in land cover will lead to a lower C_h, and hence a decreased latent heat flux.

$$LE = \rho_s \, c_p \, C_h (q_G - q_A) \tag{2.2}$$

where ρ_s is the density of the surface air (kg m−3), C_h is the surface exchange coefficient for heat and moisture (m s−1), q_G is the water vapor mixing ratio (kg kg−1) at the surface, and q_A is the water vapor mixing ratio of the air at 2 m (more details are provided in Temimi et al., 2020). As a result, reducing the Ch factor and increasing stomatal resistance in woody savannas (300 s/m) relative to forests (125 s/m) can reduce the evapotranspiration, which works to maintain soil moisture content.

The accumulated precipitation between May-1 through September-30 in the default 12-km (4-km) LULC simulation exceeded the updated LULC simulation by approximately 100 mm (45 mm) over the whole nested domain across the three study years. As shown in Figures 2.9 c-d, the areal extent and magnitude of increased precipitation (green shaded colors) was less than the decreased precipitation (red shaded colors) across the years studied. In addition, WRF simulated precipitation with the 12-km updated LULC was greater in dry and wet year, with increased

precipitation observed north of roughly 35°N in the 2014 simulation. The spatial pattern of the precipitation variations due to LULC changes showed that there is a general area of reduced precipitation over Mississippi, Alabama, Tennessee, and north Georgia in 2007 and 2014. These areas are mostly associated with decreases in mixed forests and cropland/natural vegetation mosaics and increases in woody savannas and savannas in the MODIS updated land cover data. Moreover, in northern regions of the LMRAV there was an area of increased precipitation, especially in 2014 and 2018, where the latent heat fluxes showed highest regional increases. Although the precipitation generation is a complex process and it is difficult to precisely define the causes, Dyer (2011) suggests that soil and vegetation characteristics may lead to substantial sensible and latent heat flux gradients that effectively generate and drive mesoscale convective circulations.

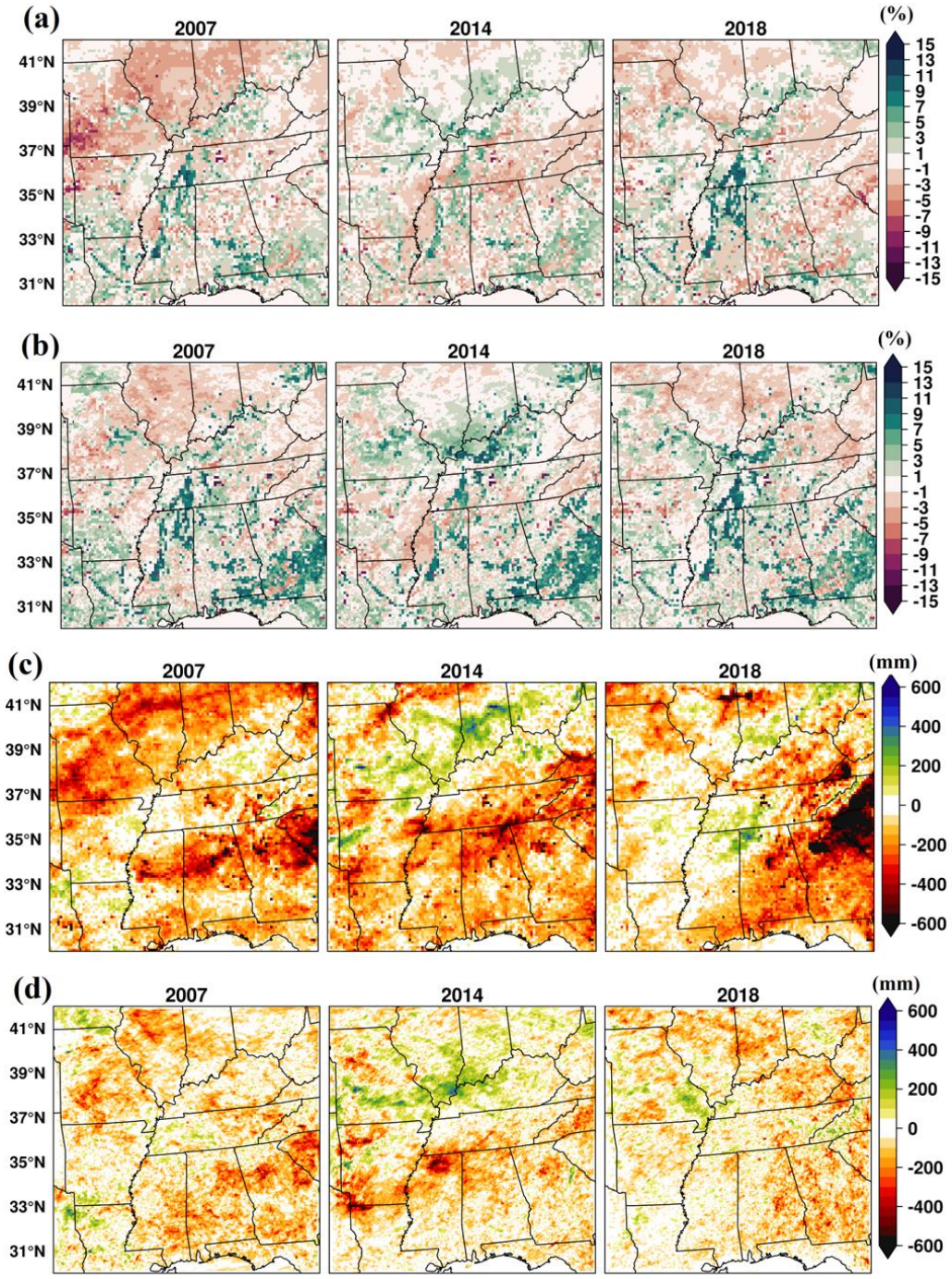

Figure 2.9 Average of 30-cm soil moisture (a-b), and accumulated precipitation (c-d) differences between the updated and default WRF land use cases (updated - default) across the three simulation years analyzed over 12-km nested domain (a and c) and 4-km parent domain (b and d). Differences were calculated as time series of the updated simulations minus the default simulations over the period from May – August.

To evaluate the temporal variability of precipitation over the study area, two subregions were selected within the study region where the difference between the accumulated precipitation with the updated and default cases was relatively high in each simulation year (Fig. 2.9). The spatially averaged precipitation was extracted for each simulation time over each sub region to create precipitation timeseries from May through September. Figure 2.9 a-f shows the cumulative probability density functions (CDF) plots of modeled precipitation for the updated and default simulations for the selected subregions. Figures 2.9-a, c, e show the CDF plots for the areas where the modeled precipitation with updated land cover is higher than the modeled precipitation with default land cover. Similarly, figures 2.9-b, d, f were generated for the areas where the WRF-simulated precipitation with the updated LULC was lower relative to the default simulation. Results indicated that above 50-mm accumulated precipitation the updated simulations showed higher precipitation relative to default cases, whereas under 50-mm the magnitude of precipitation from the updated simulation was equal or less than the precipitation from the default simulation. Similar patterns with opposite direction are shown for the CDFs of the areas where the updated simulation showed lower precipitation relative to default simulation. It is interesting to note that the detected change point around 50-mm accumulated precipitation was generally consistent across all simulations.

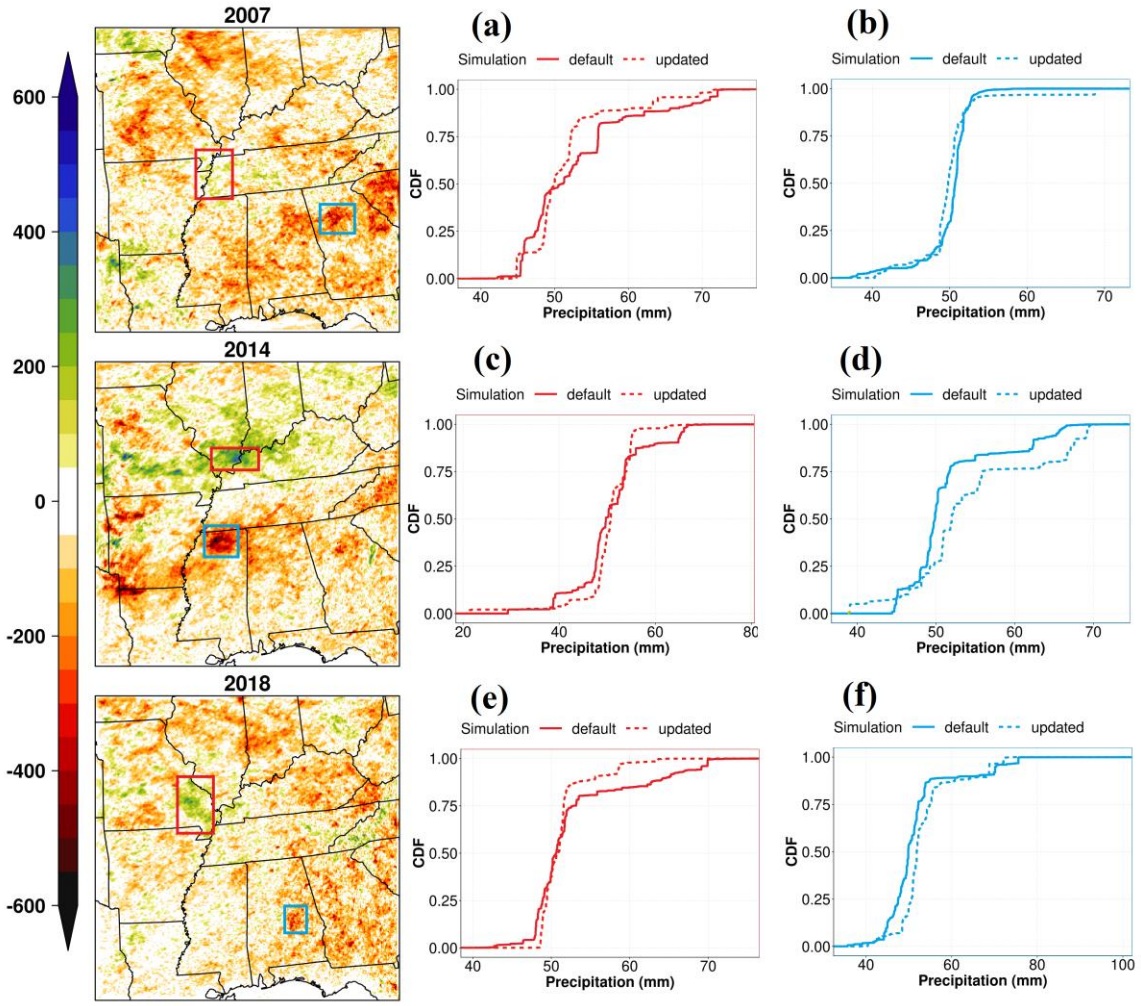

Figure 2.10 Left: same as figure 2.8d. boxes were used to extract precipitation timeseries for the areas of decreased (red) and increased (blue) precipitation due to LULC change. Right: cumulative distributed function (CDF) for each box over the study period (1-May-30-September).

To show the actual value of the model outputs for default and updated cases, histograms were plotted for each variable assessed in the study. As illustrated in the average maps, Fig. 2.10-a-b also represent an increased (decreased) accumulated sensible (latent) heat flux over the southeast. Results showed that the frequency of the sensible (latent) heat fluxes above (below) the average is shifted to the right (left) in the updated simulations. In addition, K-S test was applied

on temporally accumulated sensible heat flux to evaluate if there is difference between the updated and default simulations in spatial scale. Fig. 2.10-b shows that the histograms of accumulated latent heat fluxes for two simulations and associated average (vertical lines) which was significantly ($p < 0.05$) shifted to the left in the updated simulations. Similarly, the frequency of 2-m temperature above 27°C increased, whereas the frequency of 2-m temperature values less than ~22°C and greater than ~29°C did not substantially change between the default and updated simulations (Fig. 2.10-c-d). In addition, results showed that decreased water vapor mixing ratio by approximately 0.5 g/kg over the nested domain was obvious, and Fig. 2.10-d shows that the frequency of the extreme values (>13 g/kg) of mixing ratio is reduced. Finally, Fig. 2.10-e-d shows the average soil moisture and the accumulated precipitation for the default and updated WRF runs over the study area. Most importantly, LULC changes confirm that the extreme precipitation values were reduced in the updated simulation compared with the default simulation. These results indicate that the updated LULC may improve the bias values of the heavy rainfall simulations over the southeast; however, verification of the WRF simulations using observed precipitation is required to justify this statement.

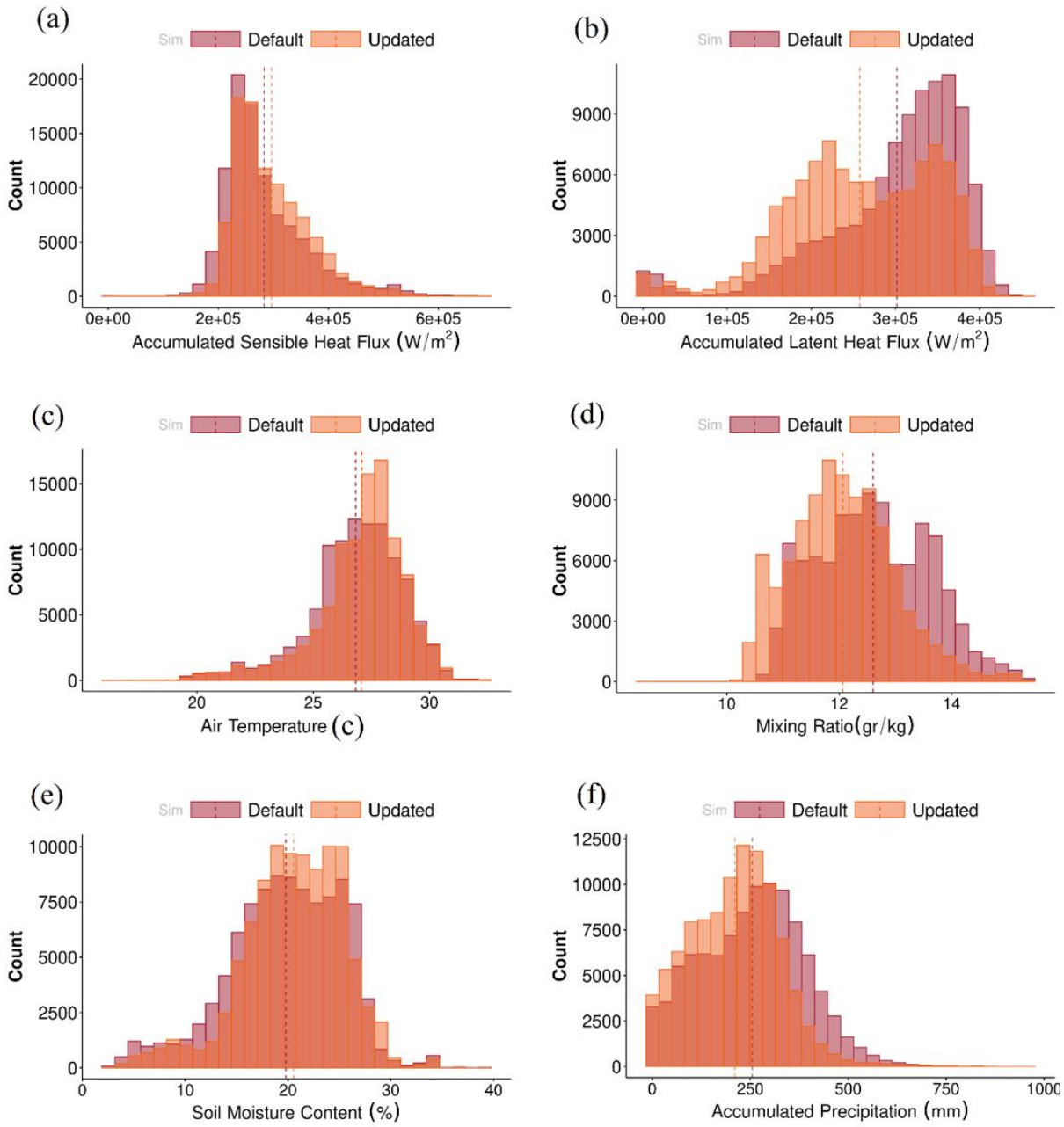

Figure 2.11 Histogram plots over the nested domain for the 4km updated and default simulations. a: accumulated sensible heat flux (W/m^2), b: accumulated latent heat flux (W/m^2), c: average 2-m air temperature (°C), d: average water vapor mixing ratio (gr/kg), e: average soil moisture content, f: accumulated precipitation (mm). Vertical lines indicate mean values for each case.

Conclusion

This study evaluated the impact of LULC changes on land surface-atmosphere interactions using WRF model simulations. Land cover is a critical factor in numerical weather simulations and land surface models; therefore, it is also important to assess the sensitivity of land surface simulations to changes in the dominant and fraction of land covers within the WRF modeling framework. This is especially important over vegetated areas and agricultural areas where land cover type is considerably affected by human activities and climate variability. Long-term averaged historical land cover data can lead to substantial uncertainty in capturing the interaction between the land surface and the atmosphere. As such, it is important to incorporate the variation of land use and land cover in weather simulations by utilizing near-real time remotely sensed land surface products such as MODIS land cover data. Results of this not only show how updated land cover information from the environmental satellites can impact mesoscale NWP model simulations, but also how important surface-based processes are on lower-atmospheric characteristics over areas with a mix of agricultural regions, natural landscapes, and human-modified environments, such as the southeast US.

The effects of spatial resolution of LULC on model-simulated surface heat fluxes, temperature, moisture, and precipitation have been examined by comparing model outputs at two different LULC horizontal resolutions. Model results showed strong sensitivity towards the model resolution and representation of surface characteristics, emphasizing the need to evaluate various horizontal resolutions of LULC over the southeast. These results indicate that modifying the dominant land use and land cover in the WRF simulations does not fully represent the impact of modifications to land use on land-atmosphere interactions independently. As a result, it is necessary to update land use fractions along with dominant land cover to understand the role of

gridded LULC information on NWP model simulations. The findings of the study are important in terms of evaluating the impacts of the land cover change on the surface heat fluxes, near surface temperature, and moisture modification at local and regional scales. This is also important in regional scale convective processes because as horizontal resolution goes higher (smaller grid spacing), more detailed LULC information is introduced to the model simulation.

Although the results of this study indicated that lower atmospheric patterns are influenced by changes in land surface properties, those influences are a result of complex interactions between surface heat fluxes, moisture variability and other land-atmosphere interactions. Thus, clearly understanding the influence of land use on specific atmospheric processes needs a comprehensive sensitivity analysis of different land surface models to define appropriate model parametrizations to improve weather predictions. By improving both model accuracy and the general scientific understanding of the role of land use on atmospheric properties, future predictions can be improved to benefit agricultural productivity and water resource management.

