UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

The Influence of Improved Land Surface and Soil Data on Mesoscale Model Predictions

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

By

CHRISTOPHER MICHAEL GODFREY
Norman, Oklahoma
2006

The Influence of Improved Land Surface and Soil Data on Mesoscale Model Predictions

A DISSERTATION APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY

David J. Stensrud (Co-Chair)

Lance M. Leslie (Co-Chair)

Kenneth C. Crawford

Peter J. Lamb

May Yuan

DEDICATION

To my parents, Albert and Barbara, who instilled in me a sense of awe and wonder for the amazing things our atmosphere can show us—Thank you for emphasizing the importance of putting forth my best effort in all of my endeavors.

To my little brother, Thomas—I look forward to seeing where your personality, talents, and brilliant mind will take you.

To my wife and best friend, Elaine, who has been and will always be there for me “in all kinds of weather”—I love you.

ACKNOWLEDGMENTS

Completing a doctoral degree certainly requires the efforts of more than one individual and I would like to thank several people for their contributions to this work. My advisor, Dr. David Stensrud, patiently led me through the long process of setting up my forecast model and guided me toward answers to difficult research questions. He was always willing to help, but he often forced me to discover my own answers, ultimately shaping me into a more knowledgeable and capable researcher. His thoughtful comments and suggestions tremendously improved this manuscript. Dr. Lance Leslie provided further guidance and suggestions and was a source of support and encouragement. Dr. May Yuan contributed the unique perspective of a geographer and pointed me toward the interdisciplinary connections between meteorology and geography. Dr. Peter Lamb directed me toward a view of the broad applications of my work and provided several insightful comments. Dr. Kenneth Crawford highlighted the policy implications of my research and requested that I quantify the impact of my research on model forecasts. In doing so, I have formed stronger arguments for continued funding of similar research.

This work would not have been possible without the Oklahoma Mesonet data provided by the Oklahoma Climatological Survey. Chris Fiebrich and Janet Martinez handled several quality control issues that I encountered and answered dozens of questions concerning the data. They and David Demko, Andrew Reader, and Rachael Sigler also quickly responded to some rather large data requests. Conversations with Dr. Jeff Basara and Brad Illston gave me a better understanding of the nature of soil moisture data. Dr. Jim Merchant and Roberto Bonifaz of the University of Nebraska–Lincoln provided the fractional vegetation coverage and leaf area index data. During my analysis of all these data, Dr. Michael Richman provided very helpful guidance with principal-component regression techniques and Dr. Kim Elmore loaned me several statistics books.

Additionally, I appreciate the helpful comments and suggestions during discussions with Dr. Ken Mitchell regarding soil errors in Eta analyses and with Dr. Alan Betts for his insight regarding surface flux observations. Dr. Michael Ek graciously provided a copy of a

manuscript that proved quite useful for deciphering the code within the Noah land surface model. Dr. Michael Baldwin furnished computer code that converts coordinates between two different map projections. Dr. Louis Wicker shared helpful advice on coding and high performance computing techniques and Dr. David Schultz answered questions regarding technical formatting. I also extend my gratitude to Dr. Fred Carr, who both funded the last semester of my research and gave me the rewarding opportunity to teach an undergraduate meteorology class.

I wholeheartedly express my appreciation to my wife, Elaine, for convincing me to continue my education when I originally had no intention of doing so. She edited my manuscripts, worked through sleepless nights, and provided thoughtful advice and unlimited encouragement. Her companionship is a blessing to me and she understands better than anyone the hurdles that must be overcome in this long process. Our strong marriage and friendship has benefited through our shared struggle to complete our education. Finally, I thank God, the rock upon which our marriage is founded and who answered innumerable prayers, for granting me with the ability to complete the long yet rewarding process of finishing my degree.

Contents

List of Tables

List of Figures

ABSTRACT

Proper partitioning of the surface heat fluxes that drive the evolution of the planetary boundary layer in numerical weather prediction models requires an accurate specification of the initial state of the land surface. The lack of observational data for characterizing these initial conditions is arguably the most difficult aspect in the evaluation of land surface models. Routine observations of fractional vegetation coverage and leaf area index (LAI) are not available at high resolution ($\sim$1 km), nor are observations of soil moisture and soil temperature. This gap in our observational capabilities seriously hampers the evaluation and improvement of land surface model parameterizations, since model errors likely relate to improper initial conditions as much as to inaccuracies in the model formulations. Two unique datasets help to overcome these difficulties. First, approximately 1-km resolution fractional vegetation coverage and LAI can be derived from biweekly maximum normalized difference vegetation index (NDVI) composites obtained from daily observations by the Advanced Very High Resolution Radiometer (AVHRR) onboard National Oceanic and Atmospheric Administration satellites. Second, the Oklahoma Mesonet supplies multiple soil temperature and soil moisture measurements at various soil depths each hour. Combined, these two unique datasets provide significantly improved initial conditions for a land surface model and allow an evaluation of the utility of the land surface model with much greater confidence and detail than previously.

Simulations from The Pennsylvania State University–National Center for Atmospheric Research fifth-generation Mesoscale Model (MM5) that both include and neglect these unique land surface observations help to evaluate the value of these two data sources to land surface model initializations. The dense network of surface observations afforded by the Oklahoma Mesonet, including surface flux data derived from special sensors available at some of the Mesonet sites, provides verification of the model results. The National Centers for Environmental Prediction (NCEP) operational Eta model, which provides initial conditions for MM5, exhibits strong biases in soil temperature and severe underestimation of soil moisture compared with observations during 2004 and 2005. Therefore, the

inclusion of soil temperature and soil moisture observations within MM5 simulations dramatically improves model performance. Including both soil and vegetation observations, however, tends to offset this improvement, indicating the necessity for adjustments to the land surface model physics.

A principal-component regression reveals simple relationships between latent heat flux and other available surface observations. Development of a new parameterization for evaporation from bare soil takes advantage of periods of very dry conditions observed across Oklahoma. Combining this with a new empirical canopy transpiration scheme within MM5 yields improved sensible and latent heat flux forecasts and better partitioning of the surface energy budget. Surface temperature and mixing ratio forecasts show improvement when compared with the dense network of observations from the Oklahoma Mesonet.

Chapter 1

Introduction

Numerical weather prediction models require an accurate representation of initial land surface conditions in order to partition properly the sensible and latent heat fluxes that drive the evolution of the planetary boundary layer. Several key components of the land surface that significantly affect surface heat and moisture fluxes include soil temperature and moisture, fractional vegetation coverage (σ_f), and green leaf area index (LAI). The lack of observational data for the accurate specification of these components in model initial conditions is arguably the most difficult aspect in the evaluation of land surface models. Soil temperature and moisture measurements are unavailable in most areas and routine remote-sensing observations of σ_f and LAI are not available at high resolution, i.e., with pixel widths on the order of 1 km and daily updates. This gap in our observational capabilities seriously hampers the evaluation and improvement of land surface model parameterizations, since improper initial conditions and inaccuracies in the model formulations very likely produce comparable model errors.

Models accomplish the exchange of energy between the land surface and the atmosphere through land surface parameterizations (e.g., Bhumralkar 1975; Blackadar 1976; Deardorff 1978; McCumber and Pielke 1981; Pan and Mahrt 1987; Noilhan and Planton 1989), which characterize the state of the land surface and forecast the evolution of the lowest layer of the model atmosphere. The surface energy balance relies strongly upon the soil

and near-surface conditions, and plays a critical role in determining the prognostic variables in land surface models. Surface energy fluxes depend heavily upon soil temperature and soil moisture conditions, as well as vegetation coverage, atmospheric conditions, and the physical properties of the soil. Soil moisture is an important component describing the land surface and provides a key link between the atmosphere and the water and energy balances at the surface of the earth (Wei 1995; Robock et al. 2000; Leese et al. 2001; Koster et al. 2004a). It influences the available water for plant transpiration, and plays a role in the mass balance for many forecast models. Soil thermal conductivity estimates, which facilitate the proper heat transfer within the soil, also strongly depend upon soil moisture specifications. For calculations of soil heat transfer, the most sophisticated land surface parameterizations require not only near-surface soil temperatures, but also temperature profiles within the soil (e.g., Viterbo and Beljaars 1995; Chen and Dudhia 2001). In addition, vegetation coverage and density provide critical information on the partitioning of total evaporation between bare soil and canopy transpiration (Chen and Dudhia 2001). Together, soil temperature, soil moisture, and vegetation affect forecasts of temperature, mixing ratio, cloud cover, and precipitation by working in concert to directly influence sensible, latent, and ground heat fluxes.

The mesoscale model employed for this study implements a monthly climatology for fractional vegetation coverage and a constant leaf area index. Studies have shown that such coarse resolution data based solely on climatology are insufficient to capture the detailed surface characteristics necessary to properly initialize a land surface parameterization (e.g., Chang and Wetzel 1991; Crawford et al. 2001; Santanello and Carlson 2001; Kurkowski et al. 2003). By using climatological values for land surface characteristics, the model does not account for short-term or annual variability in vegetation coverage and condition due to daily variations in rainfall, seasonal droughts, flooding, forest fires, irrigation, deforestation, desertification, crop harvesting, land usage, hail or tornado damage, and temporal variations in the growth and senescence of green vegetation. Large-scale atmospheric oscillations may also play a role in the interannual variability of vegetation (e.g., Jin and

Zhang 2002; Matsui et al. 2005). Modeling studies implementing near real-time land surface characteristics from satellite observations have shown great promise for improving forecasts (e.g., Oleson and Bonan 2000; Zeng et al. 2000; Crawford et al. 2001; Kurkowski et al. 2003).