Future work should focus on annually updated land cover within the model to account for climate impacts such as wildfires and drought, deforestation, and urbanization at regional scales. Inclusion of other agricultural-related activities such as plant phenology and irrigation time, which represents more realistic land information and processes, would allow for improved and/or more representative land surface simulations. Finally, including high resolution land cover data and producing simulations at high spatial resolution would allow for analysis of deep soil moisture variations at local scales and consequent surface impacts such as landslides, drought evolution, and surface water resources. This may also be beneficial for accurately resolving local meteorological variability in complex terrains.

CHAPTER III

IMPROVING WRF SIMULATIONS THROUGH NEAR REAL-TIME REMOTELY SENSED LAND USE AND LAND COVER DATA OVER THE SOUTHEAST-UNITED STATES

Introduction

Physical characteristics of the land surface are a critical component of weather and climate models, contributing to a large variety of land surface processes. Dominant land use and land cover (LULC) and land use fractions play a major role in the exchange of heat and moisture between the land surface and atmosphere by influencing evapotranspiration (ET). Accurate representation of LULC within the Weather Research and Forecast (WRF) model is critically important for an accurate simulation of the near-surface fields since LULC essentially determines the surface's physical properties within the Noah land surface model used in the WRF simulations. Due to the link between LULC, soil water content and evapotranspiration, the stability and structure of the atmospheric boundary layer is directly influenced by characteristics of the earth's surface (Dyer, 2021), which can then lead to a modification of clouds and precipitation. However, evaluating the intensity and sign of this process over the southeast United States is important due to the predominance of local-scale variations in the surface heat fluxes that play a major role in convective patterns and precipitation generation (Dyer, 2011). The accuracy of the numerical weather predictions is generally limited by imperfect initial conditions, improper representation of land surface characteristics, and simplified model approximations (J. Zhang, 2018). Accurate

representation of the physical characteristics of the land surface is important to correctly partition the energy and moisture fluxes between the land and atmosphere.

Impacts of LULC changes within numerical weather prediction (NWP) models are simulated through the land surface model; therefore, differences in land surface model parameterizations can significantly impact the simulated land surface parameters in both magnitude and spatial extent (Boisier et al., 2012). Errors and uncertainties in land surface simulations occur in response to errors in LULC classification and in associated properties such as vegetation distributions and surface albedo (Bright et al., 2018; Hartley et al., 2017). Accurate representations of these land surface parameters are especially important in regions such as southeast U.S. where regional changes in LULC due to agricultural activities and transitions to or from active agricultural land. The impact of land cover type and soil water on precipitation is also important on shorter time scales for numerical weather prediction from nowcasting to the medium range. Significant influences of land cover type on surface energy balance components and convective indices and precipitation were found due to various complex processes (Jianduo Li et al., 2018; Pérez et al., 2017; Y. Yang et al., 2019).

For the purposes of this study, we investigate whether the use of MODIS-derived annual LULC data improve the accuracy of WRF simulations and whether the 2-m temperature and precipitation are well presented and could possibly be improved by using the updated surface characteristics. The model simulated air temperature and precipitation will be compared with in-situ observations to evaluate the performance of the WRF model simulations in spatial and temporal scales with default and updated land cover. Additionally, it is important to understand the importance of the updated LULC in the WRF model to potentially improve precipitation

simulations, which are important for many topics ranging from hydrological, agricultural applications, and climate change.

Data and Methods

The Advanced Research WRF (ARW) model version 4.3 (Skamarock et al., 2021) is used in this study for evaluation of LULC changes covering the period of May through September for 2007, 2014, and 2018 over the southeast United States. The WRF domain resolutions are 12-km and 4-km spatial resolution for the parent and nested domains with the XY number of grid points as 230 × 390 grids and 380 × 380 grids, respectively (Fig. 1). A 4-km spatial resolution was chosen to provide adequate detail of LULC variability in the default and updated model runs. The model vertical grid was configured with 42 vertical levels that are logarithmically stretched to provide enhanced resolution within the PBL. The 0.5-degree Global Forecasting System (GFS) analysis forcing data provided by the National Center for Environmental Prediction (NCEP) was used to initialize the boundary conditions, which were updated every six hours. These data are used for regional weather studies (López-Bravo et al., 2018; Ngan et al., 2015) over various locations, and are available at different temporal and horizontal resolutions (i.e. 0.25, 0.5, and 1 degree); therefore, they were considered as an appropriate forcing dataset for this study.

MODIS 500-m annual LULC data was obtained from the NASA Earth sciences data and information service center website (https://earthdata.nasa.gov) for the study area over the three study years. The LULC tiles were reprojected to the WGS84 geographic projection using majority resampling (mode) method, which is commonly used for discrete data such as LULC classifications. The reprojected data then converted to binary format so that they could be used by the WRF Preprocessing System (WPS). For the control simulations, WRF 4.3 geostatic surface characteristics from the 1-km MODIS historically averaged LULC data (Kumar et al., 2014) were

used. For reference, an analysis of the differences between the default and updated LULC WRF simulations over the nested domain are provided in Chapter 2.

The simulation period was set as 1 May through 30 Sep. 2007, 2014, and 2018, with the simulation starting on 15 April to allow for proper spin-up of physical parameters within the model. This period was chosen to evaluate the potential improvements of the WRF simulations due to updated LULC over a dry year (2007), a normal year (2014), and a wet year (2018) during the growing season. In a dry (wet) year, which received substantially less (more) precipitation than a normal year during the growing season, it is expected that the surface hydrologic conditions can affect agricultural activities, such as irrigation intensity and type of crops that are cultivated; therefore, LULC characteristics can be influenced by annual precipitation conditions.

The WRF simulations were evaluated against in-situ observations to quantify the accuracy of the model results. Daily averaged 2-m air temperature and daily precipitation observations from National Weather Service (NWS) global historical climatology network (GHCN; Fig 3.1-a, -b) were used for point-scale comparison against WRF-simulated 2-m temperature and precipitation for both the updated and default simulations. GHCN data is a daily quality-controlled database that addresses the critical need for historical daily temperature and precipitation, and snow records over global land areas. GHCN-Daily is composed of daily weather observation from numerous sources that have been merged and subjected to a common suite of quality assurance (QA) reviews (Menne et al., 2012).

Validation of the WRF simulations was performed to assess the ability of the updated LULC to enhance the model simulations. To quantify the performance of the model, four statistical evaluation metrics (Parajuli et al., 2020) including mean error bias (MEB), root mean squared error (RMSE), Nash-Sutcliff Efficiency (NSE) (Knoben et al., 2019), and coefficient of determination

(R^2) were computed to ascertain whether the updated LULC is able to improve the simulations relative to default simulations. The use of NSE as a statistical evaluation metric recommended (Xiaohui Zhong & Utpal Dutta, 2015) since one drawback in using RMSE is that it can be very large in magnitude depending on the units of the predicted values; however, all of these evaluation statistics were used to provide a comprehensive model evaluation. The mean error bias represents the average of differences between the WRF simulated variable and associated in-situ observed values, where positive values indicate that the observed values are higher than the associated simulation results and negative values indicate that the observed values are lower than simulation results. The NSE is a normalized statistic that determines the relative magnitude of the residual variance compared to the measured data variance. This evaluation metric ranges from -Inf to 1, indicating how well the plot of observed versus simulated data fits the 1:1 line. The optimal value for NSE is 1.0 and therefore for a "minimally acceptable" performance of the model, values should be larger than 0.0 (Naabil et al., 2017). NSE equal to 1 corresponds to a perfect match of WRF simulations to the observed weather data, and a value of 0 indicates that the WRF simulations are as accurate as the mean of the observed data. NSE values less than 0 indicate that the observed mean is a better predictor than the WRF-simulated outputs. The R^2 explains the goodness of fit between the modeled and associated observed variables. This evaluation metric is a common statistical technique used to measure the correlation between the observed and model-simulated values. Values of R^2 vary from 0 to 1 with higher values indicating a higher correlation (given as explained variance) between the modeled and observed variables.

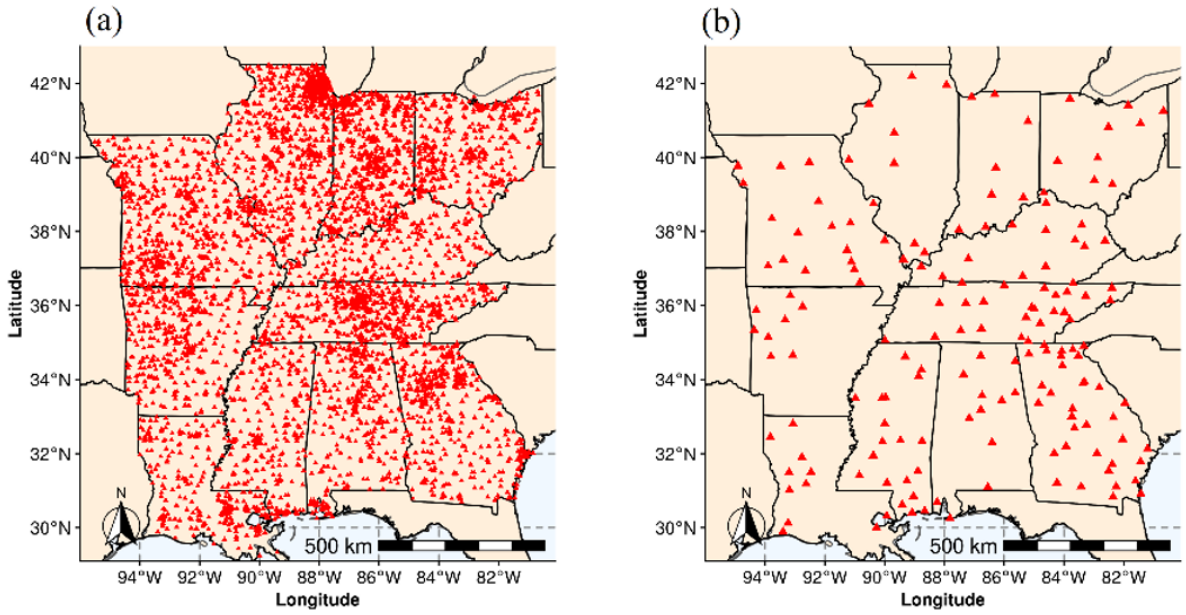

Figure 3.1 Spatial distribution of the 1300 rain gauges (a) and 151 2-m temperature observation sites (b) considered in this study for model evaluation.

Results and discussion

Updated LULC information did not improve WRF prediction performance in predicting air temperature over majority of stations.

Table 3.1 presents the RMSE, average bias, NSE, and R^2 computed between the hourly 2-m air temperature of measured with those of the default and updated simulations. Smaller RMSE values represent better agreement between measured and modeled values, whereas NSE values closer to 1 infer better model prediction. Table 3 shows that default simulation provided better result when compared with measured 2-m air temperature in 151 stations. WRF-simulated 2-m temperature values for the updated simulations showed a general warm bias compared with the associated observed data, with an average MEB of -2.24 K and a RMSE of 3.46 K. Similar results are shown for the 2-m temperature bias for the default WRF simulations, with similar MEB and

RMSE values compared with the updated simulation. This implies that the change in land use from default to updated has minimal effect on overall near-surface temperature patterns. The RMSE was relatively consistent across all stations and all simulation years, with a standard deviation less than one. For the 2-m temperature, results indicated that the RMSE values ranged from 1.9 K to 4.7 K, with the lowest (highest) RMSE occurring in May (July). Although updating LULC reduced mean error bias of WRF-simulated daily temperature by 0.1°C at 8%, 18%, and 9% of stations in 2007, 2014, and 2018, respectively. Overall, at 25 of the total 151 stations mean error bias of the updated simulations was less than the mean error bias of the default simulations. The cumulative distribution of NSE statistics were also plotted to summarize the model performance across all three simulation years for the updated and default WRF runs. Verification of WRF-simulated daily average 2-m temperature (Fig. 3.2-a) at selected stations over the study region showed that the updated LULC within the WRF simulation decreased the NSE values of temperature simulation relative to the default case. This indicates that by updating LULC, the overall performance of the temperature simulations over the southeast did not improve.

Results of the mean error bias evaluation of precipitation simulations showed that using updated LULC did not reduce the total bias of precipitation simulations relative to the default simulations over many study sites. The increased mean error bias and RMSE values in precipitation simulation due to updated LULC was relatively higher in 2007 and 2018, compared with the modeled precipitation in 2014. As shown in Fig. 3.2-b, the magnitude of the NSE values is negative in most of the stations. Further, although the updated LULC did not improve the precipitation simulation over the whole region, it decreased the frequency of the negative NSE values. Furthermore, the R^2 values of updated simulations are slightly less than the associated values of default simulations. As a result, updating LULC may not be the best way to improve the

performance of the WRF model simulations, and post processing the model outputs using statistical learning models may be a better alternative for enhancing the NWP simulations.

Table 3.1 Average values of four statistical evaluation metrics of updated (up) and default (def) WRF-simulated variables relative to 151 temperature stations across the three simulation years

Year	$RMSE_{def}$	$RMSE_{up}$	MBE_{def}	MBE_{up}	R^2_{def}	R^2_{up}	NSE_{def}	NSE_{up}
2007	2.94	3.28	-1.61	-2.04	0.66	0.64	-0.01	-0.23
2014	3.03	3.24	-1.88	-2.05	0.70	0.68	0.03	-0.12
2018	3.48	3.87	-2.33	-2.62	0.47	0.45	-1.35	-2.09
Mean	3.26	3.46	-1.94	-2.24	0.61	0.59	-0.44	-0.81

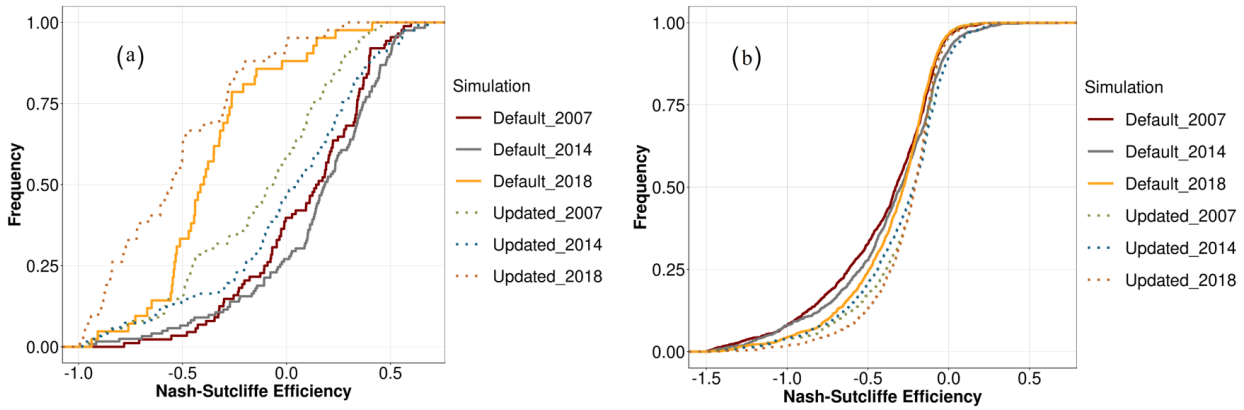

Figure 3.2 Cumulative distribution of WRF simulation performance calculated as Nash-Sutcliffe Efficiency (NSE) over 3-year simulations in 151 2-meter temperature stations (a) and 2400 rain gauge sites (b).

Performance comparison between the updated and the default WRF simulations for all stations are plotted in Figure 3.3-a, -b, -c. Difference of the RMSE values between updated and default (updated - default) simulations were calculated at each station. Negative values of change in RMSE represent improvement in temperature simulations due to updated LULC and positive

values show that decreased temperature accuracy in the updated WRF simulations. Based on the RMSE difference values at each site, the accuracy of temperature simulations decreased at almost all stations, however, over southwest Mississippi and central Texas in 2007, Missouri, Illinois, and Ohio in 2014, and Mississippi and Texas in 2018 simulations. Moreover, across many sites the magnitude of the R^2 metric decreased by approximately 2-6% due to using updated LULC. Even though the R^2 was not improved in many sites, mean error bias was reduced due to updating the land cover within the model simulations. For example, difference of the mean error bias between the updated and default simulation showed that the bias associated with the updated LULC is 0.5-1 °C less than the default temperature simulation over south Georgia in 2007; however, the R^2 values did not improve. The spatial pattern of differences in simulated temperature between the updated and default simulations over south Georgia showed that decreased RMSE and R^2 values are mostly associated with transition from croplands in the default to woody savannas in the updated LULC data (Fig. 2.3). The spatial extent of changes in LULC form the default dataset to the updated data was almost consistent across all simulations, but the change in model performance was not consistent across all simulations; therefore, it is difficult to find spatial relationships between LULC and changes in model performance.

Precipitation observations and WRF simulated precipitation were compared at each measurement point over the nested domain across the three study years (Fig. 3.3-d, -e, -f). The negative RMSE values ($RMSE_{update}$ - $RMSE_{default}$) suggested that WRF simulation of precipitation can be improved considerably with the updated information of LULC at many stations. Nevertheless, R^2 values between the observed and simulated precipitation showed that the model performance was reduced by approximately 2-6% over many other stations. Overall, the number of stations with improved simulated precipitation was higher in 2014. One reason could be because

the model struggle in simulating extreme (in dry and wet years) rainfall events, however, different microphysics and parameterizations should be evaluated with the updated LULC data. As shown in Fig. 3.2-b, the change of the NSE values between the updated and default simulations was relatively lower than the 2007 and 2018 simulations. Given the noted sensitivity in the magnitude and spatial patterns of the sensible and latent heat fluxes as a result of the updated land use (Chapter 2), results indicated that there was not an apparent spatial pattern in improved precipitation simulation over the study area.

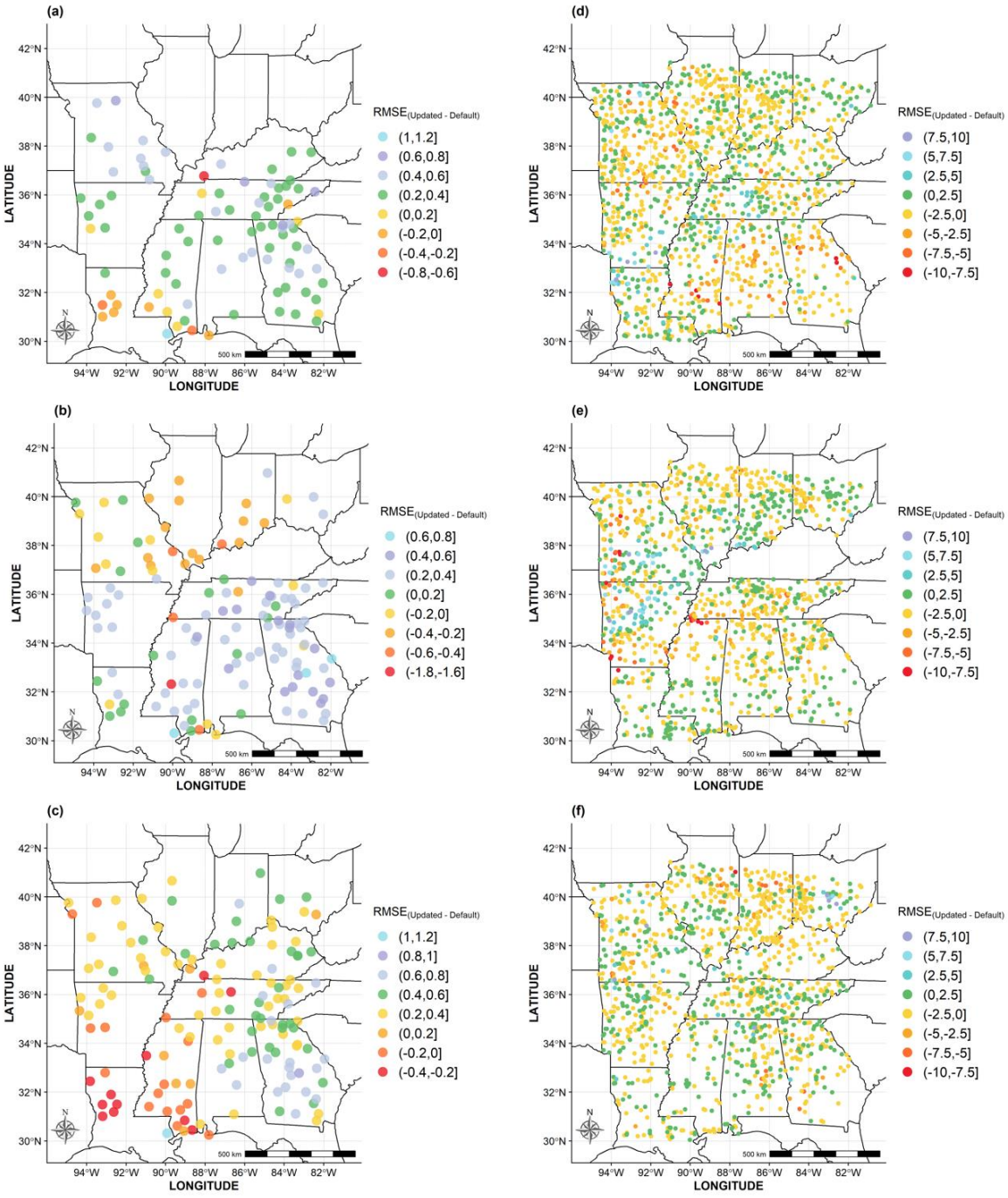

Figure 3.3 Per-station performance change between the updated and default simulations in 151 temperature stations and 2400 rain gauge sites across the southeast US. Negative values indicate stations where updated LULC improved RMSE values relative to default simulations. The first (second) column shows temperature (precipitation) stations. a, b, and c (d, e, f) are associated with 2007, 2014, and 2018 simulations.

We further investigated the performance of daily precipitation simulations by comparing the RMSE values for the updated and default model runs. Three different threshold levels (<0-, <1-, and <2-mm) were considered to compare the differences between two simulations and determine the number of stations in which the model accuracy was increased. Table 3.2 summarize the RMSE values of the default, and updated simulations in each simulation year. The percentage of number of rain gauge stations (out of 1300 sites) where the difference of the RMSE of the updated and default simulations are less than 0, 1, and 2-mm is shown. The geographical location of stations where the updated LULC increased precipitation simulations by 2-RMSE or more, is shown in Fig. 3.4. Although the number of stations in which the simulated precipitation improved by 2-mm or more according to the RMSE metric were about 15% percent of the total stations (Table 3.1), a geographical pattern was observed for each simulation year. For example, an increased precipitation simulation performance was observed over the south Great Plains and along the Arkansas into northeast in 2014 simulation, where 16% of the stations were located. Similarly, in a large area over the south Mississippi, Alabama, and north Georgia relative improvement were made in WRF precipitation simulation due to updating LULC in 2007. Additionally, almost in 10% of the stations which were located over Indiana, Illinois, and Kentucky, the model performance improved in the updated LULC simulation in 2018.

Table 3.2 Percent of rain gauge stations showed an improvement due to updated LULC

Year	$RMSE_{default}$	$RMSE_{updated}$	$RMSE_{updated}$ - $RMSE_{default}$		
			< 0	< -1	< -2
2007	10.4	9.8	65%	32%	15%
2014	12.1	11.7	60%	30%	16%
2018	12.0	11.6	63%	28%	12%

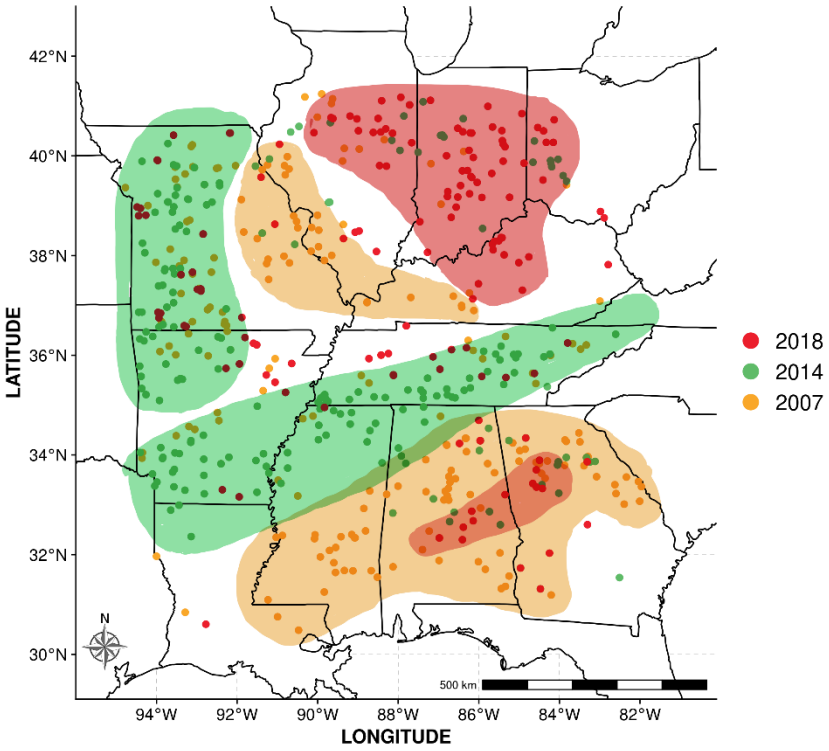

Figure 3.4 Geographical location of stations where the updated LULC increased performance of the precipitation simulations. Shading represents the spatial pattern (roughly) of enhanced precipitation simulation for each year.

Conclusion

The aim of this this work was to quantify the change in model accuracy by updating the LULC fields within the WRF modeling framework, and whether these updates might be beneficial in terms of improving regional weather predictions. In this study, the accuracy of a 4-km WRF simulation using default and updated LULC information was quantified through evaluation of 2-m temperature and precipitation predicted by the WRF model for representative wet, dry, and normal years. Simulations were conducted during the warm season (May-1 – September-30) over the Southeast US, where LULC showed substantial changes between the default and updated data.

Overall, the NSE evaluation metrics between daily observed and simulated data indicated that the 2-m temperature and precipitation forecast error from the WRF model increased due to updated LULC data across the area. Results indicated that the bias of the precipitation simulations was high in the dry year (2007) relative to the normal and wet years; however, there was not a consistent spatial pattern in model performance change for precipitation simulations. Although the performance of the temperature simulations weakened due to updated LULC across many sites, the performance of the model showed an improvement over the Southern Appalachians in 2018. Results indicated the potential importance of land cover characteristics on regional scale simulations, particularly for the precipitation estimations; therefore, future investigations should involve an ensemble simulation whereby a multi-physics ensemble is generated with updated land use to quantify potential improvements more clearly in the model simulations especially for extreme rainfall events. LULC change were proven to have substantial impacts on the inhomogeneity of climatic variables (Zhao & Wu, 2017), which was the primary objective of the research. However, other parameters such as cumulus parameterization schemes might also have influences on the simulated values, such as the cloud fractions at different levels, for which results from ensemble simulations using different parameterization schemes will be more objective. In addition, evaluating the performance of the Noah land surface model can have substantial contribution in understanding and realizing the scientific reasons of the decreased WRF performance due to providing more realistic land surface characteristics.

CHAPTER IV

IMPROVING NOAH-MP LAND SURFACE MODEL SIMULATIONS USING MACHINE LEARNING MODELS

Introduction

The land surface plays an important role in the Earth's climate system due to the exchange of energy, mass, and momentum between the land surface and atmosphere (Chang et al. 2020). These processes directly determine the repartitioning of net radiation into sensible and latent heat fluxes, thereby influencing other variables such as precipitation, evapotranspiration, soil moisture, runoff, and groundwater storage. Accurate estimation of the surface heat and moisture fluxes is a particularly challenging task for numerical weather models. On one hand, the accuracy of the geostatic information such as land cover, vegetation density, and soil texture must be sufficiently fine to adequately describe physical characteristics of the land surface (Tomasi et al. 2017; Dennis and Berbery 2020); however, due to difficulties in capturing near-real-time vegetation status and land surface conditions, multiannual means of these static fields have been used in numerical weather models (Yin et al. 2016). Research has shown that introducing dynamic crop models and prescribing more accurate vegetation information within the Noah multi-parameterization (Noah-MP) land surface model (LSM) improved the simulations of the land-atmosphere interactions (Tsvetsinskaya et al. 2001; Levis et al. 2012; Liu et al. 2016). On the other hand, the accuracy of the forcing data into such models, such as from the North American Land Data Assimilation System (NLDAS), strongly influences the regional distribution of soil moisture, atmospheric water

vapor, and surface heat fluxes (Levis et al. 2012; Pilotto et al. 2015). As a result, uncertainty in accurately capturing the land-atmosphere interactions in LSM simulations is associated with the quality of the forcing data, prescribed land surface properties, and boundary conditions (Huang et al. 2013).

Building on the Noah land surface model (Ek et al., 2003), Noah-MP has improved substantially during recent years (Godfrey & Stensrud, 2010; Mahrt & Pan, 1984; G. Niu et al., 2011; P. Wang et al., 2018). For example, some improvements include: (1) evolving snow sub-models from simple bulk-layer models to multilayer models to incorporate more physical processes (R. Yang & Friedl, 2003), (2) considering the effects of sub-grid topography on soil water distribution and runoff generation (G. Y. Niu et al., 2005), (3) including the exchange of water between an unconfined aquifer and the overlying soil column in the soil hydrology schemes, and (4) considering plant photosynthesis, respiration, and the related nitrogen cycle (Dickinson et al., 1998, 2002). Moreover, the improved physics within the Noah-NP model includes a dynamic groundwater component, an interactive vegetation canopy, and a multilayer snowpack (X. Cai, Yang, David, et al., 2014). Additionally, the multi-parameterization options allow users to choose multiple parameters in leaf dynamics, soil moisture factor for stomatal resistance, canopy stomatal resistance, runoff, and groundwater. A comprehensive description of the model improvements are given in (X. Cai, Yang, Xia, et al., 2014; Ek et al., 2003). Among the most important boundary conditions affecting partitioning the surface latent and sensible heat fluxes is soil moisture, which is one of the largest sources of uncertainty in LSM simulations (Koukoula et al. 2019). Soil moisture, which may affect precipitation and runoff simulations, can provide a “memory” of past precipitation events (Wang et al. 2018b). As a result, the lower the uncertainty of precipitation forcing data, the more accurately the LSMs can estimate soil moisture. Tuttle and Salvucci (2016)

studied the influence of soil moisture anomalies on rainfall in the United States and found a positive relationship between soil moisture and precipitation over the western United States, concluding that the relationship between these variables depends on the regional aridity; however, to date, simulated fields of soil moisture remain unreliable and the ability to observe soil moisture over large scales is limited. Fortunately, the rising interest in the LSM user's community in improving land - atmosphere interaction simulations (García-garcía et al. 2020) has coincided with an unprecedented improvement of machine learning models and availability of surface heat flux and soil moisture flux field measurements, particularly due to the efforts of the FluxNet community (Workshop 2015).

Over recent decades, the FluxNet micrometeorological tower sites have been used to measure the land surface exchange of carbon, water, and energy using eddy covariance techniques (Williams et al. 2009). The AmeriFlux network (https://ameriflux.lbl.gov) of carbon and water flux tower sites was formed more than 20 years ago by a group of scientists who were separately monitoring these fluxes at individual sites (Novick et al. 2018). With the development and introduction of a dense network of FluxNet stations with a high temporal measurement resolution, detailed observations of the surface heat and moisture fluxes are available to give more information regarding the variations in surface energy exchanges. These data have been available since roughly 1994 as 30-min quality-controlled values (Chaney et al. 2016; Pastorello et al. 2020) for studying land-atmosphere interactions and exchange of CO_2 and water and energy (K. Wang & Liang, 2009) in the biosphere-atmosphere system the to minimize the uncertainties of LSM simulations. Information about the surface heat fluxes, soil moisture fluxes, and evapotranspiration is essential for numerical weather prediction since these variables are important parameters in lower boundary layer conditions, convective processes, and precipitation generation. This network has been

considered the most reliable data source to verify LSM outputs including latent heat, sensible heat, net radiation, ground heat flux, surface friction velocity, and outgoing longwave radiation, among others (see Section 2.3).

Sensible and latent heat fluxes and radiative processes transfer heat between the air and the land surface, which are essential terms describing the physical feedbacks in coupled land-atmosphere models. Together with evaporation, precipitation, and runoff, these energy and mass exchanges modify air and water at the surface (Li et al., 2017). Diurnal variations of surface heat fluxes and energy partitioning by LSMs show systematic deviations from observations that were reported in early (Yang et al. 2015) and more recent model studies (Best et al. 2005). Consequently, the imperfect simulation of the land-atmosphere feedbacks in LSMs has been suggested as a main reason why sub-seasonal and seasonal operational weather models fail to predict extreme weather events such as extreme runoff events (Dirmeyer and Halder 2017). This is important because land-atmosphere interactions propagate to larger scales and may finally affect the model sensitivity to global changes (Renner et al. 2020).

Land surface models are developed to provide lower boundary conditions for climate and weather prediction models, which simulate partitioning of energy and moisture at the Earth's surface-atmosphere interface (Pitman 2003). Despite advancements in LSM complexity – having developed from simple bucket models to multilayer hydrology and canopy models (Liu et al. 2016; Niu et al. 2011) – they continue to struggle to accurately represent land-atmosphere processes (Dirmeyer et al. 2006). Much research has been done to optimize LSM parameters (Chaney et al. 2016; Godfrey and Stensrud 2010; Nasonova et al. 2015). Researchers usually choose the physical land surface processes options in the LSMs based on their expert knowledge, experience, and the literature, or by adopting the default values given in the manual (Chang et al. 2020). It is difficult

to use an appropriate parameterization scheme for a heterogeneous landscape due to the incompatibility of the chosen scheme for some of the various surface environments. In addition, one of the main obstacles in the parameter optimization of LSMs is that the nonlinear interactions among model parameters that are unresolvable by current LSM understanding (Sawada 2020). To improve the model estimations, an alternative is a combination of dynamical and statistical learning models for practical forecasting. For example, model output statistics (MOS) is an approach that utilizes statistical relationships between the predictand and variables from the numerical models at some projection times (Glahn and Lowry 1972) to improve model accuracy. Despite efforts to improve LSM simulations (Zhang et al. 2020; Yin et al. 2016; Chaney et al. 2016), no major study has attempted to evaluate how well different machine learning models are performing for post processing surface heat flux and moisture flux simulations of Noah-MP model.

The main objective of this study is to utilize machine learning models to improve Noah-MP surface heat flux, soil moisture and soil temperature simulations relative to available surface observations. The manuscript is organized in the following manner: Data and methods section (Section 2) will describe the dataset that has been used to force the Noah-MP model and train different machine learning models, as well as the observed surface data. In the results and discussion (Section 3), the Noah-MP simulations and associated machine learning estimations will be compared using statistical evaluation metrics for sensible and latent heat as well as soil moisture and temperature, followed by a discussion of the improvement and validity of our proposed method. The conclusions, which summarize the findings and outline future research, are provided in Section 4.

Data and Methods

Noah-MP model

Noah-MP has been developed to run in association with gridded models such as the Weather Research and Forecasting (WRF) model, or over individual site locations. Since the objective of this study is evaluating the accuracy of the simulated surface heat and moisture flux simulations in comparison with observed data, the Noah-MP model is run in a single site configuration over the AmeriFlux sites. The model was forced with 1/8° by 1/8° (roughly 11 km) hourly weather data derived from NLDAS-2 (Luo et al., 2003; Xia et al., 2012), including precipitation, air temperature, specific humidity, surface air pressure, wind speed, and incoming shortwave and longwave radiation. The superior spatiotemporal resolution of the NLDAS-2 data makes it more applicable for evaluating surface heat flux and moisture fluxes at AmeriFlux sites. In addition, Noah-MP static input data such as land surface characteristics (Strack et al., 2003; Talib et al., 2020) and soil texture data extracted from MODIS annual land cover and WRF static soil texture data, respectively. Noah-MP simulations were performed for each AmeriFlux site individually (single point) for an 18-year period (2001 – 2018) with one hour temporal resolution. Given the atmospheric forcing and static inputs (e.g., soil and land cover type), land surface models need to find their own equilibrium state before the output can be compared to actual observations; therefore, the Noah-MP simulations were initialized by running the model from 2001 through 2005 to reach an equilibrium state. More details about the spin-up time for LSMs can be found in Yang et al. (1995).