Taking advantage of a unique set of soil and vegetation observations to improve the initial specification of the land surface should lead to more accurate model forecasts of air temperature and moisture, which directly affect planetary boundary layer processes and convective development. Better short-term forecasts of these near-surface variables benefit a wide range of personal and economic activities, including aviation, electrical energy generation and trading, agricultural pest control, construction, transportation, and the weather risk market (Dutton 2002). Further societal applications to improved short-term and seasonal model forecasts include enhanced accuracy in fire weather, convective initiation, fog, and air quality predictions and positive impacts for agricultural production, land management practices, and the energy industry.

This study represents an effort to improve the specification of initial conditions and to ultimately facilitate improved model forecasts. Chapter 2 surveys the available literature concerning soil temperature, soil moisture, and vegetation and the role of each in numerical weather prediction models. Chapter 3 describes sources of soil and vegetation observations, while chapter 4 assesses the current state of initial land surface conditions provided to many numerical weather prediction models. This assessment provides a starting point for determining how to improve the specification of initial land surface variables in land surface models. Modifications to a mesoscale model discussed in chapter 5 allow the model to assimilate soil temperature, soil moisture, σ_f, and LAI observations for several case studies. In addition to supplying initial soil conditions, a dense network of surface observations over the primary study area provides a means to verify forecasts. Results show that despite improved land surface conditions, inaccuracies still exist in the model formulations. This result provides a springboard for assessing parameterization errors within the model. Chapter 6 describes the surface energy fluxes in a specific land surface model and

introduces a new empirical latent heat flux parameterization. In a novel approach to determining latent heat flux, the new parameterization derives from surface observations rather than from theoretical formulations. Chapter 7 shows the results of forecasts that implement the new parameterization and chapter 8 wraps up the discussion by presenting ways to further improve a land surface model.

Chapter 2

Background

2.1 Soil moisture

Several studies have demonstrated sensitivities of forecasts of near-surface variables to soil water content. An inspection of the relationship between soil moisture variations and surface turbulent energy fluxes for a variety of vegetation types in different land surface modeling schemes reveals that energy fluxes display more sensitivity for dry soils than for wet soils and that sparsely vegetated areas require the most accurate soil moisture information (Dirmeyer et al. 2000). Changes in soil moisture modify the balance between latent and sensible heat fluxes and can influence surface temperatures or affect turbulent transfer in the boundary layer (McCumber and Pielke 1981). Soil moisture inhomogeneities may also aide in dryline development (Ziegler et al. 1995). The importance of soil moisture is illustrated by Pan and Mahrt (1987), who couple a one-dimensional model of the planetary boundary layer (Troen and Mahrt 1986) with a two-layer soil hydrology model (Mahrt and Pan 1984) and find that surface evaporation can drive boundary-layer development.

Root-zone soil moisture impacts planetary boundary layer processes and the development of deep convection by moderating sensible and latent heat fluxes and influencing boundary layer moisture (Clark and Arritt 1995; Basara and Crawford 2002). Yan and Anthes (1988) investigate the effect of soil moisture variations on precipitation patterns

by simulating adjacent strips of moist and dry land. They find that for sufficiently wide horizontal strips under convectively unstable conditions, the inhomogeneities in surface moisture lead to gradients of ground temperature that eventually help produce sea-breeze circulations and an increase in convective rainfall. This result compliments the observations of Pielke and Zeng (1989), who show increases in available buoyant energy when irrigated land lies adjacent to natural grassland, compared with natural grassland alone. Soil moisture further affects boundary-layer cloud development by increasing cloud cover for both moist and dry soils, depending on the strength of the stability above the boundary layer (Ek and Holtslag 2004).

The influence of soil moisture on persistent seasonal climate anomalies was first suggested by Namias (1952, 1959). More extensive numerical and observational studies of soil moisture reveal that soil moisture anomalies influence regional atmospheric conditions over time scales of two to three months (Liu et al. 1993; Vinnikov et al. 1996), with variations in temporal scales of soil moisture attributable to the seasonal cycle of potential evaporation (Entin et al. 2000). After simulating soil moisture anomalies, there is evidence that soil moisture affects model forecasts of precipitation, atmospheric moisture, and temperature for several weeks (Walker and Rowntree 1977; Rowntree and Bolton 1983). Modeling studies of soil temperature and moisture conditions show that differing soil moisture initializations influence monthly or seasonal temperatures and precipitation patterns (Rind 1982; Betts et al. 1996) and that these initial conditions again possess a persistence time scale of months to seasons (Yeh et al. 1984; Walsh et al. 1985; Vinnikov and Yeserkepova 1991; Gao et al. 1996; Liu and Avissar 1999a,b). Monthly forecasts also show sensitivity to initial soil moisture conditions, displaying increased skill for precipitation and air temperature forecasts with more realistic land surface initializations (Koster et al. 2004b). Other studies report that soil moisture anomalies also affect extreme precipitation forecasts on monthly time scales (Beljaars et al. 1996; Viterbo and Betts 1999). In seasonal predictions, Fennessy and Shukla (1999) investigate the role of initial soil moisture using ensembles of global climate model simulations and find that increases in initial soil wetness lead to

increased seasonal evaporation, decreased seasonal mean surface air temperatures, and generally increased seasonal mean precipitation in many regions. Koster et al. (2004a) identify specific regions, including the Great Plains, where soil moisture anomalies are particularly important for seasonal rainfall prediction. Other authors assert that the seasonal evolution of the atmosphere in a regional atmospheric model is dependent upon initial soil moisture and landscape specification (Pielke et al. 1999). Thus, when compared with soil temperature, soil moisture clearly has more interannual variability and more strongly influences forecasts (Liu and Avissar 1999a,b; Rodell et al. 2005).

2.2 Soil temperature

While soil moisture appears to be the most important factor for land-surface initializations (Gannon 1978; McCumber and Pielke 1981; C. Smith et al. 1994), one should not underestimate the role of soil temperature in the evolution of the lower atmosphere, especially for short-range forecasts. Without accurate soil temperature information, a planetary boundary layer scheme may incorrectly distribute heat near the surface. Substrate temperatures that are too cold or warm lead to a surface cooling or warming bias (Dudhia 1996). Longwave radiation loss is a function of soil temperature and directly affects the surface radiation budget. Ground heat flux also is a function of soil temperature (Brotzge and Crawford 2003), and affects the sensible heat flux, boundary layer growth and decay, turbulence, and air temperature. Additionally, there are successful attempts at retrieving soil moisture from more easily obtained soil temperature observations (e.g., Xu and Zhou 2003).

Results from a simulation using The Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) fifth-generation Mesoscale Model (MM5) version 3.6 (Dudhia 1993; Grell et al. 1995; Dudhia 2003) illustrate the importance of initial soil temperature conditions. A test compares a 48-hour control forecast over Oklahoma initialized with a 1200 UTC National Centers for Environmental Prediction (NCEP) Eta

analysis on 3 May 2004 with a second forecast with the same initial conditions except that the soil temperature at each grid point in the 0–10 cm layer is perturbed by a uniform random number (Bratley et al. 1987, chapter 6) bounded by ±2°C. For consistency with the lower soil layers, the soil temperature in the 10–40 cm layer is perturbed by half the magnitude of the perturbation in the top soil layer. The root-mean squared difference between the perturbed forecast compared with the control forecast (Wilks 2006, p. 308) shows that the magnitude of the difference between the perturbed soil temperatures and those in the control forecast decreases over the length of the forecast period (Fig. 2.1a). Because of external influences on the top soil layer, perturbed soil temperatures in the 0–10 cm layer return to control forecast soil temperatures more quickly over time than temperatures in the 10–40 cm layer. An anomaly correlation, given by

$$AC = \frac{\sum_{n=1}^{m} \left[(T_p - T_c)\Big|_{t=0} (T_p - T_c)\Big|_{t=h} \right]}{\left[\sum_{n=1}^{m} (T_p - T_c)^2 \Big|_{t=0} \sum_{n=1}^{m} (T_p - T_c)^2 \Big|_{t=h} \right]^{1/2}}, \tag{2.1}$$

where T_p is the perturbed soil temperature and T_c is the control soil temperature at grid point n, summed over m grid points for each forecast hour t over h forecast hours, provides a measure of association between the control and perturbed forecast fields (Wilks 2006, p. 311). Here, the sign and magnitude of the perturbation at forecast time h is compared against the value of the initial perturbation at each grid point. The anomaly correlation at each forecast hour for both soil levels reveals that the sign of each perturbation strongly persists throughout the forecast period (Fig. 2.1b). Perturbations persist because horizontal diffusion between adjacent gridded soil temperature values is negligible compared with horizontal diffusion in the atmosphere. This test indicates that inaccurate soil temperatures provided as initial conditions from gridded analysis fields may adversely affect the resulting short-term model forecasts of near-surface variables.