FluxNet observations and site selection

The AmeriFlux dataset (Novick et al., 2018) provides critical linkages between terrestrial ecosystem processes and climate-relevant responses at regional and continental scales, and there are more than 260 flux tower sites in the U.S. that support continuous observations of water and energy fluxes and ecosystem carbon exchanges between the Earth's surface and atmosphere. Alongside many other meteorological variables, the AmeriFlux eddy covariance sites measure latent heat flux (W/m^2), sensible heat flux (W/m^2), soil moisture content (gr/m^3), and soil temperature at 30-minute intervals. Two main factors limited the number of sites to be used to evaluate LSM simulations: (1) the available period of record of observations and (2) the height of the sensors collecting the data at each station. To ensure that the seasonal cycles of surface heat fluxes are captured, the years of near-continuous observed data were kept. Therefore, the sites used for this study were selected based on the availability of the surface heat flux measurements and the depth of soil moisture and soil temperature observations at 0.1, 0.3, 0.6, and 1 meter to align with the Noah-MP simulation model levels. For example, for developing machine learning models for 0.1-m soil moisture, all sites that had at least 3 years consecutive years observed soil moisture data at 0.1m were selected, because a separate model for each site were developed. The selected sites were representative of 11 land cover types (Fig. 4.1) according to the International Geosphere-Biosphere Program (IGBP) classification and different climate conditions. Due to complexity of temporal autocorrelation of time series data, the temporal structure of the data should be considered in machine learning model development. In other words, consistency of the time series data provides additional structure like trends and seasonality that can be leveraged to improve the skill of the machine learning models. Figure 4.1 shows the geolocation of AmeriFlux sites which were used for evaluating the Noah-MP and machine learning estimations in this study. Further, Table

4.1 and 4.2 show the AmeriFlux sites measuring soil moisture content and soil temperature and their respective dates at which the data are available.

Table 4.1 AmeriFlux sites measuring soil moisture content and their respective dates at which the data are available.

	Observation period			Observation period	
Site ID	Start	End	Site ID	Start	End
US-A32	2015	2017	US-Me2	2007	2019
US-A74	2016	2017	US-Me6	2011	2019
US-AR1	2009	2012	US-MMS	2001	2015
US-AR2	2009	2012	US-Moz	2004	2017
US-Aud	2003	2008	US-MRf	2006	2012
US-Bar	2005	2017	US-NC2	2005	2019
US-Bkg	2004	2010	US-NR1	2002	2005
US-Blk	2004	2008	US-Rwf	2014	2018
US-Dk1	2001	2007	US-SdH	2004	2010
US-Dk2	2001	2007	US-Snd	2007	2015
US-FmF	2005	2010	US-SRG	2008	2019
US-FR3	2004	2012	US-SRM	2004	2019
US-Fuf	2005	2010	US-Tw3	2013	2018
US-Fwf	2005	2010	US-UMB	2009	2019
US-GLE	2001	2014	US-Whs	2007	2019
US-IB1	2005	2018	US-Wkg	2004	2019
US-IB2	2004	2018	US-Wrc	2001	2018

Table 4.2 AmeriFlux sites measuring soil temperature content and their respective dates at which the data are available.

	Observation period			Observation period	
Site ID	Start	End	Site ID	Start	End
US-Cop	2001	2007	US-MOz	2004	2017
US-CPk	2009	2013	US-NC2	2005	2019
US-Dk2	2001	2005	US-Ne1	2001	2019
US-Elm	2008	2014	US-Ne2	2001	2019
US-Esm	2008	2015	US-Ne3	2001	2019
US-Fmf	2005	2010	US-Ro1	2004	2012
US-FR3	2004	2012	US-Ro4	2015	2019
US-Fuf	2005	2010	US-Rwf	2014	2018
US-Fwf	2005	2010	US-SdH	2004	2009
US-GLE	2001	2014	US-Skr	2004	2011
US-Ha2	2004	2019	US-SRM	2004	2019
US-IB1	2005	2018	US-Syv	2001	2008
US-IB2	2004	2018	US-WCr	2001	2019

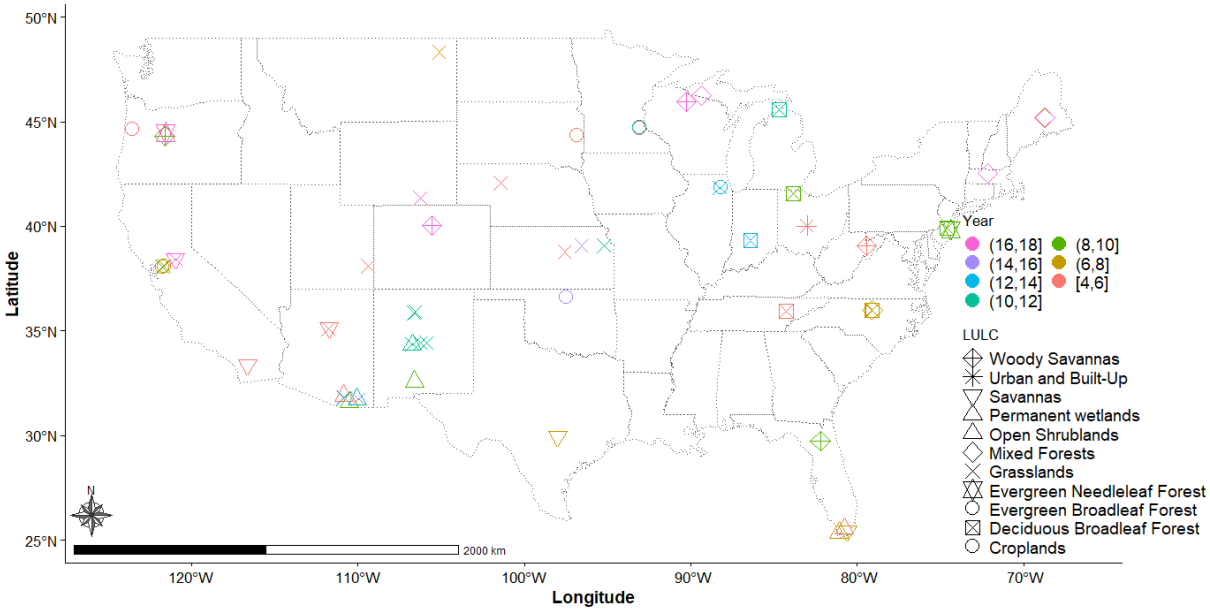

Figure 4.1 Map of AmeriFlux sites over the United States used in this study to evaluate surface heat flux simulations. Stations are colored according to the length of the period of record.

Preprocessing data

Prediction of surface heat and moisture flux (predictands) was essentially a supervised regression problem in which the predictand (i.e., latent heat flux) was a continuous variable. For the first step in the study, the data were sectioned into k-folds (time-splitting) to perform K-fold-Cross-Validation. The train-test split is a technique for evaluating the performance a machine learning algorithm, which can be used for classification or regression problem and can be used for any supervised learning algorithm. The models were trained on a subset of folds as training set (e.g., 3 years out of 5 years), validated on a small subset of folds as validation fold (e.g., year 3 out of 5 years), and test on remaining folds as test set (e.g., last year). It should be noted that the validation and test data sets were completely withheld from the entire data process and only the training segment was used to fit the model, and then evaluated on the validation set. The validation

data points are then included as part of the next training dataset, so the length of training data set growth over time, and subsequent folds were divided into a new validation and test sets (Fig. 4.2). Precisely, the workflow of how the data is being used to train the models, tune the hyperparameters and select the best model is as below:

1- Training dataset is used to train models for each hyperparameter.

2- Validation dataset is used to evaluate the models.

3- One of the models is chosen based on adjusted R^2.

4- The trained model is evaluated with the test dataset.

These steps were repeated to test the model on each training fold (Fig. 4.2), and the skill of the model predictions was assessed based on the average of evaluation statistics obtained in every holdout testing segment. The number of folds (K) was different for each site depending on the number of available data at each site. It should be noted that splitting timeseries data is different from non-time series data, since the characteristics of time series data, such as autoregressive nature, trend and seasonality would not allow to randomly shuffle the data and split into training and testing groups. Instead, the data must be split with respect to the temporal order (blocks of time) in which values were observed. Additionally, since seasonal patterns are inherent in the data, the data were grouped into annual sets to avoid introducing bias for a particular season, so the size of blocks of time for each segment (train, validation, and test) was at least one year.

Training data sets were standardized using the *z*-score technique such that each new feature had a zero-mean and unit-variance distribution. The mean and standard deviation of the training set were then used to standardize the validation and test data frames. Noah-MP-simulated Shortwave downward radiation, total absorbed solar radiation at ground, sensible heat flux, latent heat flux, ground heat flux, canopy sensible heat flux, water vapor pressure, evaporation, soil

moisture, and soil temperature were used as input features, and AmeriFlux observations (e.g. latent heat flux) were used as dependent variable (predictand) for the machine learning models.

Feature selection is primarily focused on removing non-informative or redundant predictors from the machine learning model and maximizing the adjusted R^2 for the training phase. The selection methodology utilized Recursive Feature Elimination (RFE) method as a baseline/control methodology, which has been shown to be a useful approach (B. Zhang et al., 2022) to isolating important predictors before optimizing machine learning configurations. This method is a sequential backward selection algorithm which uses a specific underlying algorithm to select features by evaluating the association between each variable with the target variable and removing non-informative predictors from the model and reducing the size of the feature set. This is a useful approach to isolating important predictors before optimizing machine learning configurations (Mercer, 2020). Having irrelevant features can decrease the accuracy of the model; therefore, feature selection can reduce overfitting, reduce computation time, and improve the accuracy of the model. It should be notice that the chosen predictors were different for each ML model configuration at a given site. Table 4.3 summarize the frequency (%) of the kept predictors for sensible heat flux (H), latent heat flux (LE), soil moisture at 0.1m, and 0.3m (SWC_10, SWC_30, respectively) and soil temperature at 0.1m (TS_10) across all sites.

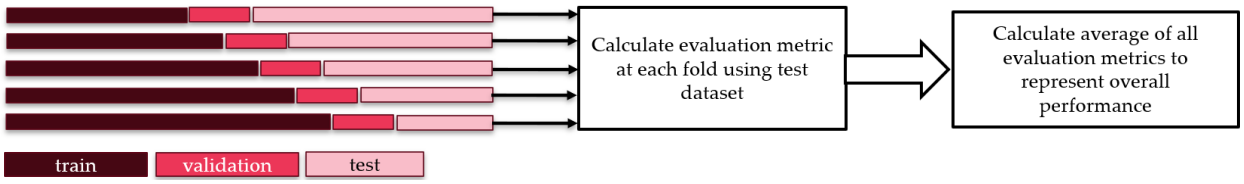

Figure 4.2 Flowchart illustrating cross-validation procedure used for training machine learning models to post-process time series of Noah-MP model outputs in Ameriflux sites.

Table 4.3 Percent of kept predictors for all 5 variables (predictands) in the stations based on the best selected model at each site.

Noah-MP-simulated variables (Predictors)		AmeriFlux observed variables (Predictands)				
Variable	long name	H	LE	SWC_10	SWC_30	TS_10
HFX	sensible heat flux (W/m^2)	100	100	0	0	4
LH	latent heat flux (W/m^2)	100	100	24	36	7
LWFORC	shortwave radiation (W/m^2)	18	0	0	0	0
SWFORC	longwave radiation (W/m^2)	0	73	28	14	0
SOIL_M10	0.1m soil moisture (%)	36	36	100	57	14
SOIL_M30	0.3m soil moisture (%)	9	36	92	100	4
SOIL_T10	0.1m soil temperature (K)	0	27	32	36	100
SOIL_T30	0.3m soil temperature (K)	9	0	44	50	100
EVG	ground evaporation heat (W/m^2)	0	0	32	36	11
EAH	canopy air vapor pressure (Pa)	0	0	20	7	4
GRDFLX	ground heat flux (W/m^2)	82	24	24	21	0
FSA	total absorbed solar radiation	64	64	0	0	0
SHC	Canopy sensible heat (W/m^2)	45	27	0	0	0
CHLEAF	leaf exchange coefficient	18	55	0	0	0
CANLIQ	intercepted liquid water (mm)	0	0	16	21	4

Machine learning models

In this study, five machine learning algorithms were tested, including: (1) penalized regression (Ridge regression), (2) support vector regression (SVR), (3) random forest, (4) gradient boosting and (5) neural network. The standard linear model (or the ordinary least squares method) performs poorly in a situation where there is high complexity in a large multivariate data set. An alternative is the regularized regression (or penalized regression) (DelSole & Banerjee, 2017) allowing for the generation of a penalized linear regression model by adding a constraint in the equation, referred to as regularization. This penalty forces the model to reduce (i.e., shrink) the

coefficient values towards zero, which allows the less contributive variables to have a coefficient close to zero or equal zero.

The Support Vector Regression (SVR) gives the flexibility of defining how much error is acceptable in the model and will find an appropriate line (or hyperplane in higher dimensions) to fit the data. The SVR model has parameters that requires a robust optimization methodology to obtain the best possible configuration. Both radial basis function and polynomial kernels were tuned where polynomial kernel degree from 1 to 4, and the γ values on the radial basis function kernels included 0.001, 0.01, 0.05, 0.1, 0.2, and 0.3. Further, the cost function was tested for 1, 10, 100, and 500 and tested values for the ε-loss function included 0.1, 0.05, and 0.01 (which were tested for every kernel configuration).

A random forest (RF) is a machine learning technique that is widely used to solve real-world problems (Wu & Li, 2019; Xu et al., 2018). The backbone of the random forest method is decision trees, with each of the data partitions in classification or regression trees based on if-else rules (Liaw & Wiener, 2002). The aggregate of the results of multiple predictors gives a better prediction than the best individual predictor. There is a strategy to predict the response variable by getting them from the predictions of all individual trees. The Random Forest model works based on the strategy of the mean ensemble technique (Liaw & Wiener, 2002). The Gini index or the Entropy index are used for splitting or partitioning the data into regression or classification trees. In practice, the Gini index measures the purity of each node, such that the value of the Gini index for a child node should be less than the parent node in order to predict the class. Conversely, the Entropy index is a measure of node impurity. For the RF model number of predictors (2, 3, 4, and 5), number of trees (5, 10, 15, 100, 500, and 1000), and maximum number of levels in each decision tree (2, 3, 4, 5, and 6) were used to grow each tree.

Gradient boosting (GB) is another method of converting weak learners into strong learners, the same strategy as random forest, but with a different method to train the model. In a random forest the trees are trained independently of each other, but in gradient boosting the k^{th} tree is trained from the first *k*-1 trees and update the residual for the i^{th} example of difference between prediction and target (real value) (Friedman, 2002). There are several ways to control the overfitting problem in random forest and gradient boosting models, such as limiting the depth of trees for splitting the input samples and finding the optimal number of samples for each node and each leaf for each tree.

The artificial neural network (ANN) is a powerful data modeling tool that can capture complex relationships between inputs and outputs, which is developed by the motivation of implementing artificial systems that can perform intelligent tasks similar to those performed by the human brain. In general, ANN can approximate any nonlinear relationship between the predictors and predictand variables. Neural networks map output variable to input features by propagating the input through a net of nodes. Next to the input layer with the input nodes, several hidden layers with various numbers of nodes and an output layer can be set up. At each node, an activation function is applied to modify the incoming signal. Along each path between two nodes a weight factor is applied to the signal. These weights are modified during the training process in a way to minimize a defined loss function (Lecun et al., 2015). The loss function provides a measure of error by comparing the final output of the network and the target values. Hyperparameters including different combination of hidden layers ([32, 32, 32], [15, 10, 5], [32, 16, 8]), activation function (Hyperbolic tangent or Tanh, sigmoid or logistic, and Rectified Linear Unit or Relu), L1 and L2 (0.0001, 0.001, 0.01) regularization which add stability and improve generalization, were used to tune the model.

Evaluation metrics

To quantify the performance of the Noah-MP model and machine learning models, mean error bias (MEB), and coefficient of determination (R^2) were used to compare the model outputs for both the Noah-MP simulations and machine learning models. In addition, MEB of Noah-MP outputs was calculated for the training set, then it was used to correct the Noah-MP simulations in the testing sets (Noah-MP_BC). Same evaluation metrics (MEB and R^2) were then used to assess how well the bias correction can improve the Noah-MP simulations. The MEB represents the average of differences between the Noah-MP simulated (and machine learning estimations) variable and associated Fluxnet observations and in which positive values indicate that the observed values are higher than the associated model outputs and negative values indicate that the observed values are lower than model results. The R^2 describes the goodness of fit between the modeled and observed variables and the magnitude of R^2 varying from 0 to 1 and the higher values indicates a closer association between the modeled and observed variables.

Results and discussion

Sensible heat flux

Figure 4.3-a shows the mean error bias values of cross validation of sensible heat flux simulations derived from Noah-MP model, bias-corrected Noah-MP (Noah-MP_BC), and the four different machine learning models. Results showed that the sensible heat flux residuals associated with Noah-MP simulations vary between -75.2 to 64.4 (W/m^2) and the Noah-MP_BC residuals vary between -61 to 53 (W/m^2) over the 69 different stations. The machine learning models substantially reduce the sensible heat flux residuals in most of the stations. The mean error bias was relatively consistent across all stations for the random forest and gradient boosting than the

Noah-MP_BC. Figure 4.3-b compare the cumulative distribution of coefficient of determinations (R^2) for the machine learning models versus Noah-MP simulations in AmeriFlux sites, with average R^2 values of 0.74, 0.77, 0.78 and 0.76 between the true values and predicted values for regularized regression, random forest, gradient boosting, and neural network models, respectively. For comparison, this value is 0.64 for the Noah-MP simulations. This result indicate that machine learning improved the R^2 score of the Noah-MP hourly sensible heat flux simulation and reduced the mean error bias in most of the Ameriflux sites. According to the mean error bias evaluation metric, results revealed that an average, the regularized regression, random forest, and gradient boosting outperformed with respect to the neural network, Noah-MP and bias corrected Noah-MP results. In addition, the average of R^2 values associated with neural network models (0.76) was relatively higher than the regularized regression models (0.74). Random forest, gradient boosting and neural network models' performance was similar according to the average of R^2 evaluation metric in all Ameriflux sites.