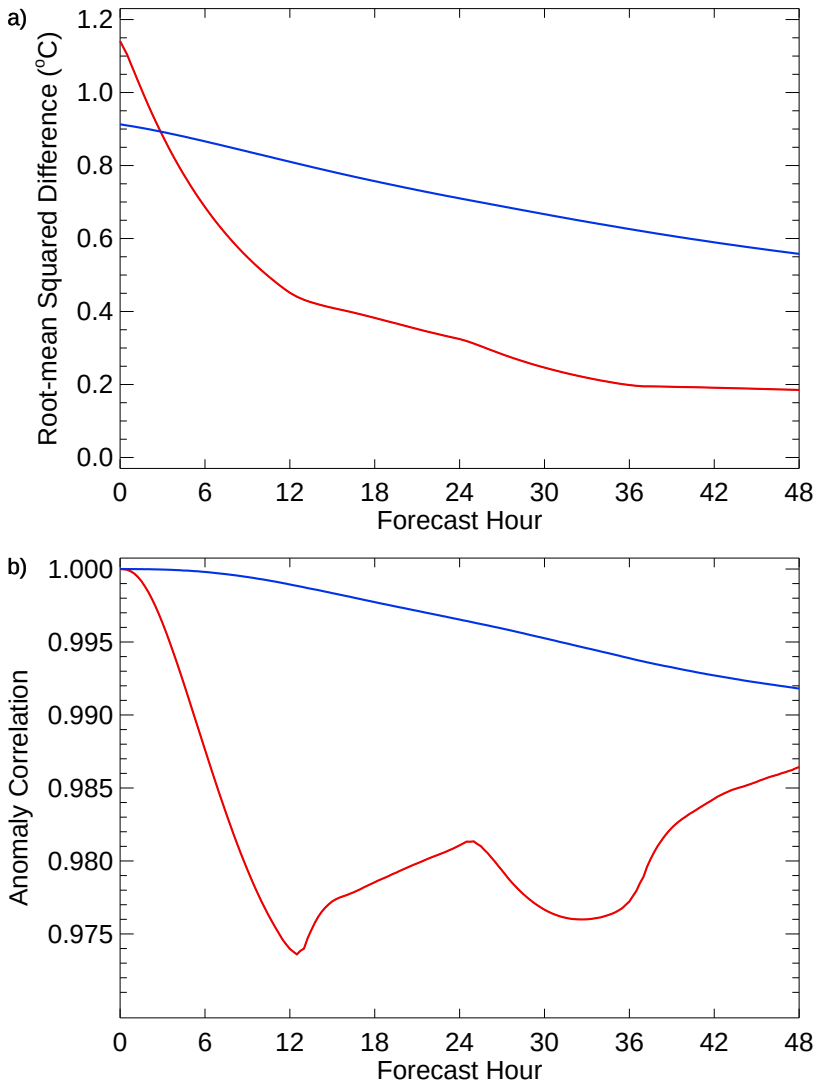

FIG. 2.1: Comparative measures between a control model simulation and a model simulation with randomly perturbed soil temperatures showing a) root-mean squared difference and b) anomaly correlation for the 0–10 cm layer (red) and the 10–40 cm layer (blue).

2.3 Vegetation

The concept that areas of vegetation may influence precipitation through increased moisture availability has circulated for a number of years. Anthes (1984) notes that with the rapid colonization and deforestation of islands in the Caribbean Sea in the 18th century, legislators noticed detrimental environmental effects and enacted an ordinance to create a protected forest reserve on the island of St. Vincent "for the purpose of attracting the clouds and rain" (Beard 1949, p. 30). More recently, scientific studies of the effect of vegetation

density and coverage indicate the absolute necessity for including vegetation parameterizations within numerical weather prediction models (e.g., Pielke et al. 1991; McPherson et al. 2004).

An assessment of the spatial variability of observed latent heat fluxes reveals a strong relationship between evaporation and the distribution of soil moisture and vegetation (Chen and Brutsaert 1995). Lower albedo and decreased infrared emission over vegetated surfaces increases the net radiation absorbed at the surface. This energy feeds evaporation, transpiration, and sensible heating and results in a moister lower atmosphere with higher equivalent potential temperature than over bare soil (Anthes 1984). The effect of vegetation and soil moisture on soil heat capacity and thermal inertia can produce pronounced soil temperature, skin temperature, and air temperature gradients in response to observed vegetation density gradients (e.g., E. Smith et al. 1994). Compared with areas with healthy vegetation, high temperatures and drier soils in areas with struggling vegetation lead to high sensible heat fluxes and suppressed latent heat fluxes. These gradients and mesoscale heterogeneities may induce perturbation boundary-layer circulations as the vegetation increases available low-level moist static energy and water vapor (Anthes 1984). Strong mesoscale latent heat fluxes imply the presence of well-developed mesoscale circulations that spawn intense cloud activity and transport water vapor into the atmosphere (Chen and Avissar 1994b).

Several modeling studies have demonstrated that inhomogeneities in spatial landscape variability may induce mesoscale circulations (e.g., Ookouchi et al. 1984; Avissar and Pielke 1989; Pielke et al. 1991; Chen and Avissar 1994a; Segele et al. 2005). Development of these circulations requires a discontinuity of soil moisture, vegetation type, vegetation coverage, or land usage, such as dense, irrigated, extended crop areas adjacent to bare soil areas (Segal et al. 1988) or zones of devegetation produced by hail swaths (Segele et al. 2005). Vegetation breezes operate on relatively small scales of perhaps 30 km (Sellers et al. 1986), though Anthes (1984) finds enhanced convective precipitation under certain atmospheric conditions over semiarid land by modeling strips of vegetation on the order

of 50–100 km wide. Using a model with a land surface parameterization calibrated with observations, Pinty et al. (1989) show that a significant mesoscale circulation forms over moist soil, but dry soil cuts off the vegetation breeze, even with dense vegetation cover.

Observations support the assertion that vegetation breezes only appear under relatively weak synoptic forcing and light winds (e.g., Segal and Arritt 1992). Doran et al. (1995) observe a mesoscale circulation due to thermal contrasts between semiarid grassland and irrigated farmland, but only under relatively calm conditions. One study documents a temperature reduction of 10°C over irrigated cropland compared with adjacent dry soil, with an associated rise in low-level moisture over the crops that penetrates well into the boundary layer (Segal et al. 1989). This same study, however, did not observe any well-defined mesoscale circulations, though terrain effects and strong synoptic flow may have overwhelmed the circulation.

Vegetation also affects the diurnal range of temperatures. A documented seasonal variation in the annual march of the diurnal temperature range, which displays a minimum in the summer and winter months, results from the cooling effect of high rates of transpiration from vegetation during the summer and from changes in insolation and cloudiness during the winter (Durre and Wallace 2001). This finding corroborates a modeling study that first simulates forest cover, then pasture. The change from forest to pasture reduces evapotranspiration by 30%, increases the surface temperature by 4°C, and increases the summertime diurnal temperature range (Silberstein et al. 1999).

Low-level horizontal temperature gradients such as those caused by surface inhomogeneities in soil moisture or vegetation may generate enough lift to release potential instability (e.g., Sun and Ogura 1979). Numerical simulations have shown that vegetation initiates and enhances convection both by shading the soil, which reduces ground heat flux and increases the energy available for sensible and latent heating, and by extracting moisture from the soil (Sud et al. 1993; Clark and Arritt 1995). The increased roughness length associated with vegetation compared with bare soil reduces the low-level wind speed and increases the strength of turbulent eddies, which may also lead to convective initiation (Anthes 1984).

The preferred location for convective initiation by mesoscale circulations is at the interface between vegetation and bare soil (Anthes 1984; Mahfouf et al. 1987). In one study, the greatest rainfall occurs over moist, vegetated surfaces and the authors note that vegetation may moderate the sensitivity of models to initial soil moisture conditions (Clark and Arritt 1995). Models generally perform best at forecasting convection when given realistic parameterizations of both soil moisture and vegetation, particularly under weak synoptic flow conditions (Garrett 1982; Chang and Wetzel 1991; Xue et al. 1996). To improve the soil moisture specification and to prevent excessive drying in a multilayer soil hydrological model, some authors recommend including the vegetation-dependent process of hydraulic lift, a soil water redistribution process that releases water from root systems into dry soil at night (Ren et al. 2004).

Observational studies report measurements of the effect of vegetation on the boundary layer. Fiebrich and Crawford (2001) trace a case of anomalously cool air temperatures at a single Oklahoma Mesonet site to its proximity to an irrigated cotton field. Similarly, growing winter wheat can develop dewpoint anomalies, while distinct warm anomalies appear over areas of harvested wheat (Haugland and Crawford 2002; McPherson et al. 2004). Under weak synoptic forcing and when the atmosphere is relatively dry, these warm anomalies over harvested wheat adjacent to growing vegetation may induce cloud formation, while areas with high latent heat fluxes such as heavy tree cover and lakes tend to suppress clouds (Rabin et al. 1990).

Chapter 3

Observations

An accurate specification of initial conditions for model forecasts relies on observations of soil temperature, soil moisture, surface fluxes, and standard variables available from the Oklahoma Mesonet and vegetation conditions based on satellite observations. The availability of this unique set of observations across Oklahoma makes this an ideal region for studying potential improvements to model forecasts.

3.1 Oklahoma Mesonet

The Oklahoma Mesonet is an integrated network of automated surface observing stations, with at least one site in each of Oklahoma's seventy-seven counties (Fig. 3.1). Measurements of atmospheric variables occur every five minutes at each of the 116 sites.[1] All Mesonet sites report soil temperature at one or more depths every 15 minutes. Infrared temperature sensors (Fiebrich et al. 2003) record the skin temperature at 86 sites. Over 100 sites also record soil moisture every thirty minutes at levels of 5, 25, 60, and 75 cm below the surface. Approximately 75 sites measure ground heat flux and total net radiation every five minutes. A special suite of instruments augments the standard instrumentation at

[1] A newly commissioned site at Fittstown, Oklahoma briefly brought the total number of sites to 117 on 12 May 2005, but the Oklahoma Climatological Survey decommissioned the Bee, Oklahoma site on 13 July 2005.