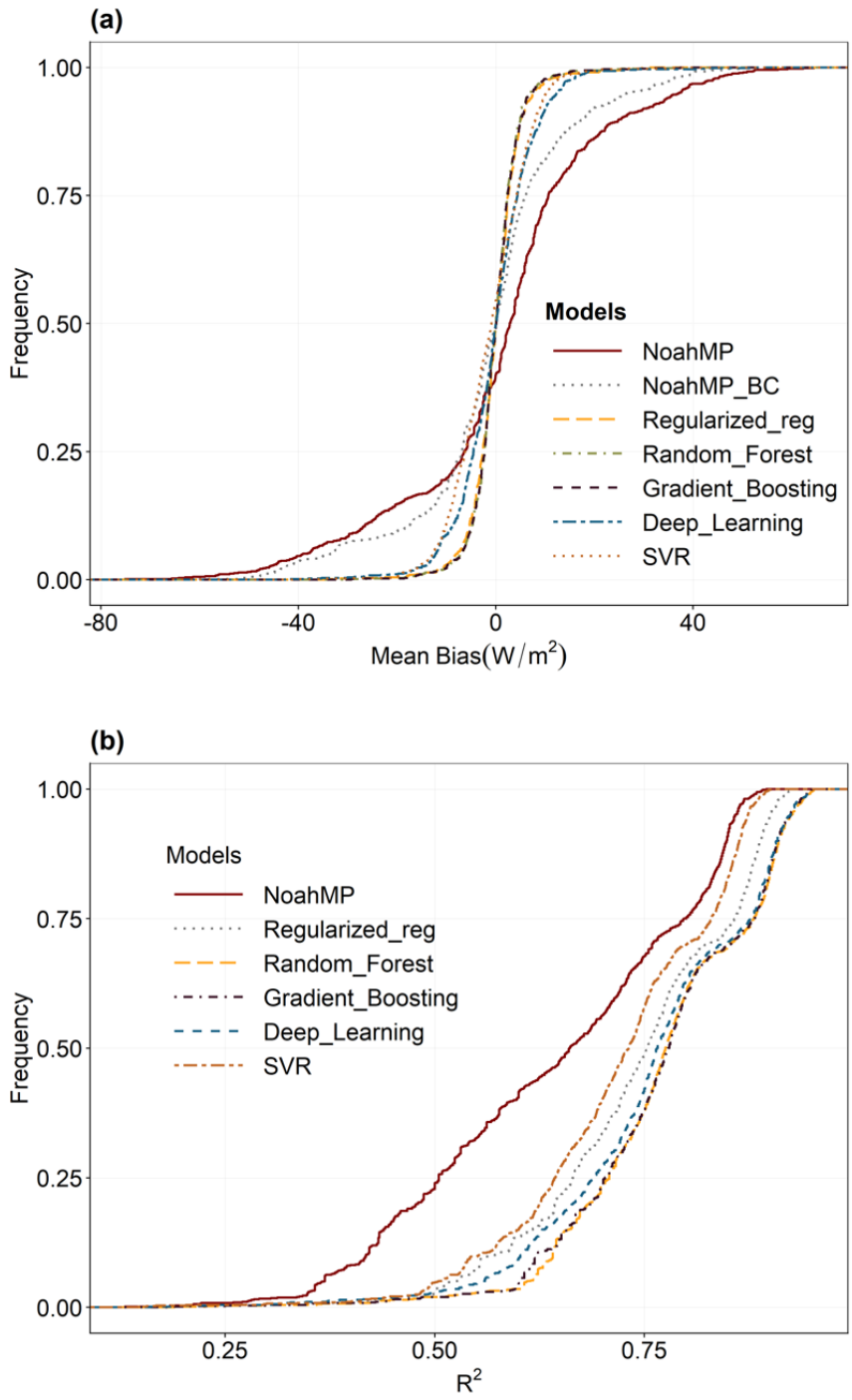

Figure 4.3 Results showing the cumulative distributions of model performance calculated as mean error bias (a) and coefficient of determinations (b) of sensible heat flux of Noah-MP, Noah-MP bias corrected (NoahMP_BC), and five machine learning models in Ameriflux sites.

Although reducing bias does not mean an improvement necessarily, to illustrate spatial pattern of the Noah-MP simulation bias versus Gradient Boosting bias, Figure 4.4 presents the mean error bias values of cross validations over all the AmericFlux sites. It should be noted that, these values were obtained from evaluating the models on validation sets. High bias values in arid regions over the southwest represent the underestimations of sensible heat fluxes at these sites (red color). Low mean error bias values indicate that the Noah-MP overestimates sensible heat fluxes at some stations such as eastern U.S. and California (green color). Furthermore, results indicated that the gradient boosting, and random forest models not only improved the R^2 metric, but the models also substantially reduced the mean error bias values (3.4 W/m^2 compared with 14.9 W/m^2) at most of the FluxNet stations. In general, findings suggested that model results with machine learning post-processing are an improvement over raw Noah-MP simulations at nearly all stations.

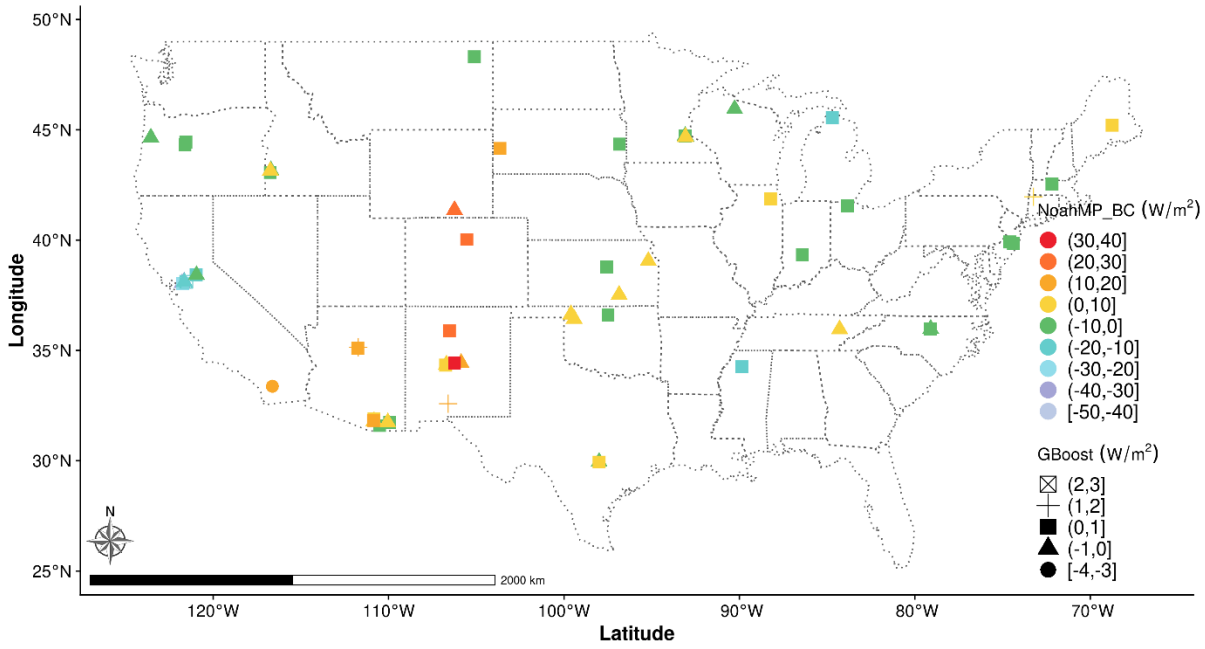

Figure 4.4 Mean residuals of sensible heat flux (W/m^2) of bias corrected NoahMP (NoahMP_BC) and gradient boosting in AmeriFlux sites across the Continental U.S. Colors indicates the residuals of NoahMP_BC and shapes indicate the residuals of gradient boosting model.

Latent heat flux

Analysis of Noah-MP-simulated latent heat flux versus observed values at the 96 FluxNet stations showed that the magnitude of mean error bias values varied between -44 to 137 (W/m^2), with the associated bias-corrected residuals (Noah-MP_BC) varying between -37 to 91 (W/m^2). The machine learning estimations show that the magnitude of biases range from -34 to 71 over the 64 FluxNet sites. Although the magnitude of the latent heat flux residuals are small at some stations, it does not necessarily indicate that the hourly and diurnal variability of latent heat flux is well-captured by the Noah-MP model. Instead, it can be related to the lower annual mean and standard deviation of latent heat flux at those specific stations. For example, the US-Rwf and US-

Aud sites, which show lower simulation bias among the other stations, have lower annual vegetation coverage and the maximum value of the normalized difference vegetation index (NDVI) is less than 0.4 in both stations (0.1<NDVI<0.4). As a result, the magnitude of annual mean latent heat flux is lower at these sites (40 and 20 W/m^2, respectively) and the standard deviation is about 60 W/m^2 and 38 W/m^2 at US-Rwf site and US-Aud, respectively. In contrast, in the US-Twt site the higher annual mean (83 W/m^2) and standard deviation (112 W/m^2) of latent heat flux is associated with high annual vegetation coverage and variability (0.25<NDVI<0.75).

The machine learning models improved the Noah-MP-simulated latent heat fluxes over the AmeriFlux sites. Mean error bias (Fig. 4.5-a) of regularized regression, random forest, and gradient boosting models were relatively lower than the mean error bias associated with neural network, Noah-MP, and Noah-MP_BC models. This improvement was substantial at some stations such as US-Twt, US-Tw4, US-Tw3, US-Tw1, US-Myb, and US-Snd; Although neural network with tangent hyperbolic activation function, and 3 hidden layers and 0.01 of L1 configuration showed better performance (lower mean error bias and higher R-squared) relative to other configurations, the estimated error tend to slightly larger (6.4 W/m^2) relative to the other machine learning models (4.5 W/m^2). Figure 4.5-b summarizes the cumulative R^2 values for each model for over the AmeriFlux stations, representing how well model estimates agree with hourly surface flux observations at each station. Overall, according to the R^2 of testing data set, Noah-MP simulations adequately captured the latent heat fluxes with average R^2 value of 0.57. At the US-Var, US-Ton, US-Whs, US-SRG, US-Ho1 sites, the Noah-MP simulations show larger (>0.7) R^2 values, whereas at the US-Tw1, US-Tw4, US-Myb, US-GLE, and US-GBT sites this evaluation metric was lower than 0.3. The dominant soil class in these sites are Silty and Organic Materials, and results showed that Noah-MP surface heat flux estimates are highly biased in these soil types. Investigating the

effects of soil texture or error associated with the forcing data is beyond the scope of current study; however, soil texture can potentially plays an important role in driving this bias. Overall, these results suggested that, relative to the Noah-MP model outputs, the gradient boosting machine learning model was able to explain much of the variance of latent heat fluxes at all stations with average R^2 values of 0.77.

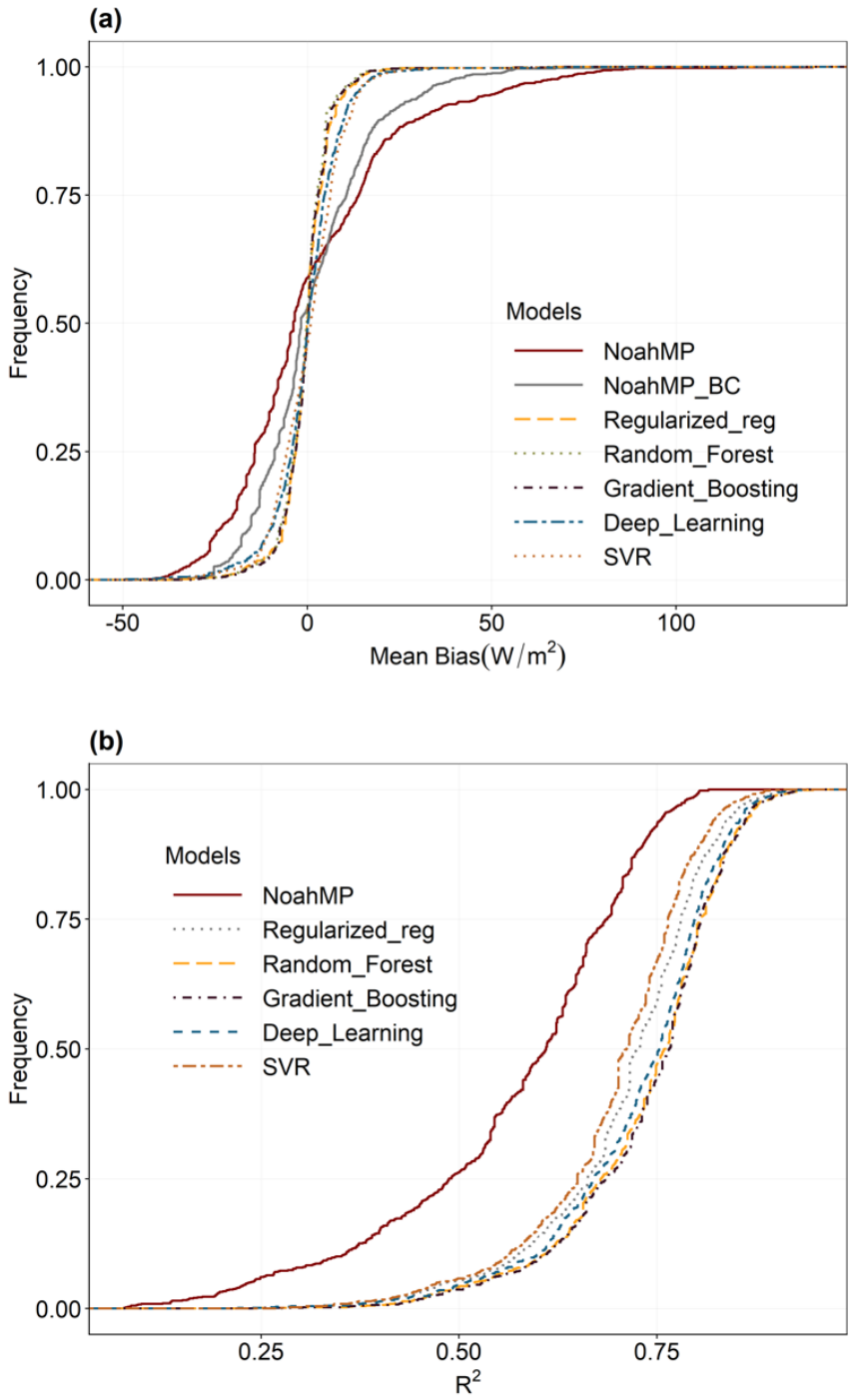

Figure 4.5 Results showing the cumulative distributions of model performance calculated as mean error bias (a) and coefficient of determinations (b) of latent heat flux of Noah-MP, Noah-MP bias corrected (NoahMP_BC), and five machine learning models in Ameriflux sites. These results are based on the average of evaluation statistics obtained in every holdout testing segment.

To represent the spatial distribution of bias associated with Noah-MP_BC and gradient boosting models, Fig. 4.6 shows the spatial distribution of latent heat flux biases for these two models. Noah-MP shows the larger overestimation (green shaded) of latent heat fluxes at the northeast sites. This is not surprising because this area is covered by vegetation and the evapotranspiration is higher in this area, therefore the residual values are relatively higher. In addition, the underestimations are shown at the central region, mountain region, Florida, California, and New Jersey. Most notably, the bias values of latent heat flux can be larger than the 55 (W/m^2) during the daytime. (Zheng et al., 2015) found similar results for Noah model simulations, which verifies that within the model simulations the largest amount of latent heat is released to the atmosphere during the daytime.

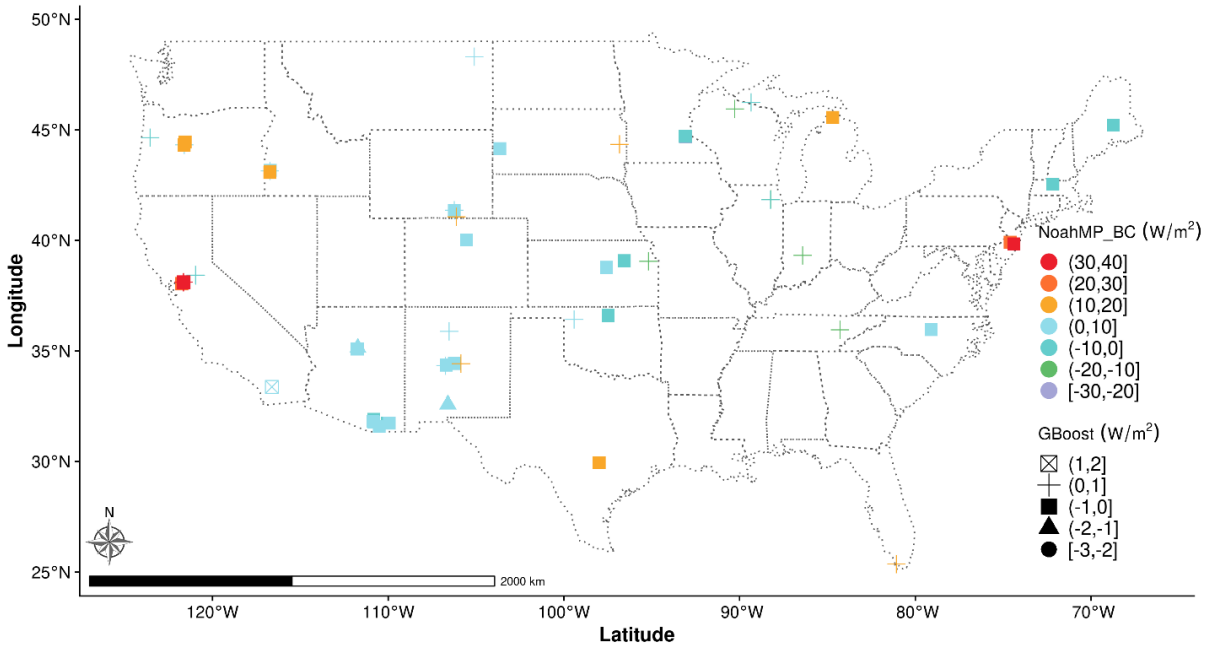

Figure 4.6 Mean residuals of latent heat flux (W/m^2) of bias corrected NoahMP (NoahMP_BC) and gradient boosting in AmeriFlux sites across the Continental U.S. Colors indicates the residuals of NoahMP_BC and shapes indicate the residuals of gradient boosting model.

It can be concluded that, compared with Noah-MP bias-corrected latent heat flux estimates, the gradient boosting models have lower mean error bias values over all 64 stations used for evaluating latent heat flux simulations. Considering mean error bias values and R^2 altogether, between the four machine learning models used, the gradient boosting, and random forest models provided the best fit to observations, meaning that these models are capable of reducing Noah-MP-simulated latent heat flux bias in terms of diurnal and seasonal variability over all stations. On the whole, latent heat mean error bias values averaged across all stations for the Noah-MP, Noah-MP bias corrected, and gradient boosting models were 1.64, 1.56, and -0.06 (W/m^2) respectively. As a result, in the majority of stations, machine learning provides a substantial improvement over the Noah-MP surface heat flux simulations, which are important for the other numerical modeling frameworks such as Weather Research and Forecasting (WRF) and WRF-Hydro model, the latter of which forms the basis for the National Water Model (NWM).