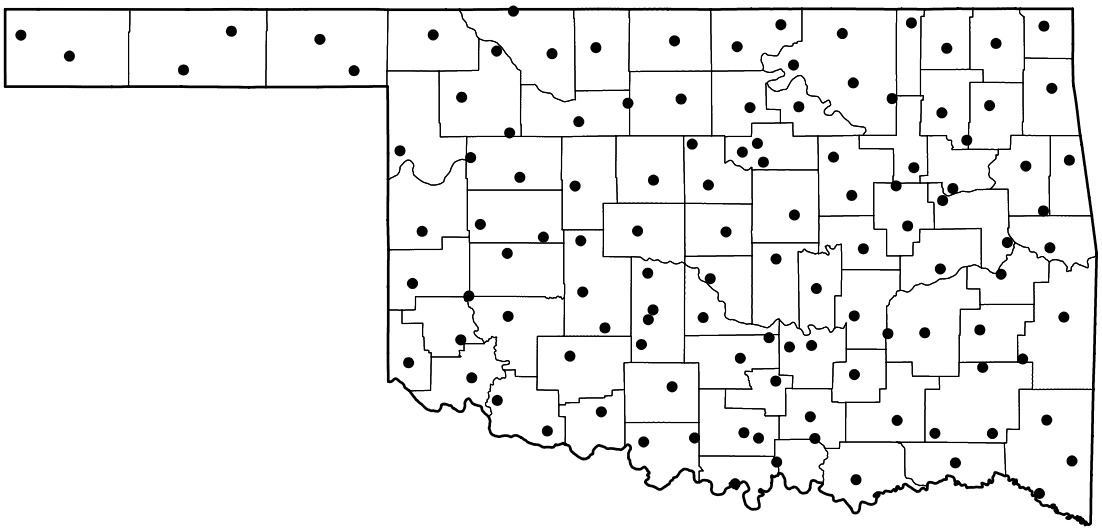

FIG. 3.1: Site locations for each of the 117 Mesonet sites providing data between 1 March 2004 and 1 October 2005.

ten sites, measuring sensible heat flux and the four components of net radiation every five minutes. All data fall subject to rigorous quality assurance procedures in order to produce reliable research-quality data (Shafer et al. 2000). A complete description of the Oklahoma Mesonet, including sensor specifications, appears in Brock et al. (1995), while Basara and Crawford (2000) describe the soil moisture instrumentation.

3.1.1 Soil moisture measurements

Matric potential is a pressure potential arising from the interaction of water with the colloidal matrix of soil particles. Water molecules undergo attractive forces due to capillary suction and surface adsorption (Marshall et al. 1996, p. 34). Plants must overcome this attractive force within the soil in order to maintain water transport from roots to leaves. Values of matric potential are negative, with larger absolute values of matric potential indicating drier soil. Depending on whether a unit quantity of water has volume, mass, or weight units, expressions of matric potential may appear with a variety of units attached. Potential, or more generally energy, per unit volume is in J m^{-3} or equivalent pressure units. Potential per unit mass is in J kg^{-1} and potential per unit weight is in meters (Marshall et

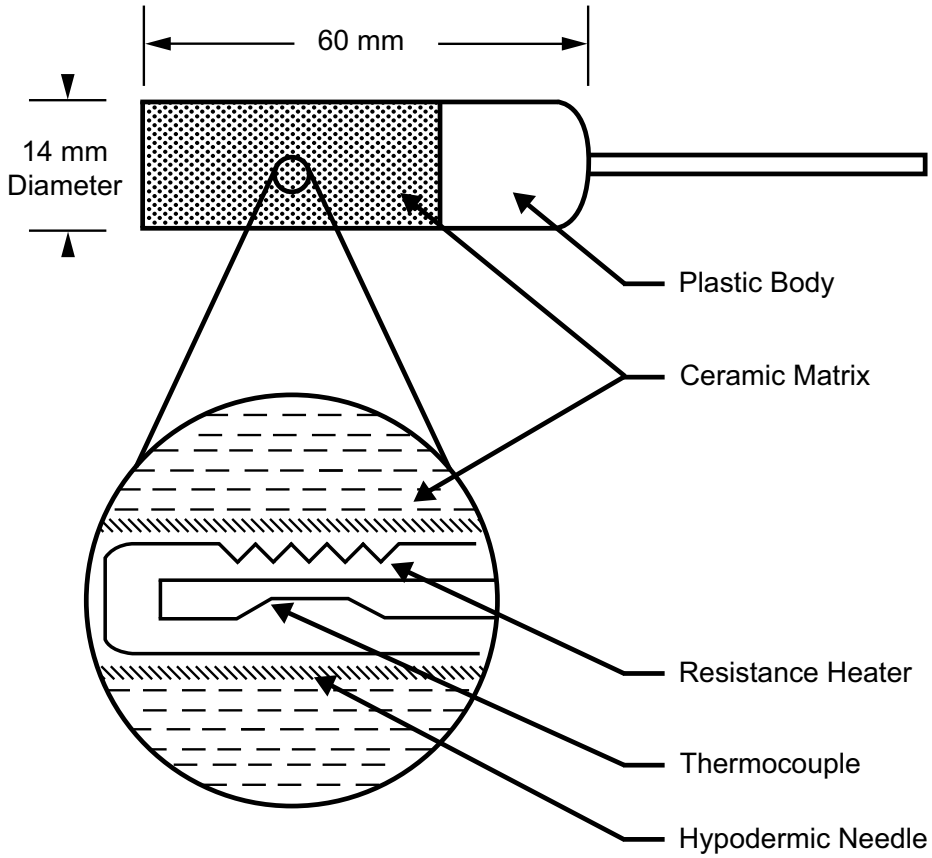

FIG. 3.2: The Campbell Scientific, Inc. 229-L heat dissipation matric potential sensor. The inset represents a cross section through the ceramic matrix (after Basara and Crawford 2000).

al. 1996, p. 37).

Scientists have recognized for some time that the rate of heat dissipation in soil directly relates to the matric potential (e.g., Shaw and Baver 1939). The Campbell Scientific, Inc. (CSI) 229-L heat dissipation matric potential sensor (Fig. 3.2) installed at Oklahoma Mesonet sites takes advantage of this principle. Encased within a porous ceramic matrix resides a hypodermic needle that houses a resistor as a heating element and a thermocouple as a temperature sensor. The instrument measures an initial soil temperature with the thermocouple, applies a small voltage to the resistance heater for several seconds, and again measures the resulting soil temperature. The difference ΔT between the initial and final temperatures depends upon the amount of water in the surrounding soil. To remove variability between sensors across the Mesonet, ΔT relates to a normalized reference temperature for all sensors

ΔT_{ref} according to

$$\Delta T_{\text{ref}} = m\Delta T + b, \tag{3.1}$$

where m and b are sensor-specific calibration coefficients (Basara and Crawford 2000). Data from vacuum, pressure chamber, and tensiometer measurements of soils (Reece 1996) yield an empirical relationship between the normalized reference temperature and matric potential given by

$$\psi = -c\exp(a\Delta T_{\text{ref}}), \tag{3.2}$$

where ψ is the matric potential (kPa) and a and c are calibration constants equal to 1.788 $^{\circ}\text{C}^{-1}$ and 0.717 kPa, respectively. Compared with both the original formulation that appears in Reece (1996) and the modified version from Basara and Crawford (2000), this relationship is simpler and more accurate (B. G. Illston 2005, personal communication).

While matric potential provides an important measure of soil moisture for modeling water movement within the soil and from the soil to plants, volumetric water content provides forecast models with important information regarding the volume of water present within the soil as a fraction of the total soil volume. Land surface models rely on measures of volumetric water content to determine soil thermal conductivity and model hydrology. A soil water retention curve describes the relationship between volumetric water content and matric potential for a given soil type (e.g., Clapp and Hornberger 1978; Rawls et al. 1982). Due to the large number of sensors at different depths and different observing sites, the Oklahoma Climatological Survey (OCS) decided not to determine a soil water retention curve for each sensor at each site. Instead, an empirical relationship based on detailed soil characteristics and bulk density measurements at each observing site provides coefficients α (kPa^{-1}) and n characteristic to each soil texture (Arya and Paris 1981). This same methodology also provides estimates of the residual water content, Θ_r, and the saturated water content, Θ_s, both measured in units of $\text{m}^3_{\text{water}}\ \text{m}^{-3}_{\text{soil}}$. The residual water content represents the volumetric water content of very dry soil and the saturated water content, or porosity, represents the maximum amount of water that a given soil volume can hold. These quanti-

ties provide estimates of the soil volumetric water content from calculated values of matric potential using

$$\Theta = \Theta_r + \frac{\Theta_s - \Theta_r}{\left(1 + \left[\alpha\left(-\psi/100\right)\right]^n\right)^{1-1/n}}\,, \tag{3.3}$$

where Θ is the volumetric water content and Θ_r is specifically defined as the water content for which the gradient $\partial\Theta/\partial\psi$ becomes zero (van Genuchten 1980).

3.1.2 Soil temperature measurements

All Mesonet sites employ a Fenwal thermistor to measure soil temperature at 30-second intervals at a depth of 10 cm under both bare soil and native vegetation. Recorded soil temperature observations represent an average of these measurements over 15 minutes. Approximately half of the Mesonet sites measure soil temperature at a depth of 5 cm under both bare soil and native vegetation and at a depth of 30 cm under native vegetation. Brock et al. (1995) note that the shadow of the solar panel from the Mesonet tower occasionally affects soil temperature readings at the 5-cm depth. In addition, vegetation cover may moderate the response of soil temperature sensors (Fiebrich and Crawford 2001). Allowing for the inherent difficulty with consistently maintaining a completely vegetation-free area over thermistors buried under bare soil, and because numerical weather prediction models typically account for vegetation, all soil temperatures in this study represent those measured under native vegetation.