Analyzing Noah-MP simulation bias

Noah-MP-modeled surface heat fluxes are compared with the observed values at AmeriFlux sites. To explore the Noah-MP simulation biases associated with each land use and land cover, we further analyzed the magnitude of surface heat flux simulations at different land cover types. Table 4.4 summarizes the average of bias values of surface heat flux simulations for different land use and land cover types over all the FluxNet sites. Results indicated that, on average, Noah-MP overestimates sensible heat fluxes with the mean error bias values of -12.5 and -11.8 (W/m^2) for Deciduous Broadleaf Forest and Evergreen Needleleaf Forest land covers respectively, but underestimates by 17.4 (W/m^2) at Woody Savannas land cover. Furthermore, the largest amount of latent heat overestimation is shown at the Mixed Forest (-12.3 W/m^2) sites and

the highest values of underestimations were observed over Evergreen Broadleaf Forest (24.3 W/m^2). The results show that the magnitude of the bias values associated with sensible heat fluxes are larger at Deciduous Broadleaf Forest, Evergreen Needleleaf Forest, and Woody Savannas, and the latent heat flux bias values are larger at the permanent wetlands and mixed forest sites. It should be noted that only one site is located at permanent wetlands land cover which is not satisfactory for statistical evaluation of the Noah-MP simulations. Figure 4.7 shows cumulative distributions of mean error bias (W/m^2) comparing the performance of the Noah-MP sensible heat flux (Fig. 4.7-a) and latent heat flux (Fig. 4.7-b) simulations associated with different land use and land cover (LULC) classes across the AmeriFlux sites. Overall, the cumulative plots show that how negative and positive values of residuals of surface heat flux simulations distributed at each LULC types. For instance, approximately at 35% (80%) of the sites, sensible (latent) heat flux simulations showed bias values less than or equal zero in croplands LULC (Fig. 4.7-a, -b, respectively).

Table 4.4 Average mean error bias values of sensible and latent heat flux simulations at different land use and land covers over the AmeriFlux sites. n represents the number of years used for cross validation at all sites associated with each land cover type.

Land Use and Land Cover	Sensible Heat (W/m^2)		Latent Heat Flux (W/m^2)	
	Mean	n	Mean	n
Deciduous Broadleaf Forest	-12.5	75	-0.1	24
Evergreen Needleleaf Forest	-11.5	26	-10.0	82
Savannas	-7.2	96	-11.1	94
Evergreen Broadleaf Forest	-6.8	4	-10.0	4
Croplands	-2.1	206	-2.9	255
Mixed Forests	-2.0	81	-12.3	30
Grasslands	7.4	240	9.5	273
Open Shrublands	13.6	46	-2.0	46
Woody Savannas	17.4	71	6.1	65
Permanent wetlands	--	--	22.0	5

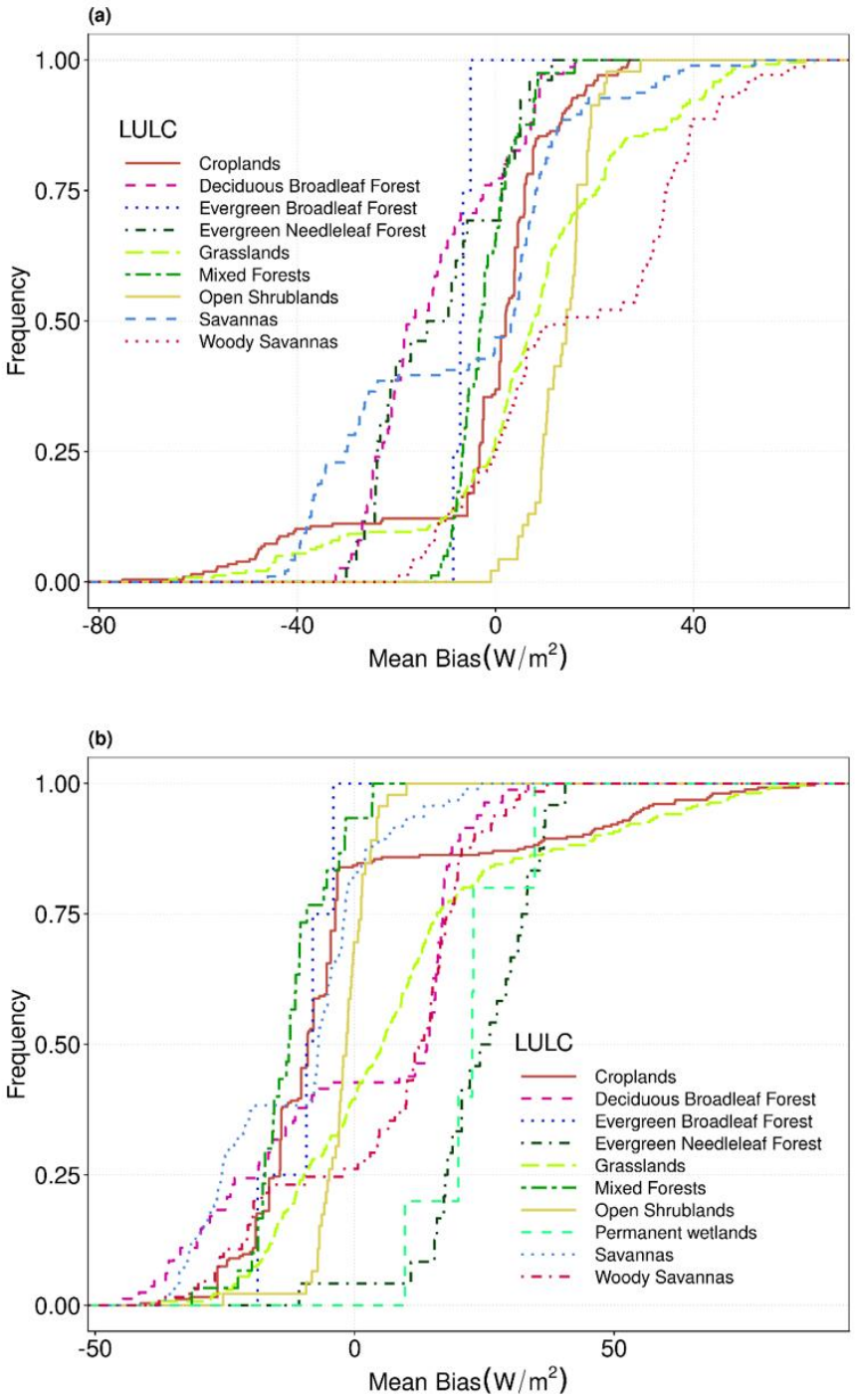

Figure 4.7 Cumulative mean error bias (W/m^2) values of sensible heat flux (a) and latent heat flux (b) of Noah-MP simulations at different land use land covers (LULC).

To explore the possible reasons for the range in bias of the surface heat flux simulations, we further investigated the Noah-MP simulation bias values at different soil categories (Table 4.4). Results indicate that the higher values of overestimations (approximately -44.5 W/m^2) are associated with Organic Material and the model underestimates the sensible heat fluxes by 14.5 (W/m^2) at the Sandy Loam soil type. Similarly, latent heat fluxes are overestimated by 19.1 (W/m^2) at silty Clay Loam soil type but underestimated by approximately 55.7 (W/m^2) at the Organic Material soil type. The lowest values of sensible heat fluxes are shown at Silt Loam (-2.4 W/m^2) and Clay (+2.4 W/m^2). Likewise, the lowest values of latent heat fluxes are linked to Loam soil category (~ 0.0 W/m^2). To summarize these results, Table 4.4 illustrates the mean error bias of Noah-MP simulation for different soil classes. To illustrate the distribution of Noah-MP-simulated surface heat fluxes at different soil categories, figure 4.8 summarize cumulative residuals comparing the performance of sensible and latent heat flux simulations across all sites.

Table 4.5 Average mean error bias values of surface heat flux simulations at different soil classes over the AmeriFlux sites. n represents the number of years used for cross validation at all sites associated with each soil type.

Soil Categories	Sensible Heat (W/m^2)		Latent Heat Flux (W/m^2)	
	Mean	n	Mean	n
Organic Materials	-44.5	33	-19.1	36
Sand	-10.2	74	-9.4	406
Silt Loam	-2.4	355	0.0	133
Clay	2.4	12	6.8	12
Silty Clay Loam	6.3	31	9.0	159
Loam	6.8	130	17.6	79
Sandy Loam	14.5	210	55.7	53

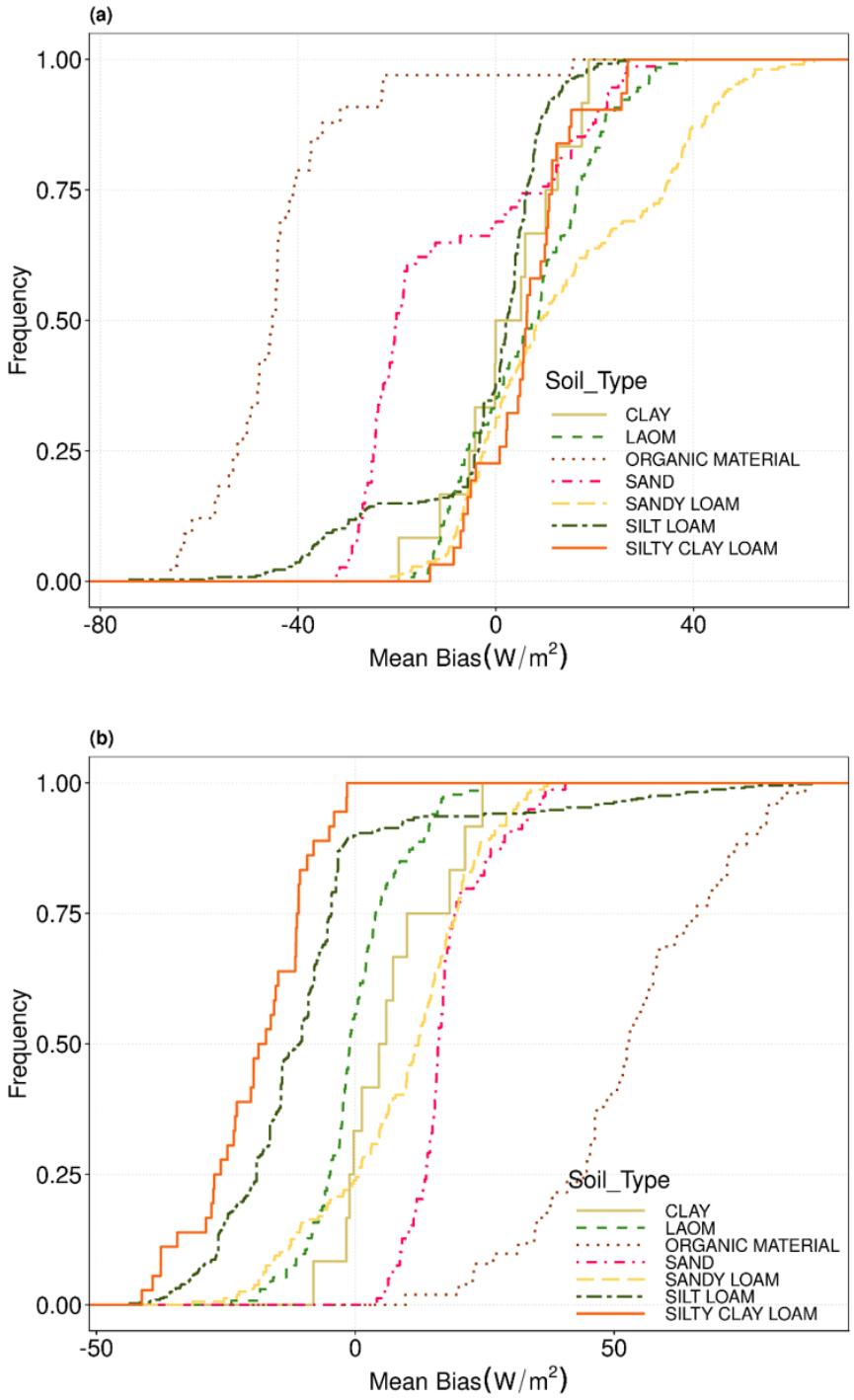

Figure 4.8 Cumulative distribution of mean error bias (W/m^2) of sensible heat flux (a) and latent heat flux (b) of Noah-MP simulations associated with different soil categories.

As the data used in this study are time series, it is critically important to capture the timing of extreme events. As timing information is not represented in metrics showing central tendency, other statistical evaluation metrics should be considered that account for the time component of the data. We further investigated the time series of surface heat flux observations and associated model output bias at different sites to explore the discrepancy of sensible and latent heat flux simulations at different land cover types (Fig. 4.9). Time series of the AmeriFlux observations and associated simulation's bias values show different patterns during different seasons and over different vegetation types. During the warm season, the Noah-MP sensible and latent heat flux simulations are substantially higher (>200 W/m^2) than the associated AmeriFlux observations at Woody Savannas vegetation type (Fig. 4.9i-j). In addition, in cropland and close shrubland sites (Fig. 4.9a-h), when surface heat flux magnitudes are higher than roughly 200 (W/m^2), Noah-MP yields lower sensible heat flux estimations (positive bias values), indicating that Noah-MP underestimates the maximum values (Pilotto et al., 2015). Similarly, Noah-MP yields larger sensible heat values resulting in negative bias values where the surface heat fluxes are lower than ~200 (W/m^2), such as from January through July in the CRT site (Fig. 4.9a-h), and from July through September at the Ro1 site (Fig. 4.9c), The pattern of the surface heat fluxes changes around July, and consequently the Noah-MP bias is shifted at this time (Fig. 4.9a-d). These results suggest that surface energy partitioning and associated land-atmosphere interaction are more strongly controlled by the state of the vegetation at cropland areas. This is an important result because if vegetation characteristics forcing the atmosphere and near surface atmosphere are connected to land cover and vegetation type, then we can conclude that land management and crop-growth timing could play a significant role in the strength of that forcing (Bagley et al., 2017).

Figure 4.9 Time series of NoahMP-simulated sensible heat flux (left) and latent heat flux (right), observed (orange), and associated bias (blue) of surface heat fluxes at different AmeriFlux sites with different vegetation types.

Soil Moisture

Noah-MP simulations and AmeriFlux observations have coinciding soil data at 0.1, and 0.3 meters. As a result, following the observed data availability, the model's outputs verified at only 0.1-, and 0.3-, depths. Fig. 4.10 presents the root mean squared error (RMSE) values associated with the Noah-MP simulations, and five machine learning models at 0.1-m, and root zone (0.3-m). Among the 36 stations observing soil moisture at 0.1-m, 25 stations were used to evaluate 0.1-m soil moisture model outputs and 11 stations were used to assess soil moisture simulations at 0.3-m soil depth. The difference between the number of stations for each depth was due to the availability of the data at that site. Results from the Noah-MP model in association with data from soil moisture observations shows that the error of 0.1-m soil moisture simulations varies between -10% to 10% over all stations, with values being slightly higher at 0.3-m (figures not shown). Overall, results showed that the lowest errors (as quantified by RMSE) were seen in machine learning models (5%) relative to Noah-MP simulations (9.2%) over all stations.

In addition, comparing observed and simulated soil moisture by Noah-MP and the five machine learning models indicate that the machine learning estimations agree well with the observed maximum and minimum soil moisture values. Since the soil moisture time series are not fluctuating substantially in hourly and daily scale, especially at the 0.3-m layers, the machine learning models show smaller bias values at most of the sites. At lower depths the soil moisture estimations show lower bias values and generally agree well with that of the observed data; however, the performance of the machine learning models can be affected by the lack of the hydro-climate diversity in the training data (O & Orth, 2021). As a result, evaluating the trained machine learning models for the next years with different hydro-climate conditions could be a promising solution to estimate more reliable soil moisture data. Additionally, comparing the Noah-MP

simulation versus observed data indicated that the simulated volumetric soil moisture at 0.3-m is overpredicted. For example, at the Ar2, Ar1, Wkg, SRG, SRM, and Blk sites, the mean values of Noah-MP model outputs were two times higher than the observations; however, the machine learning-based estimations, especially neural network, and gradient boosting models tended to capture the minimum, maximum values better than Noah-MP. In some stations (i.e., AR2, Me6, Blk, and UMd sites), neural network outputs stand out with higher coefficients of determination and lower RMSE values compared to the Noah-MP outputs.

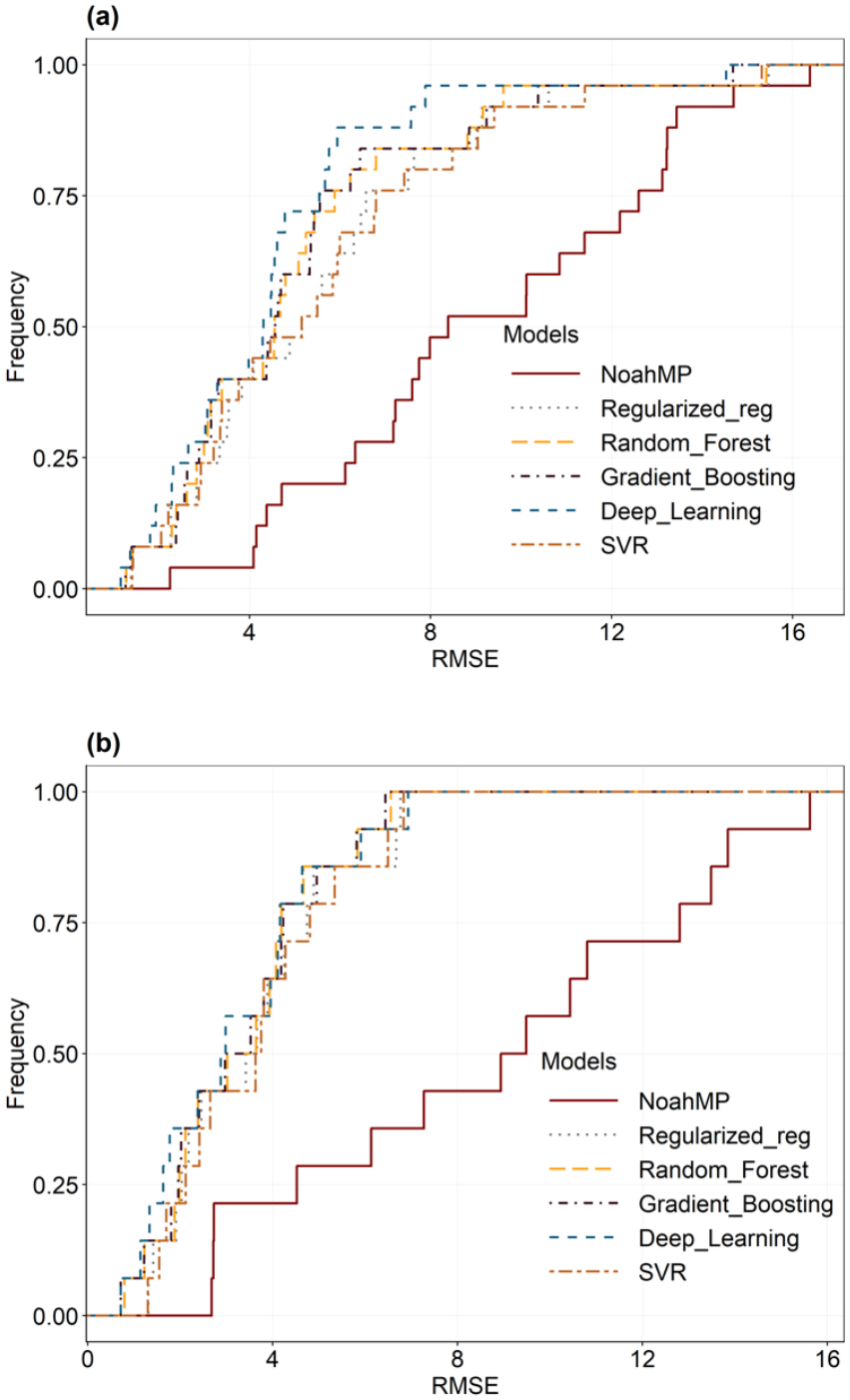

Figure 4.10 Accumulated RMSE values of Noah-MP-simulated soil moisture content (%) and associated estimations from five machine learning models at 0.1m (a) and 0.3m (b) depths.