3.1.3 Surface energy flux measurements

Measurements of surface energy fluxes rely on instrumentation installed as part of the Oklahoma Atmospheric Surface-layer Instrumentation System (OASIS) project (Brotzge et al. 1999). Standard OASIS sites, of which there were approximately 75 during the course of this study, observe total net radiation and ground heat flux. An NR-Lite domeless net radiometer manufactured by Kipp & Zonen measures the net radiation at each of these

sites. While a four-component net radiometer likely provides measurements with greater accuracy, cost and maintenance requirements make the NR-Lite the most suitable net radiometer for wide distribution across the Mesonet (Brotzge and Duchon 2000). Values of net radiation from the NR-Lite radiometer have a correction applied based on a sensor-specific factory calibration coefficient and the observed 2-meter wind speed. Corrections for wind speed only apply if the observed net radiation is greater than 20 W m^{-2} and the 2-meter wind speed exceeds 5 m s^{-1} (Brotzge 2000). The ground heat flux is the sum of the conductive ground heat flux and the storage ground heat flux. Each site directly measures the conductive ground heat flux using the arithmetic mean of two Radiation and Energy Balance Systems, Inc. (REBS) HFT3.1 heat flux plates installed at a depth of 5 cm. Estimates of the storage ground heat flux derive from measurements of the volumetric water content at 5-cm, an average volume fraction of minerals and organic matter, and soil temperature from two REBS integrating platinum resistance temperature detectors installed diagonally within the 0–5 cm soil layer (Brotzge and Crawford 2003).

A suite of OASIS instruments measures sensible heat flux and the four components of net radiation at ten super sites.[2] A Kipp & Zonen CNR1 four-component net radiometer measures the incoming and outgoing components of solar and far infrared radiation. The CNR1 measures net radiation more stably and accurately than the NR-Lite (Brotzge 2004). Sensible heat fluxes derive from CSAT3 sonic anemometer measurements of the covariance of the z-axis wind speed with sonic temperature, computed by an eddy covariance processing algorithm from CSI, and multiplied by the specific heat at constant pressure and the observed air density. Due to the nature of sonic anemometer observations, these sensible heat fluxes are unavailable or invalid during periods of precipitation. A gradient profile technique (e.g., Brotzge and Crawford 2000) can produce estimates of sensible heat flux during precipitation, though comparisons of sensible heat flux between numerical models and observations occur in this study during only synoptically quiescent

[2]The Oklahoma Climatological Survey discontinued OASIS super site measurements at the Burneyville site on 21 June 2005, bringing the total number of super sites to nine.

conditions. Many of the super sites retain the capability of directly estimating latent heat fluxes from measurements, corrected for the density of atmospheric oxygen, obtained from a CSI KH-20 Krypton hygrometer (Brotzge and Crawford 2003). However, OCS does not distribute latent heat flux data because maintenance ceased on this instrumentation prior to the beginning of this study. The residual of the surface energy balance instead provides a proxy for latent heat flux estimates and may actually provide a more reasonable estimate of the latent heat flux than direct estimates from an eddy covariance system (Brotzge 2004). Brotzge (2000) provides a thorough discussion of OASIS instrumentation, measurements, and calculations of surface energy fluxes.

3.2 Satellite-derived vegetation indexes

The Advanced Very High Resolution Radiometer (AVHRR) subsystem resides aboard each of the six currently active NOAA Polar Orbiting Environmental Satellites (POES) referred to as NOAA-12, -14, -15, -16, -17, and -18. The Office of Satellite Operations, an organizational component of NOAA's National Environmental Satellite, Data, and Information Service (NESDIS), manages these operational environmental satellites. Each satellite flies at an altitude of 833 $\pm$ 19 km in a sun-synchronous orbit with a period of 101.6 $\pm$ 0.5 minutes. The local solar time of the satellite's passage is constant for any latitude. Thus, multiple images of the same location show the same sun angle, excepting changes in illumination over long periods due to orbital drift (Kidwell 1998). The AVHRR subsystem measures six spectral channels with a field of view of 1.3 milliradians by 1.3 milliradians, giving a ground resolution of 1.09 km. Complete details of the newest AVHRR/3 system appear in Goodrum et al. (2001).

Two primary satellites work in tandem to sample the entire planet daily, with several backup satellites available should a primary satellite fail. Other earth-observing satellites carry AVHRR sensors and possess the capability of obtaining higher resolution measurements, e.g., the 30-m resolution land remote sensing satellite system (Landsat) and the

10-m resolution System Probatoire d'Observation de la Terre (SPOT). However, the cost of acquiring the higher resolution data may be prohibitive for large-scale applications, both financially and in terms of storage space. The readily available AVHRR data require only nominal cost. In addition, the revisit period of these high-resolution satellites is 16 days for Landsat and roughly 26 days for SPOT. Since cloudless observations may not be possible for a given location, more than a month may pass between successive observations. The Earth Observing System (EOS) Moderate-Resolution Imaging Spectroradiometer (MODIS) provides daily coverage in 36 spectral bands at 250–100-m resolution (Justice et al. 1998). However, this satellite is still experimental and has an uncertain future.

Several types of vegetation indexes derive from AVHRR data (e.g., Viña et al. 2004), but NDVI is the most popular vegetation index for estimating σ_f and LAI. NDVI is a function of the reflectance of different wavelengths of the solar spectrum:

$$\text{NDVI} = \frac{\rho_2 - \rho_1}{\rho_2 + \rho_1}, \tag{3.4}$$

where ρ_1 and ρ_2 are reflectance measurements by a silicon detector in AVHRR channels 1 (0.58–0.68 μm) and 2 (0.725–1.00 μm), respectively (Goodrum et al. 2001; Gutman and Ignatov 1998). The high reflectance of near-infrared light (ρ_2) and the low reflectance of visible red light (ρ_1) on vegetation produce larger values of NDVI (Walter-Shea et al. 1992). Conversely, the low reflectance of near-infrared light and high reflectance of red light from clouds, snow, water, and bare soil produce low (typically negative) values of NDVI (Yin and Williams 1997). Fortunately, NDVI partially compensates for changes in illumination, surface slope, and viewing angle, all of which strongly affect observed radiances (Gutman et al. 1995). Composite maximum NDVI images over a period of weeks effectively eliminate low NDVI values due to cloud contamination and provide for appropriate parameterizations of vegetation coverage (e.g., Crawford et al. 2001; Kurkowski et al. 2003). The current study employs a 15- or 16-day observation window for computing maximum NDVI composites.

The characterization of vegetation in numerical models requires two parameters. The model grid cell fraction where a photosynthetically active green canopy intercepts downward solar radiation at midday defines σ_f (Chen et al. 1996). The vegetation fraction acts as a weighting coefficient between direct evaporation from the top soil layer, evaporation of precipitation intercepted by the canopy layer, and transpiration from the vegetation. Depending on the season and the area of interest, σ_f could conceivably range from 0% to 100%. The ratio of total green leaf area to its covered ground area (Curran 1983; Yin and Williams 1997) defines the LAI, which is a measure of the vegetation biomass. Typical values of LAI vary depending on the biome represented in a satellite pixel, but may have maxima between 6 and 8 for deciduous forests and between 2 and 4 for annual crops. Desert and tundra yield low LAI values near 0.1, while LAI for coniferous forests may exceed 15. Area-averaged LAI values such as those measured by satellite display lower maxima and a narrower range of values than point measurements (Scurlock et al. 2001).

Depending upon the sub-pixel structure of vegetation, σ_f and LAI have different relationships with NDVI. For example, if a satellite image pixel contains non-uniform dense vegetation, the relationship is

$$\mathrm{NDVI} = \sigma_f \mathrm{NDVI}_\infty + (1-\sigma_f)\mathrm{NDVI}_0, \tag{3.5}$$

where NDVI_0 and NDVI_∞ are the signals from bare soil and dense green vegetation, respectively. However, if a satellite pixel contains non-uniform non-dense vegetation, the relationship becomes

$$\mathrm{NDVI} = \sigma_f \mathrm{NDVI}_\infty(\mathrm{NDVI}_\infty - \mathrm{NDVI}_0)\exp(-kL_g) + (1-\sigma_f)\mathrm{NDVI}_0, \tag{3.6}$$

where k is an extinction coefficient and L_g is a leaf area index defined as the number of leaf layers over only the vegetated part of the pixel (Gutman and Ignatov 1998), in contrast to the conventional definition of LAI. Clearly, extracting both σ_f and LAI from NDVI is a

non-trivial process. An alternative approach prescribes one variable and derives the second. Some land surface models prescribe a constant value of LAI for all land use categories and allow σ_f to vary each month. Yin and Williams (1997) assume a linear relationship between LAI and NDVI, such that

$$\mathrm{LAI}_i = \mathrm{LAI}_{\max} \frac{\mathrm{NDVI}_i - \mathrm{NDVI}_{\min}}{\mathrm{NDVI}_{\max} - \mathrm{NDVI}_{\min}}, \tag{3.7}$$

where the subscripts *max*, *min*, and *i* refer to the climatological maximum, climatological minimum, and period values observed at a particular location, respectively. However, Gutman and Ignatov (1998) indicate that the derivation of σ_f from NDVI should be more accurate than the derivation of LAI from NDVI.

Chang and Wetzel (1991) introduced a two-line-segment method in which the linear relationship between σ_f and NDVI changes where the NDVI value exceeds a certain threshold:

$$\sigma_f = \begin{cases} 1.5(\mathrm{NDVI} - 0.1), & \mathrm{NDVI} \leq 0.547 \\ 3.2(\mathrm{NDVI}) - 1.08, & \mathrm{NDVI} > 0.547 \end{cases}, \tag{3.8}$$

where the values of σ_f are bounded by 0 and 1. This method has been used in several studies (e.g., Crawford et al. 2001, Kurkowski et al. 2003) and provides an optimal fit to field validation data under the assumption that σ_f and the fraction of photosynthetically active radiation (i.e., the fraction of visible light used by the green canopy for photosynthesis) are linearly related.

The Center for Advanced Land Management Information Technologies (CALMIT) at the University of Nebraska, Lincoln, Nebraska receives AVHRR High Resolution Picture Transmission (HRPT) data from a receiving station and provides processed σ_f and LAI data over the central United States. Location information embedded within the AVHRR Level 1b HRPT data allow for geometric correction of NDVI images to a Lambert Azimuthal map projection. A two-line-segment method (Eq. 3.8) determines σ_f and an empirical model (Eq. 3.7) provides estimates of LAI. Global field measurements of LAI (Scurlock et al.

2001) provide a basis for assigning a maximum LAI to each land cover classification in the National Land-Cover Database (NLCD, Homer et al. 2004). NLCD classifications derive from Landsat Thematic Mapper (TM) observations over the period 1992–2003. CALMIT mosaics and resamples the NLCD cell sizes from 30 m to 1 km to match the pixel resolution of AVHRR measurements. Maximum and minimum NDVI values for the growing season (April–September) are generated from AVHRR data spanning the period 1989–2002 from the United States Geological Survey (USGS) Earth Resources Observation Systems (EROS) Data Center.

When compared with a 5-year climatology for fractional vegetation coverage (Gutman and Ignatov 1998) implemented in several operational forecast models, a systematically low bias exists for σ_f derived from 15-day maximum NDVI composites over the period 15 April–15 September 2004. After alerting CALMIT scientists to the problem, they responded by applying a radiometric enhancement to both the σ_f and LAI data using calibration coefficients for each spectral channel embedded within the HRPT data stream (Goodrum et al. 2001). Radiometric enhancement is a standard digital image processing technique that customizes an image for a particular application and serves to adjust radiance measurements for changes in atmospheric conditions or instrument response characteristics. This modification produced more reasonable σ_f and LAI values characteristic of the unusually dry spring and wet summer observed in 2004.

Chapter 4

Soil temperature and moisture errors in Eta analyses

Clearly, forecast models require both accurate soil temperature and soil moisture initializations. Though efforts are under way to provide more extensive networks of soil moisture data from a variety of remote sensing and direct observational sources (Entekhabi et al. 1999; Leese et al. 2001; Seuffert et al. 2004; Crawford and Essenberg 2006), routine *in situ* observations of soil temperature and moisture suitable for data assimilation are currently unavailable over large areas of the continental United States and the world.

Due to the absence of a large observational soil-monitoring network, many forecast models implement complex land surface models to realistically determine soil hydrology. The NCEP operational Eta model (Black 1994) produces land surface analyses by continuously cycling temperature and moisture fields within the National Centers for Environmental Prediction–Oregon State University–Air Force–Hydrologic Research Lab (Noah) land surface model (LSM, Chen et al. 1996; Koren et al. 1999). In the past, these fields evolved only in response to radiation budget constraints and modeled precipitation, but NCEP recently upgraded the self-cycling process so that soil fields respond instead to radiation budget constraints and adjusted precipitation observations from both radar and gauge data over the United States.

Many modeling efforts have used NCEP Eta model analyses and forecasts over the continental United States as initial and boundary conditions for a variety of applications (e.g., Colle et al. 2001; Bright and Mullen 2002; Stensrud and Weiss 2002; Westrick et al. 2002; Zhender 2002; Brennan et al. 2003; Chen et al. 2004; Hart et al. 2004; Hoadley et al. 2004; Galewsky and Sobel 2005; Zamora et al. 2005; Zhong et al. 2005). The Eta model therefore provides very important initial land-surface conditions that strongly influence forecasts for both operational and research purposes. Unfortunately, many land surface models, including the Noah LSM, do not capture observed soil moisture variations when forced with atmospheric observations or cycled model output (Robock et al. 2000). Marshall et al. (2003) find a strong positive bias in soil moisture from the Eta model in comparison to Oklahoma Mesonet observations, but also noted that a change in the Eta model initialization procedure to a continuous self-cycling initialization for soil moisture significantly mitigated this bias. Marshall et al. (2003) also report a warm bias in soil temperatures at a depth of 5 cm in the late afternoon and a cool bias in the early morning. On the other hand, Robock et al. (2003) find good agreement when comparing soil temperature and moisture output from a more recently implemented version of the Noah LSM with observations from the Oklahoma Mesonet averaged over all of Oklahoma during 1998–99.

This portion of the study compares Eta model analyses of soil temperature and moisture at 0000 UTC and 1200 UTC with observations from the Oklahoma Mesonet between 1 March 2004 and 1 October 2005. In contrast to the findings of Robock et al. (2003), strong biases in model soil temperature exist, as well as a severe underestimation of soil moisture at all depths.

4.1 Eta model description

The NCEP Eta model (Black 1994) is initialized from analyses provided by the Eta Data Assimilation System (EDAS, Rogers et al. 1996; Nelson 1999). The EDAS first produces a 3-h forecast from its own analysis over the continental United States. The system then

uses this forecast as a background field for assimilating subsequent observations over this 3-h period and produces a new analysis valid at the end of the 3-h window. This process continues indefinitely, with forecasts out to 84 hours produced from the most recent EDAS analysis every six hours. The Eta model produces each EDAS forecast, and consequently the initial atmospheric and soil conditions are consistent with the forecast model and match its resolution, physics, and dynamics (Rogers et al. 1996). The absence of a complete set of observations of soil temperature and soil moisture necessitates continuously self-cycling soil fields within the EDAS without observational corrections or soil moisture nudging toward climatology. These soil fields evolve only in response to external forcing from model physics and surface forcing in the form of precipitation and the surface radiation balance within the EDAS.

Prior to a modification on 16 March 2004, the EDAS assimilated hourly precipitation data consisting of radar and gauge observations from NCEP Stage II and Stage IV analyses (Fulton et al. 1998; Lin et al. 2005). These analyses exhibit a systematic dry bias which, when used as the driver for soil moisture, leads to drier soil. After an adjustment on this date, comparisons of the cumulative 24-h precipitation from EDAS against daily gauge analyses, inflated by 10% to correct for catchment errors, yield a long history of net deficits or surpluses in precipitation. Adjustments to the EDAS hourly precipitation input based on this history attempt to eliminate the deficit or surplus over 24 hours. Adjustments remain limited to $\pm$20% of the hourly precipitation analysis values and only apply to grid points in the analysis with non-zero precipitation. The EDAS assimilates the adjusted hourly precipitation input and then models the precipitation field. This modeled precipitation drives the land surface physics, though the modeled precipitation does not necessarily match the bias-adjusted observations (Lin et al. 2005).

A more extensive modification to the land surface scheme occurred on 3 May 2005 in the operational Eta model, now termed the North American Mesoscale (NAM) model. Previously, the EDAS would create precipitation during the assimilation process in regions where the Eta model did not forecast precipitation. The renamed NAM Data Assimila-

tion System (NDAS) no longer adjusts precipitation totals in locations where the precipitation from the NAM model is less than the bias-adjusted observations. However, the latent heat and moisture fields are reduced where the modeled precipitation is greater than the bias-adjusted observations. More importantly, the NDAS drives the land surface physics directly with the bias-adjusted observations rather than with the NDAS modeled precipitation, resulting in moister soil. The previous version tended toward a dry bias during the assimilation because the modeled precipitation did not exactly replicate observed precipitation coverage and intensities. The new method allows for a more robust and more accurate precipitation assimilation that increases soil moisture. Additionally, there is no longer an upper limit for cloud water mixing ratios when computing optical depths, which improves radiation absorption, and modifications to the cloud cover parameterization allow for more fractional cloudiness (DiMego and Rogers 2005).

Simultaneous upgrades to the Noah LSM addressed low-level temperature and humidity biases. Vegetation and soil databases have more classes with higher spatial resolution. A 1-km resolution, USGS 24-class vegetation type database replaced the 13-class, 1-degree resolution simple biosphere (SiB) vegetation types (Sellers et al. 1986). For soil characteristics, the 1-km resolution, 16-class State Soil Geographic Database (STATSGO, Miller and White 1998) data eclipsed the 1-km resolution, nine-class Zobler soil types (Zobler 1986). A 1-degree database of soil temperatures at the lower boundary at 300 cm depth replaced an old 2.5-degree soil temperature database. In addition, model developers lowered the leaf area index and compensated for the effect of the new precipitation assimilation procedures on the existing soil moisture bias by tuning the canopy conductance and other vegetation parameters within the Noah LSM. A lowered roughness length for heat reduces the skin temperature, thereby lowering the 2-m temperature forecasts and reducing the warm bias, though this does not change latent or sensible heat fluxes significantly. Overall, these modifications reduce drying trends and increase the low-level moisture (DiMego and Rogers 2005).

4.2 Comparison with observations

Gridded 40-km Eta model analyses of soil temperature and moisture at 0000 UTC and 1200 UTC are bilinearly interpolated to Oklahoma Mesonet observation sites allowing for direct model verification. While Eta model soil analyses are available at the present operational grid spacing of 12 km, researchers seldom use these analyses for initializing forecast models. Comparisons span the period from 1 March 2004 through 1 October 2005. This period is sufficient to characterize the performance of the EDAS soil temperature and moisture schemes both before and after the change from continuously self-cycling modeled precipitation and radiation to assimilation of precipitation observations on 3 May 2005.