In this study five machine learning models were evaluated vs. the Noah-MP-simulated soil moisture content at different AmeriFlux sites. Comparing to observed soil moisture content, Noah-MP simulations show an overall better agreement with the observed soil moisture data over the study stations (X. Cai, Yang, Xia, et al., 2014). Machine learning provided improved soil moisture estimations in each site. The error associated with the Noah-MP simulations could be due to error in NLDAS forcing data, systematic error in the observed data, and complexity of physical processes involved in the Noah-MP model that makes it difficult to identify the interaction effects within the model simulations. Furthermore, the Noah-MP surface heat flux simulations showed higher discrepancy at some land covers and soil categories; therefore, the errors associated with the partitioning of moisture fluxes could potentially be due to inconsistency in latent heat flux simulations.

Soil Temperature

Figure 4.11 shows the results of accumulated RMSE values of simulated soil temperature at 0.1 meter using Noah-MP and five machine learning models against observations. Results indicate that the post-processing Noah-MP-estimated soil temperature data using machine learning models well agreed with observed data at 0.1-m. Although the mean error bias values show that the Noah-MP model overpredicts by approximately -1.5 C° at most locations, overall, results highlight the impressive performance of temperature simulations by Noah-MP model. Likewise, all machine learning models yield lower mean error bias values (-0.03 – 0.05 C°) than Noah-MP for all the study sites. Compared with observed values, results indicate that Noah-MP overpredicts soil temperature values across almost all the study sites, where the post-processing results showed a good match with observed temperatures, especially at minimum and maximum soil temperate values. These results show that overestimating high temperatures and underestimating low

temperatures within the model simulations could be the main source of the biases in the Noah-MP model estimations. The lower RMSE values, especially for gradient boosting model, reveal the performance of machine learning models in post processing soil temperature predictions.

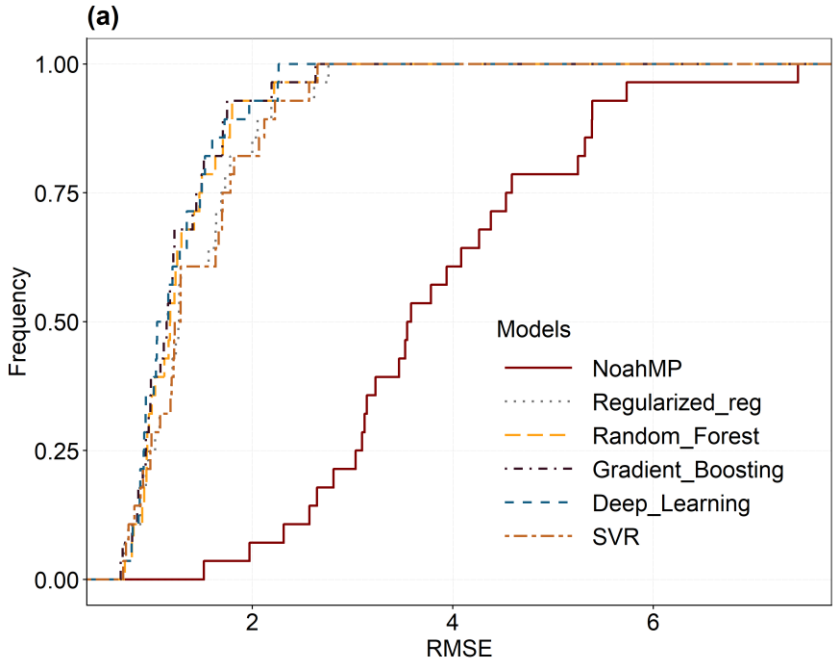

Figure 4.11 Illustration of accumulated RMSE of 0.1-m soil temperature of Noah-MP simulations and five machine learning model estimations.

Conclusion

This paper focuses on estimation of surface heat flues and soil moisture and temperature at selected AmeriFlux sites using the Noah-MP land surface model and post-processing Noah-MP outputs using five machine learning models, including: (1) penalized regression (Ridge regression), (2) support vector regression (SVR), (3) random forest, (4) gradient boosting, and (5) neural network. The Noah-MP simulations were forced by NLDAS-2 data and the simulation outputs were subsequently used as input features for the machine learning models. The best

machine learning model for surface heat flux estimates was found to be gradient boosting, based on the high coefficient of determination for sensible heat flux and latent heat flux (R^2=0.75 for both variables) over all sites, along with lower bias values (-0.1 W/m^2 for both latent heat flux and sensible heat flux, respectively). On average, residuals of Noah-MP bias corrected values were -3.7 W/m^2 for the latent heat flux and 2.8 W/m^2 for the sensible heat flux. In addition, results indicate that stand-alone machine learning algorithms were capable of estimating the surface heat fluxes with a higher level of accuracy compared with the Noah-MP model; however, unlike the Noah-MP simulations of extreme values, the machine learning models were not able to capture the extreme values of surface heat fluxes. This result is opposite for soil temperature simulations, which reveal a potential misrepresentation of soil temperature processes by the Noah-MP model.

Among the five machine learning models, gradient boosting resulted in higher R^2 in calibrating surface heat fluxes. Moreover, the findings of the current study showed that the accuracy of surface heat flux simulations by Noah-MP are influenced by land cover and soil type. The results of this research also showed that the Noah-MP model simulation is not able to accurately capture soil moisture at deeper soil layers. Our findings have an important implication that machine learning models are able to reduce the bias of the physics-based models and improve their overall performance. Finally, it is important to note that the machine learning models developed in this study are valid for individual stations only. Advanced deep learning models like Long Short-Term Memory (LSTM) in combination with Convolutional Neural Network (CNN) are recommended so that it may be possible to estimate spatiotemporal surface heat flux variables over large areas.

CHAPTER V

CONCLUSION

Accurate representation of land surface properties is essential to conduct reliable and precise numerical modelling experiments. Because land cover types change over space and time, land use related parameters such as vegetation, green vegetation fraction, stomatal resistance, albedo, and surface roughness will lead to differences in simulated sensible heat flux, latent heat flux and many other variables depending on vegetation and land use parameters. Our analysis suggested higher sensitivity of the WRF simulations such as surface heat fluxes and precipitation over the southeast United States, which indicated the importance of utilizing near-real-time LULC within the NWP simulations. Therefore, annually updated satellite observations of land use products such as MODIS MCD12Q1 data provides accurate representation of Earth's surface on different spatial resolutions and is a desirable source of land cover data. MCD12Q1 data was used in this study for regional WRF modelling and investigated the sensitivity and accuracy of a land cover dataset that has been used in several atmospheric modelling systems for the southeast United States.

Comparisons of default WRF land cover data with new land cover dataset derived from MODIS images showed that the WRF default data are not up-to-date and accurate, especially for southeast. We were able to show that surface heat and moisture fluxes are more sensitive to land cover change. Additionally, the model simulations showed that the spatial extent and intensity of the surface heat and moisture flux simulations were different for the dry, normal, and wet years.

Two simulation results were evaluated using point measurements of daily accumulated precipitation and daily average 2-m temperature data over the southeast. Results indicated that the updated land cover data did not improve the temperature and precipitation simulations over the many sites. This result implies the importance of developing machine learning models to post-process the NWP model outputs to enhance the accuracy of the weather predictions in regional and local scales.

The results of this work showed that land surface energy and moisture flux simulations were impacted due to change in land cover change. Because the WRF default LULC data were extracted based on long-term historical land cover data, the differences in default and updated land cover data were substantial, especially over the southeast. Transitions from cropland in the default LULC data to the woody savannas in the updated land cover data should impact the climatology of the region, especially in terms of maximum and average summer temperatures. The result of WRF-simulated surface heat fluxes with updated LULC data showed that latent heat flux, sensible heat flux, temperature, and precipitation indicated more sensitivity to local land cover changes; however, validation of the WRF-simulated precipitation and near-surface temperature did not show an appreciable improvement in the model simulations due to updating the land surface characteristics. More investigations should involve an ensemble simulation whereby a multi-physics is generated with updated land use to realize the scientific reasons of the decreased WRF performance by providing more realistic land surface characteristics.

The post-processing Noah-MP surface heat flux and moisture flux procedure presented here is one of the techniques currently available for combining physics-based models and data-driven models. Several other methods of combining the benefits of machine learning techniques with the benefits of physically realistic are in development. For example, Pelissier et al., (2020)

use Gaussian Processes to predict error between modeled and observed soil moisture, which allows machine learning models to be used dynamically within a land surface model to correct the soil moisture state at each timestep of a simulation. Results of this study indicated that data-driven methods, especially gradient boosting, are able to substantially improve the land surface model simulations such as surface heat and moisture fluxes. Implementing post-processing Noah-MP model outputs is relatively straightforward compared to other techniques such as adding physics into machine learning code or using machine learning to dynamically update the state variables.

REFERENCES

Al, N. E. T. (2017). The Sensitivity of WRF Daily Summertime Simulations over West Africa to Alternative Parameterizations . Part II : Precipitation. 215–233. https://doi.org/10.1175/MWR-D-15-0294.1

Aylas, Y. G. R., de Souza Campos Correa, W., Santiago, A. M., Reis Junior, N. C., Albuquerque, T. T. de A., Santos, J. M., & Moreira, D. M. (2020). Influence of land use on the performance of the WRF model in a humid tropical climate. Theoretical and Applied Climatology, 141(1–2), 201–214. https://doi.org/10.1007/s00704-020-03187-3

Bagley, J. E., Kueppers, L. M., Billesbach, D. P., Williams, I. N., Biraud, S. C., & Torn, M. S. (2017). The influence of land cover on surface energy partitioning and evaporative fraction regimes in the U.S. Southern Great Plains. Journal of Geophysical Research, 122(11), 5793–5807. https://doi.org/10.1002/2017JD026740

Bin Li, & Avissar, R. (1994). The impact of spatial variability of land-surface characteristics on land-surface heat fluxes. In Journal of Climate (Vol. 7, Issue 4, pp. 527–537). https://doi.org/10.1175/1520-0442(1994)007<0527:tiosvo>2.0.co;2

Boisier, J. P., De Noblet-Ducoudré, N., Pitman, A. J., Cruz, F. T., Delire, C., Van Den Hurk, B. J. J. M., Van Der Molen, M. K., Mller, C., & Voldoire, A. (2012). Attributing the impacts of land-cover changes in temperate regions on surface temperature and heat fluxes to specific causes: Results from the first LUCID set of simulations. Journal of Geophysical Research Atmospheres, 117(12). https://doi.org/10.1029/2011JD017106

Bright, R. M., Eisner, S., Lund, M. T., Majasalmi, T., Myhre, G., & Astrup, R. (2018). Inferring Surface Albedo Prediction Error Linked to Forest Structure at High Latitudes. Journal of Geophysical Research: Atmospheres, 123(10), 4910–4925. https://doi.org/10.1029/2018JD028293

Cai, S., Liu, D., Sulla-Menashe, D., & Friedl, M. A. (2014). Enhancing MODIS land cover product with a spatial-temporal modeling algorithm. Remote Sensing of Environment, 147, 243–255. https://doi.org/10.1016/j.rse.2014.03.012

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, 119(1), 23–38. https://doi.org/10.1002/2013JD020792

Cai, X., Yang, Z., Xia, Y., Huang, M., Wei, H., & Leung, L. R. (2014). Assessment of simulated water balance from Noah, Noah-MP, CLM, and VIC over CONUS using the NLDAS test bed. Journal of Geophysical Research: Atmospheres, 119(24), 13–751. https://doi.org/10.1038/175238c0

Chang, M., Liao, W., Wang, X., Zhang, Q., Chen, W., Wu, Z., & Hu, Z. (2020). An optimal ensemble of the Noah-MP land surface model for simulating surface heat fluxes over a typical subtropical forest in South China. Agricultural and Forest Meteorology, 281(October 2019), 107815. https://doi.org/10.1016/j.agrformet.2019.107815

Chen, F., & Dudhia, J. (2001a). Coupling an advanced land surface-hydrology model with the Penn-State-NCAR MM5 modeling system. Part II: Preliminary model validation. Monthly Weather Review, 129(4), 587–604. https://doi.org/10.1175/1520-0493(2001)129<0587:CAALSH>2.0.CO;2

Chen, F., & Dudhia, J. (2001b). Coupling and advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. Monthly Weather Review, 129(4), 569–585. https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2

Chen, M., Vernon, C. R., Graham, N. T., Hejazi, M., Huang, M., Cheng, Y., & Calvin, K. (2020). Global land use for 2015–2100 at 0.05° resolution under diverse socioeconomic and climate scenarios. Scientific Data, 7(1), 1–11. https://doi.org/10.1038/s41597-020-00669-x

DelSole, T., & Banerjee, A. (2017). Statistical seasonal prediction based on regularized regression. Journal of Climate, 30(4), 1345–1361. https://doi.org/10.1175/JCLI-D-16-0249.1

Dickinson, R. E., Berry, J. A., Bonan, G. B., Collatz, G. J., Field, C. B., Fung, I. Y., Goulden, M., Hoffmann, W. A., Jackson, R. B., Myneni, R., Sellers, P. J., & Shaikh, M. (2002). Nitrogen controls on climate model evapotranspiration. Journal of Climate, 15(1), 278–295. https://doi.org/10.1175/1520-0442(2002)015<0278:ncocme>2.0.co;2

Dickinson, R. E., Shaikh, M., Bryant, R., & Graumlich, L. (1998). Interactive canopies for a climate model. Journal of Climate, 11(11), 2823–2836. https://doi.org/10.1175/1520-0442(1998)011<2823:ICFACM>2.0.CO;2

Dyer, J. (2011). Analysis of a Warm-Season Surface-Influenced Mesoscale Convective Boundary in Northwest Mississippi. Journal of Hydrometeorology, 12(5), 1007–1023. https://doi.org/10.1175/2011jhm1326.1

Dyer, J. L., & Rigby, J. R. (2020). Assessing the sensitivity of lower atmospheric characteristics to agricultural land use classification over the Lower Mississippi River Alluvial Valley. Theoretical and Applied Climatology, 142(1–2), 305–320. https://doi.org/10.1007/s00704-020-03318-w

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., & Tarpley, J. D. (2003). Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. Journal of Geophysical Research: Atmospheres, 108(22), 1–16. https://doi.org/10.1029/2002jd003296

Esteve, B. J. (2015). Land use influence in WRF model. A high resolution mesoscale modeling over Oriental Pyrenees. Universitat de Barcelona. http://diposit.ub.edu/dspace/bitstream/2445/105800/1/TFM_Bernat_Jimenez.pdf

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sensing of Environment, 114(1), 168–182. https://doi.org/10.1016/j.rse.2009.08.016

Friedman, J. H. (2002). Stochastic gradient boosting. Computational Statistics and Data Analysis, 38(4), 367–378. https://doi.org/10.1016/S0167-9473(01)00065-2

Godfrey, C. M., & Stensrud, D. J. (2010). An empirical latent heat flux parameterization for the Noah land surface model. Journal of Applied Meteorology and Climatology, 49(8), 1696–1713. https://doi.org/10.1175/2010JAMC2180.1

Gogoi, P. P., Vinoj, V., Swain, D., Roberts, G., Dash, J., & Tripathy, S. (2019). Land use and land cover change effect on surface temperature over Eastern India. Scientific Reports, 9(1), 1–10. https://doi.org/10.1038/s41598-019-45213-z

Hartley, A. J., MacBean, N., Georgievski, G., & Bontemps, S. (2017). Uncertainty in plant functional type distributions and its impact on land surface models. Remote Sensing of Environment, 203(September), 71–89. https://doi.org/10.1016/j.rse.2017.07.037

Holt, T. R., Niyogi, D., Chen, F., Manning, K., LeMone, M. A., & Qureshi, A. (2006). Effect of land-atmosphere interactions on the IHOP 24-25 May 2002 convection case. Monthly Weather Review, 134(1), 113–133. https://doi.org/10.1175/MWR3057.1

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. Journal of Geophysical Research Atmospheres, 113(13), 113. https://doi.org/10.1029/2008JD009944

Janjic, Z., Gall, R., & Pyle, M. E. (2010). Scientific Documentation for the NMM Solver. In NCAR technical note (Vol. 53). https://doi.org/10.1017/CBO9781107415324.004

Janjić, Z. I. (1994). The step-mountain eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes. Monthly Weather Review, 122(5), 927–945. https://doi.org/https://ui.adsabs.harvard.edu/link_gateway/1994MWRv..122..927J/doi:10.1175/1520-0493(1994)122%3C0927:TSMECM%3E2.0.CO;2

Jefferson, J. L., Maxwell, R. M., & Constantine, P. G. (2017). Exploring the Sensitivity of Photosynthesis and Stomatal Resistance Parameters in a Land Surface Model. Journal of Hydrometeorology, 18(3), 897–915. https://doi.org/10.1175/jhm-d-16-0053.1

Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., Roy, D. P., & Morisette, J. T. (2002). An overview of MODIS Land data processing and product status. Remote Sensing of Environment, 83(1–2), 3–15. https://doi.org/10.1016/S0034-4257(02)00084-6

Kirthiga, S. M., & Patel, N. R. (2018). Impact of updating land surface data on micrometeorological weather simulations from the WRF model. Atmosfera, 31(2), 165–183. https://doi.org/10.20937/ATM.2018.31.02.05

Kishtawal, C. M., Niyogi, D., Tewari, M., Pielke, R. A., & Shepherd, J. M. (2010). Urbanization signature in the observed heavy rainfall climatology over India. International Journal of Climatology, 30(13), 1908–1916. https://doi.org/10.1002/joc.2044

Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. Hydrology and Earth System Sciences, 23(10), 4323–4331. https://doi.org/10.5194/hess-23-4323-2019

Kumar, A., Chen, F., Barlage, M., Ek, M. B., & Niyogi, D. (2014). Assessing impacts of integrating MODIS vegetation data in the weather research and forecasting (WRF) model coupled to two different canopy-resistance approaches. Journal of Applied Meteorology and Climatology, 53(6), 1362–1380. https://doi.org/10.1175/JAMC-D-13-0247.1

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Levis, S., Gordon, B. B., Erik Kluzek, Thornton, P. E., Jones, A., Sacks, W. J., & Kucharik, C. J. (2012). Interactive Crop Management in the Community Earth System Model (CESM1): Seasonal influences on land-atmosphere fluxes. Journal of Climate, 25(14), 4839–4859. https://doi.org/10.1175/JCLI-D-11-00446.1

Li, H., Zhang, H., Mamtimin, A., Fan, S., & Ju, C. (2020). A New Land-Use Dataset for the Weather Research and Forecasting (WRF) Model. April 1992.