Point measurements of soil temperature and moisture are not as spatially representative as atmospheric measurements, primarily due to spatial heterogeneities in vegetation coverage and soil types (Marshall et al. 2003; Brotzge and Crawford 2003). For this reason, spatial and temporal averaging of observations reduces small-scale noise and enables model validation and intercomparisons (e.g., Marshall et al. 2003; Robock et al. 2003). However, interpolating observations to a model grid yields comparisons that are partly a function of the interpolation scheme rather than the underlying observations. An analysis scheme cannot account for small spatial variations in the observations and thus analyzed and observed values may differ considerably (Schlatter 1975). Moreover, individual observation points, and not areal averages, provide the raw data for objective analysis schemes that produce gridded initial conditions for models. As shown in chapter 2, initial soil conditions strongly influence model forecast results. It is therefore important to correctly estimate point values of soil temperature and soil moisture in the Eta model so that these values can provide meaningful initial conditions for other numerical models with different grid sizes. The choice to average point comparisons in this study rather than interpolate the observations to the model grid permits a bulk characterization of the model performance over all of Oklahoma without introducing errors via an objective analysis scheme. This approach is similar to other studies that compare model output and observations (e.g., Crawford et

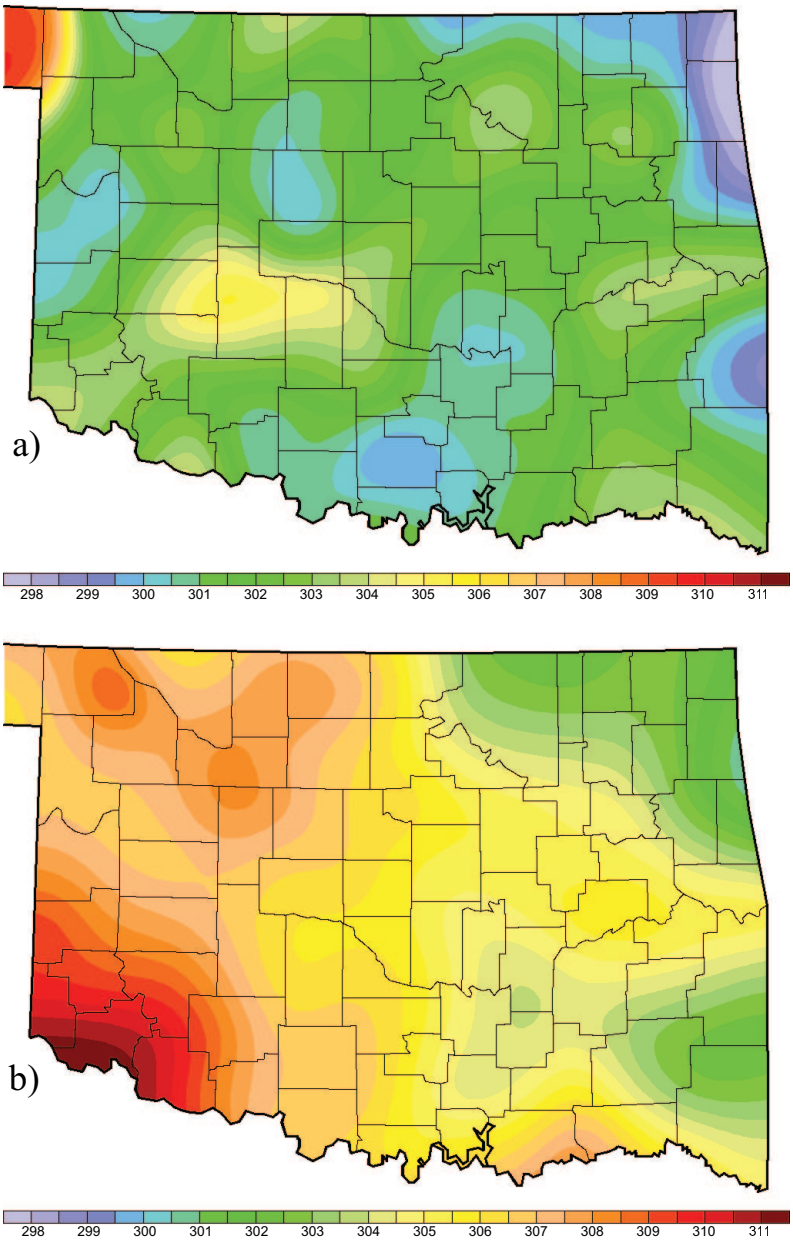

FIG. 4.1: Soil temperature (K) at 0000 UTC 15 July 2005 from a) Oklahoma Mesonet observations at a depth of 5-cm under sod and b) the 0–10 cm soil layer of the 0000 UTC Eta analysis.

al. 2000, 2001; Santanello and Carlson 2001; Robock et al. 2003).

Though comparisons between the Eta model and observations only include point measurements, Figures 4.1 and 4.2 provide informative visualizations of the geographic variability of Oklahoma Mesonet 5-cm soil temperature and moisture observations compared with Eta model 0–10 cm soil temperature and moisture analyses for a representative sum-

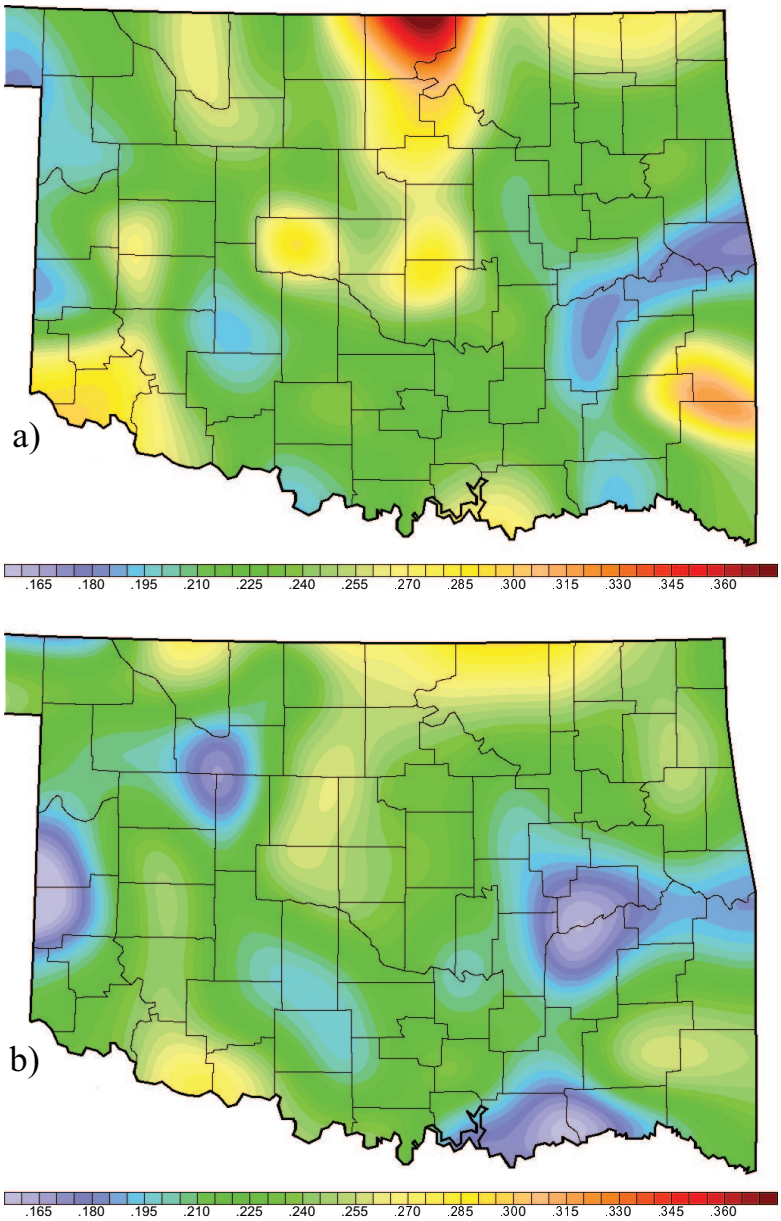

FIG. 4.2: Soil moisture ($m^3\ m^{-3}$) at 0000 UTC 15 July 2005 from a) Oklahoma Mesonet observations at a depth of 5-cm and b) the 0–10 cm soil layer of the 0000 UTC Eta analysis.

mer day. The Oklahoma Mesonet observations are interpolated to a 3-km horizontal grid using a two-pass Barnes analysis (Barnes 1973). The Eta analyses, shown here interpolated to the same 3-km horizontal grid, display a cool and dry bias typical of many 0000 UTC analyses. In addition, the differences in the patterns of each field can influence the subsequent forecast.

The Noah LSM model within the EDAS contains five soil layers representing depths of 0–10 cm, 10–40 cm, 40–100 cm, and 100–200 cm, and a constant reservoir temperature at 300 cm. The physical equations in the Noah LSM predict the soil temperature and soil moisture at the midpoint of each soil layer. Soil temperatures in the 0–10 cm model layer are compared with Oklahoma Mesonet observations at a depth of 5 cm and soil temperatures in the 10–40 cm model layer are compared with observations at a depth of 30 cm. For soil moisture, the values from the Eta analysis in the 0–10 cm, 10–40 cm, and 40–100 cm layers are compared with observations at depths of 5 cm, 25 cm, and 60 cm, respectively. These direct comparisons allow computation of root-mean squared error (rmse) and bias (Wilks 2006, 279–280) across the entire Oklahoma Mesonet.