Li, Jianduo, Chen, F., Zhang, G., Barlage, M., Gan, Y., Xin, Y., & Wang, C. (2018). Impacts of Land Cover and Soil Texture Uncertainty on Land Model Simulations Over the Central Tibetan Plateau. In Journal of Advances in Modeling Earth Systems (Vol. 10, Issue 9, pp. 2121–2146). https://doi.org/10.1029/2018MS001377

Li, Jianduo, Miao, C., Wei, W., Zhang, G., Hua, L., Chen, Y., & Wang, X. (2021). Evaluation of CMIP6 Global Climate Models for Simulating Land Surface Energy and Water Fluxes During 1979–2014. Journal of Advances in Modeling Earth Systems, 13(6), 1–32. https://doi.org/10.1029/2021MS002515

Li, Jiayang, Zheng, X., Zhang, C., & Chen, Y. (2018). Impact of Land-Use and Land-Cover Change on Meteorology in the Beijing – Tianjin – Hebei Region from 1990 to 2010. https://doi.org/10.3390/su10010176

Li, X., Mitra, C., Dong, L., & Yang, Q. (2018). Understanding land use change impacts on microclimate using Weather Research and Forecasting (WRF) model. Physics and Chemistry of the Earth, 103, 115–126. https://doi.org/10.1016/j.pce.2017.01.017

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. R News 2. 3(December 2002), 18–22.

Liu, X., Chen, F., Barlage, M., Zhou, G., & Niyogi, D. (2016). Noah-MP-Crop: Introducing dynamic crop growth in the Noah-MP land surface model. Journal of Geophysical Research, 121(23), 13,953-13,972. https://doi.org/10.1002/2016JD025597

López-Bravo, C., Caetano, E., & Magaña, V. (2018). Forecasting summertime surface temperature and precipitation in the Mexico city metropolitan area: Sensitivity of the wrf model to land cover changes. Frontiers in Earth Science, 6(February), 1–14. https://doi.org/10.3389/feart.2018.00006

Luo, L., Robock, A., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., Lohmann, D., Cosgrove, B. A., Wen, F., Sheffield, J., Duan, Q., Higgins, R. W., Pinker, R. T., & Tarpley, J. D. (2003). Validation of the North American Land Data Assimilation System (NLDAS) retrospective forcing over the southern Great Plains. Journal of Geophysical Research D: Atmospheres, 108(22), 1–10. https://doi.org/10.1029/2002jd003246

Mahrt, L., & Pan, H. (1984). A two-layer model of soil hydrology. Boundary-Layer Meteorology, 29(1), 1–20. https://doi.org/10.1007/BF00119116

Mallard, M. S., Spero, T. L., & Taylor, S. M. (2018). Examining WRF's sensitivity to contemporary land-use datasets across the contiguous united states using dynamical downscaling. Journal of Applied Meteorology and Climatology, 57(11), 2561–2583. https://doi.org/10.1175/JAMC-D-17-0328.1

Martin, M., Dickinson, R. E., & Yang, Z. L. (1999). Use of a coupled land surface general circulation model to examine the impacts of doubled stomatal resistance on the water resources of the American Southwest. Journal of Climate, 12(12), 3359–3375. https://doi.org/10.1175/1520-0442(1999)012<3359:UOACLS>2.0.CO;2

Matsui, T., Lakshmi, V., & Small, E. E. (2005). The effects of satellite-derived vegetation cover variability on simulated land-atmosphere interactions in the NAMS. Journal of Climate, 18(1), 21–40. https://doi.org/10.1175/JCLI3254.1

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., & Houston, T. G. (2012). An overview of the global historical climatology network-daily database. Journal of Atmospheric and Oceanic Technology, 29(7), 897–910. https://doi.org/10.1175/JTECH-D-11-00103.1

Mercer, A. (2020). Predictability of common atmospheric teleconnection indices using machine learning. Procedia Computer Science, 168(2019), 11–18. https://doi.org/10.1016/j.procs.2020.02.245

Morisette, J. T., Privette, J. L., & Justice, C. O. (2002). A framework for the validation of MODIS Land products. Remote Sensing of Environment, 83(1–2), 77–96. https://doi.org/10.1016/S0034-4257(02)00088-3

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G. R., Lotsch, A., Friedl, M., Morisette, J. T., Votava, P., Nemani, R. R., & Running, S. W. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. Remote Sensing of Environment, 83(1–2), 214–231. https://doi.org/10.1016/S0034-4257(02)00074-3

Naabil, E., Lamptey, B. L., Arnault, J., Kunstmann, H., & Olufayo, A. (2017). Water resources management using the WRF-Hydro modelling system: Case-study of the Tono dam in West Africa. Journal of Hydrology: Regional Studies, 12(May), 196–209. https://doi.org/10.1016/j.ejrh.2017.05.010

Ngan, F., Cohen, M., Luke, W., Ren, X., & Draxler, R. (2015). Meteorological modeling using the WRF-ARW Model for Grand Bay intensive studies of atmospheric mercury. Atmosphere, 6(3), 209–233. https://doi.org/10.3390/atmos6030209

Niu, G. Y., Yang, Z. L., Dickinson, R. E., & Gulden, L. E. (2005). A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models. Journal of Geophysical Research Atmospheres, 110(21), 1–15. https://doi.org/10.1029/2005JD006111

Niu, G., 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., Niu, G., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Longuevergne, L., Manning, K., … Xia, Y. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Evaluation over Global River Basins, Journal of Geophysical Research: Atmospheres, 116, 12. https://doi.org/10.1029/2010JD015140%0Ahttps://doi.org/10.1029/2010JD015139

Niyogi, D. S., Raman, S., & Alapaty, K. (1998). Comparison of four different stomatal resistance schemes using FIFE data. Part II: Analysis of terrestrial biospheric-atmospheric interactions. Journal of Applied Meteorology, 37(10 PART II), 1301–1320. https://doi.org/10.1175/1520-0450(1998)037<1301:cofdsr>2.0.co;2

Noilhan, J., and S. P. (1989). A simple parameterization of land surface processes for meteorological models. Monthly, 536–549. https://doi.org/https://doi.org/10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2

Novick, K. A., Biederman, J. A., Desai, A. R., Litvak, M. E., Moore, D. J. P., Scott, R. L., & Torn, M. S. (2018). The AmeriFlux network: A coalition of the willing. Agricultural and Forest Meteorology, 249(September 2017), 444–456. https://doi.org/10.1016/j.agrformet.2017.10.009

O, S., & Orth, R. (2021). Global soil moisture data derived through machine learning trained with in-situ measurements. Scientific Data, 8(1), 1–14. https://doi.org/10.1038/s41597-021-00964-1

Parajuli, K., Jones, S. B., Tarboton, D. G., Hipps, L. E., Zhao, L., Sadeghi, M., Rockhold, M. L., Torres-Rua, A., & Flerchinger, G. N. (2020). Stone content influence on land surface model simulation of soil moisture and evapotranspiration at reynolds creek watershed. Journal of Hydrometeorology, 21(8), 1889–1904. https://doi.org/10.1175/JHM-D-19-0075.1

Paul, S., Ghosh, S., Oglesby, R., Pathak, A., Chandrasekharan, A., & Ramsankaran, R. (2016). Weakening of Indian Summer Monsoon Rainfall due to Changes in Land Use Land Cover. Scientific Reports, 6(August), 1–10. https://doi.org/10.1038/srep32177

Pelissier, C., Frame, J., & Nearing, G. (2020). Combining Parametric Land Surface Models with Machine Learning. International Geoscience and Remote Sensing Symposium (IGARSS), 3668–3671. https://doi.org/10.1109/IGARSS39084.2020.9324607

Pérez, A., Santamaria, E. K., Operario, D., Tarkang, E. E., Zotor, F. B., Cardoso, S. R. de S. N., Autor, S. E. U., De, I., Dos, A., Vendas, O. D. E., Empresas, D. A. S., Atividades, P. O., Artigo, N., Gest, G. N. R. M. D. E., Para, D. E. F., Miranda, S. F. da R., Ferreira, F. A. A., Oliver, J., Dario, M., … Volk, J. E. (2017). No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析Title. BMC Public Health, 5(1), 1–8. https://ejournal.poltektegal.ac.id/index.php/siklus/article/view/298%0Ahttp://repositorio.unan.edu.ni/2986/1/5624.pdf%0Ahttp://dx.doi.org/10.1016/j.jana.2015.10.005%0Ahttp://www.biomedcentral.com/1471-2458/12/58%0Ahttp://ovidsp.ovid.com/ovidweb.cgi?T=JS&P

Pielke, R. A., Pitman, A., Niyogi, D., Mahmood, R., McAlpine, C., Hossain, F., Goldewijk, K. K., Nair, U., Betts, R., Fall, S., Reichstein, M., Kabat, P., & de Noblet, N. (2011). Land use/land cover changes and climate: Modeling analysis and observational evidence. Wiley Interdisciplinary Reviews: Climate Change, 2(6), 828–850. https://doi.org/10.1002/wcc.144

Pilotto, I. L., Rodríguez, D. A., Tomasella, J., Sampaio, G., & Chou, S. C. (2015). Comparisons of the Noah-MP land surface model simulations with measurements of forest and crop sites in Amazonia. Meteorology and Atmospheric Physics, 127(6), 711–723. https://doi.org/10.1007/s00703-015-0399-8

Qu, Weiqing, et al. (1998). Sensitivity of Latent Heat Flux from PILPS Land-Surface Schemes to Perturbations of Surface Air Temperature. JOURNAL OF THE ATMOSPHERIC SCIENCES, 55, 1909–1927.

Ran, L., Gilliam, R., Binkowski, F. S., Xiu, A., Pleim, J., & Band, L. (2015). Sensitivity of the Weather Research and Forecast/Community Multiscale Air Quality modeling system to MODIS LAI, FPAR, and albedo. Journal of Geophysical Research : Atmospheres, 120(16), 8491–8511. https://doi.org/https://doi.org/10.1002/2015JD024406

Running, S. W., Baldocchi, D. D., Turner, D. P., Gower, S. T., Bakwin, P. S., & Hibbard, K. A. (1999). A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. Remote Sensing of Environment, 70(1), 108–127. https://doi.org/10.1016/S0034-4257(99)00061-9

Schneider, N., Eugster, W., & Schichler, B. (2004). The Impact of Historical Land-Use Changes on the Near-Surface Atmospheric Conditions on the Swiss Plateau. Earth Interactions, 8(12), 1–27. https://doi.org/10.1175/1087-3562(2004)008<0001:tiohlc>2.0.co;2

Skamarock, W. C., Klemp, J. B., Dudhia, J. B., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., & Powers, J. G. (2021). A Description of the Advanced Research WRF Model Version 4.3. NCAR Technical Note, July, 1–165.

Strack, J. E., Pielke, R. A., & Adegoke, J. (2003). Sensitivity of model-generated daytime surface heat fluxes over snow to land-cover changes. Journal of Hydrometeorology, 4(1), 24–42. https://doi.org/10.1175/1525-7541(2003)004<0024:SOMGDS>2.0.CO;2

Sulla-Menashe, D., Gray, J. M., Abercrombie, S. P., & Friedl, M. A. (2019). Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product. Remote Sensing of Environment, 222(December 2018), 183–194. https://doi.org/10.1016/j.rse.2018.12.013

Talib, J., Taylor, C. M., Duan, A., & Turner, A. G. (2020). Intraseasonal soil moisture-atmosphere feedbacks on the Tibetan Plateau circulation. Journal of Climate, 1–51. https://doi.org/10.1175/jcli-d-20-0377.1

Temimi, M., Fonseca, R., Nelli, N., Weston, M., Thota, M., Valappil, V., Branch, O., Wizemann, H. D., Kondapalli, N. K., Wehbe, Y., Hosary, T. Al, Shalaby, A., Shamsi, N. Al, & Naqbi, H. Al. (2020). Assessing the impact of changes in land surface conditions on wrf predictions in arid regions. Journal of Hydrometeorology, 21(12), 2829–2853. https://doi.org/10.1175/JHM-D-20-0083.1

Thompson, G., Field, P. R., Rasmussen, R. M., & Hall, W. D. (2008). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. Monthly Weather Review, 136(12), 5095–5115. https://doi.org/10.1175/2008MWR2387.1

Trail, M., Tsimpidi, A. P., Liu, P., Tsigaridis, K., Hu, Y., Nenes, A., Stone, B., & Russell, A. G. (2013). Potential impact of land use change on future regional climate in the Southeastern U.S.: Reforestation and crop land conversion. Journal of Geophysical Research Atmospheres, 118(20), 11577–11588. https://doi.org/10.1002/2013JD020356

Tsvetsinskaya, E. A., Mearns, L. O., & Easterling, W. E. (2001). Investigating the effect of seasonal plant growth and development in three-dimensional atmospheric simulations. Part I: Simulation of surface fluxes over the growing season. Journal of Climate, 14(5), 692–709. https://doi.org/10.1175/1520-0442(2001)014<0692:ITEOSP>2.0.CO;2

Wang, K., & Liang, S. (2009). Estimation of daytime net radiation from shortwave radiation measurements and meteorological observations. Journal of Applied Meteorology and Climatology, 48(3), 634–643. https://doi.org/10.1175/2008JAMC1959.1

Wang, P., Niu, G. Y., Fang, Y. H., Wu, R. J., Yu, J. J., Yuan, G. F., Pozdniakov, S. P., & Scott, R. L. (2018). Implementing Dynamic Root Optimization in Noah-MP for Simulating Phreatophytic Root Water Uptake. In Water Resources Research (Vol. 54, Issue 3, pp. 1560–1575). https://doi.org/10.1002/2017WR021061

Workshop, D. O. E. B. (2015). An AmeriFlux network perspective on urban and managed systems 1. 1–11.

Wu, H., & Li, W. (2019). Downscaling Land Surface Temperatures Using a Random Forest Regression Model with Multitype Predictor Variables. IEEE Access, 7, 21904–21916. https://doi.org/10.1109/ACCESS.2019.2896241

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(3). https://doi.org/10.1029/2011JD016048

Xiaohui Zhong, & Utpal Dutta. (2015). Engaging Nash-Sutcliffe Efficiency and Model Efficiency Factor Indicators in Selecting and Validating Effective Light Rail System Operation and Maintenance Cost Models. Journal of Traffic and Transportation Engineering, 3(5), 255–265. https://doi.org/10.17265/2328-2142/2015.05.001

Xu, T., Guo, Z., Liu, S., He, X., Meng, Y., Xu, Z., Xia, Y., Xiao, J., Zhang, Y., Ma, Y., & Song, L. (2018). Evaluating Different Machine Learning Methods for Upscaling Evapotranspiration from Flux Towers to the Regional Scale. Journal of Geophysical Research: Atmospheres, 123(16), 8674–8690. https://doi.org/10.1029/2018JD028447

Yang, J., Ji, Z., Chen, D., Kang, S., Fu, C., Duan, K., & Shen, M. (2018). Improved Land Use and Leaf Area Index Enhances WRF-3DVAR Satellite Radiance Assimilation: A Case Study Focusing on Rainfall Simulation in the Shule River Basin during July 2013. Advances in Atmospheric Sciences, 35(6), 628–644. https://doi.org/10.1007/s00376-017-7120-4

Yang, R., & Friedl, M. A. (2003). Modeling the effects of three-dimensional vegetation structure on surface radiation and energy balance in boreal forests. Journal of Geophysical Research D: Atmospheres, 108(16), 1–11. https://doi.org/10.1029/2002jd003109

Yang, Y., Hu, X. M., Gao, S., & Wang, Y. (2019). Sensitivity of WRF simulations with the YSU PBL scheme to the lowest model level height for a sea fog event over the Yellow Sea. Atmospheric Research, 215(December 2017), 253–267. https://doi.org/10.1016/j.atmosres.2018.09.004

Yang, Z., & Dominguez, F. (2019). Investigating land surface effects on the moisture transport over South America with a moisture tagging model. Journal of Climate, 32(19), 6627–6644. https://doi.org/10.1175/JCLI-D-18-0700.1

Yang, Z. L., Dickinson, R. E., Henderson-Selles, A., & Pitman, A. J. (1995). Preliminary study of spin-up processes in land surface models with the first stage data of project for intercomparison of land surface parameterization schemes phase 1(a). Journal of Geophysical Research, 100(D8). https://doi.org/10.1029/95jd01076

Yoon, D., Cha, D. H., Lee, M. I., Min, K. H., Jun, S. Y., & Choi, Y. (2021). Comparison of regional climate model performances for different types of heat waves over south korea. Journal of Climate, 34(6), 2157–2174. https://doi.org/10.1175/JCLI-D-20-0422.1

Zhang, B., Zhang, Y., & Jiang, X. (2022). Feature selection for global tropospheric ozone prediction based on the BO-XGBoost-RFE algorithm. Scientific Reports, 12(1), 1–10. https://doi.org/10.1038/s41598-022-13498-2

Zhang, J. (2018). Evaluation of the quality of precipitation products: A case study using WRF and IMERG data over the central United States. Journal of Hydrometeorology, 19(12), 2007–2020. https://doi.org/10.1175/JHM-D-18-0153.1

Zhang, W., Xu, Z., & Guo, W. (2016). The impacts of land-use and land-cover change on tropospheric temperatures at global and regional scales. Earth Interactions, 20(7), 1–23. https://doi.org/10.1175/EI-D-15-0029.1

Zhao, D., & Wu, J. (2017). Inclusion of land use changes in long-term regional climate simulations over East Asia. Atmospheric Science Letters, 18(4), 187–192. https://doi.org/10.1002/asl.744

Zheng, D., van der Velde, R., Su, Z., Wang, X., Wen, J., Booij, M. J., Hoekstra, A. Y., & Chen, Y. (2015). Augmentations to the Noah model physics for application to the Yellow River source area. Part II: Turbulent heat fluxes and soil heat transport. Journal of Hydrometeorology, 16(6), 2677–2694. https://doi.org/10.1175/JHM-D-14-0199.1