4.2.1 Soil temperature

There is a strong positive soil temperature bias (forecasts minus observations) in the 0–10 cm layer from 0000 UTC Eta model analyses compared with observations of 5-cm soil temperatures from all Oklahoma Mesonet sites (Fig. 4.3). Twelve hours later at 1200 UTC, there is a predominately negative bias. The conspicuous spike representing the 1200 UTC 6 March 2005 bias is a notable exception that depicts an analysis problem for the Oklahoma-wide soil temperatures in the 0–10 cm layer. Overall, the bias for this most shallow soil layer is 4.1°C (-1.0°C) and the rmse is 5.0°C (2.4°C) for 0000 UTC (1200 UTC) Eta analyses. Ground temperature errors of this magnitude can produce errors in upward longwave radiation of over 20 W m^{-2} during the summer. Errors appear reduced in magnitude in the deeper 10–40 cm soil layer, and the 0000 UTC and 1200 UTC soil temperature analyses differ only slightly (Fig. 4.4). There is a temporally coherent pattern of errors throughout the year such that errors of the same sign persist for multi-week periods. This trend appears to follow the more variable pattern of daily biases in the upper soil layer. Modifications to the land surface model on 3 May 2005 do not appear to affect significantly the magnitude of subsequent soil temperature errors.

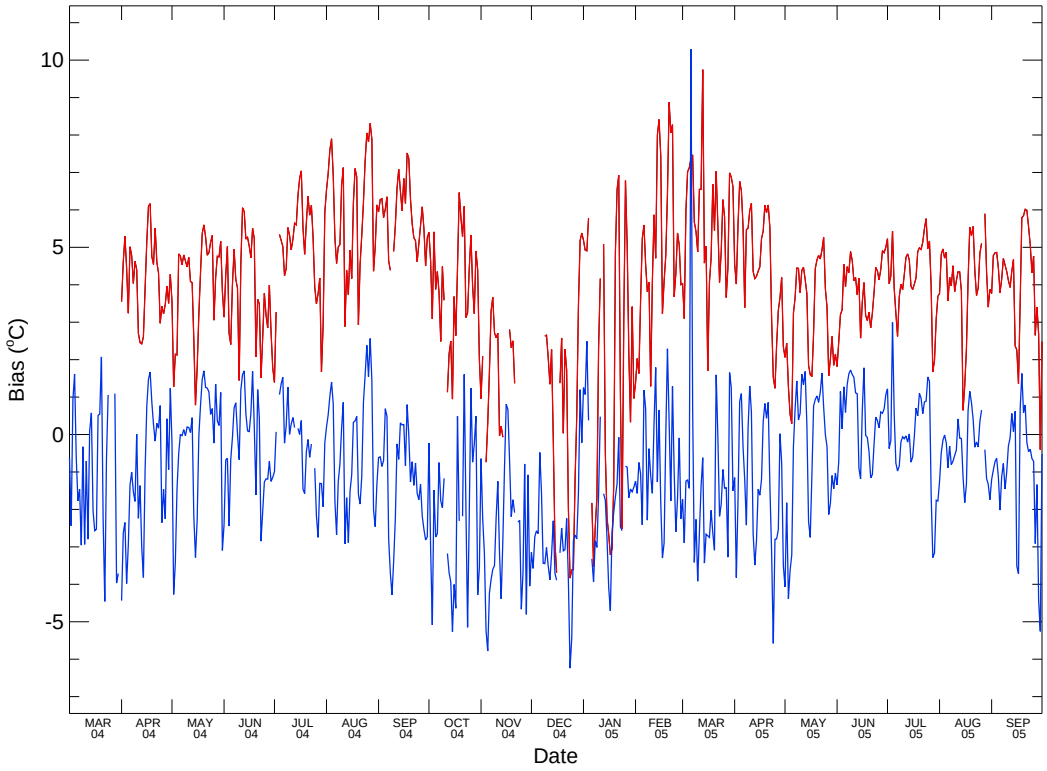

FIG. 4.3: Point calculations of daily soil temperature bias (°C) averaged over all of Oklahoma in the 0–10 cm layer from 0000 UTC (red) and 1200 UTC (blue) Eta analyses compared with 5-cm soil temperature observations from the Oklahoma Mesonet.

While the physical equations predict the soil temperature at the midpoint of a given soil layer, soil temperatures in the Eta model physically represent an average in that layer. A more strict comparison with observations therefore requires an integrated soil temperature throughout a layer rather than point measurements at a specific depth. A cubic spline interpolation between observations of skin temperature and soil temperature at depths of 5 and 10 cm, summed over 5 mm increments, allows an estimate of the soil temperature in the 0–10 cm layer. This integrated soil temperature compares well with the 5-cm soil temperature observations. The daily difference between the Oklahoma-wide 0–10 cm soil temperature bias in Eta analyses calculated from either direct measurements at 5-cm or an integrated temperature in the layer may reach as high as 2°C. However, the overall error statistics for the 0–10 cm layer calculated using an integrated soil temperature do not change substantially compared with the direct measurements shown in Fig. 4.3.

FIG. 4.4: Point calculations of daily soil temperature bias (°C) averaged over all of Oklahoma in the 10–40 cm layer from 0000 UTC (red) and 1200 UTC (blue) Eta analyses compared with 30-cm soil temperature observations from the Oklahoma Mesonet.

4.2.2 Soil moisture

There is a pervasive and persistent dry bias in both the 0000 UTC and 1200 UTC Eta soil moisture analyses. For each day, the Oklahoma-wide average soil moisture in the 0–10 cm model layer of the Eta analyses is generally drier than the observations at 5 cm (Fig. 4.5). In the 10–40 cm layer, the soil moisture bias slightly exceeds zero for only a single 0000 UTC Eta analysis and in the 40–100 cm layer, the soil moisture bias never becomes positive over the period of study (Figs. 4.6 and 4.7). Overall, the bias for each soil layer is -0.03 m^3 m^{-3}, -0.05 m^3 m^{-3}, and -0.09 m^3 m^{-3} for the 0–10 cm, 10–40 cm, and 40–100 cm Eta model layers, respectively. In the 40–100 cm Eta model layer, the daily average soil moisture error across all of Oklahoma reaches as large as 35% of the typical range of soil moisture when compared with observations at a depth of 60 cm.

There is notable improvement in the analyzed soil moisture fields after the change from

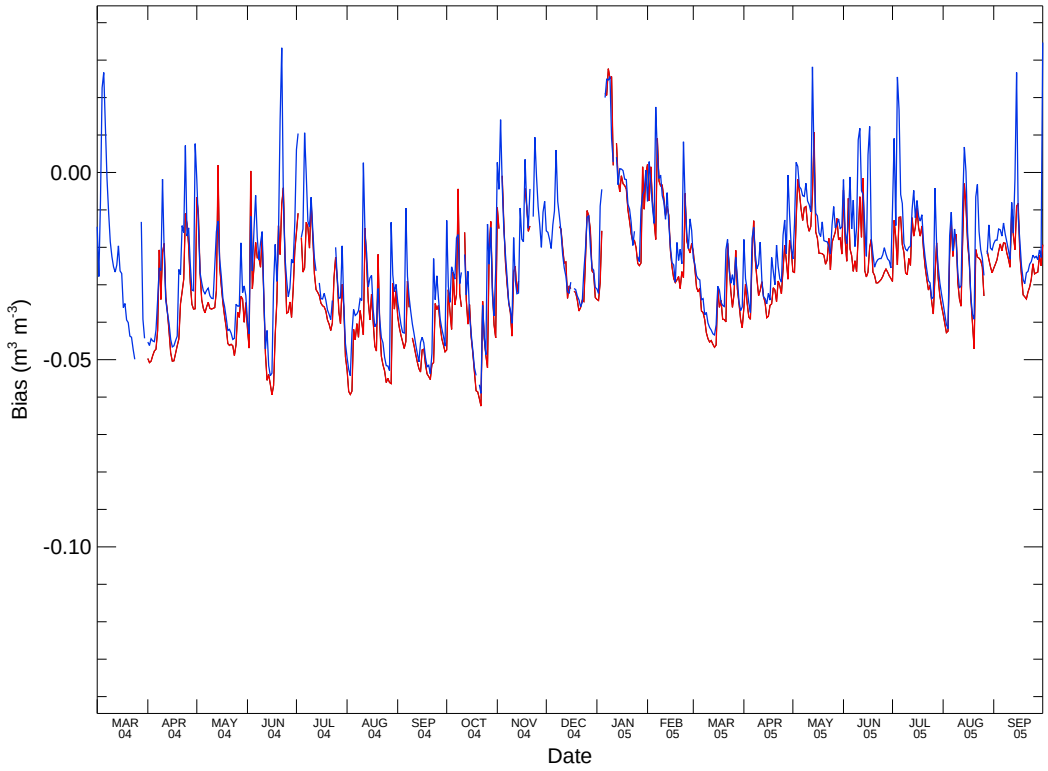

FIG. 4.5: Point calculations of daily soil moisture bias ($m^3\ m^{-3}$) averaged over all of Oklahoma in the 0–10 cm layer from 0000 UTC (red) and 1200 UTC (blue) Eta analyses compared with 5-cm soil moisture observations from the Oklahoma Mesonet.

self-cycling precipitation to observed precipitation assimilation on 3 May 2005. While this change reduced the magnitude of the errors, and evidences itself as a large discontinuity in the bias time series of Figures 4.6 and 4.7, a strong dry bias persists in the soil moisture field.

4.3 Discussion

Systematic biases clearly exist in Eta analyses of both soil temperature and soil moisture. Consistent with the results of Marshall et al. (2003), soil temperatures in the most shallow soil layer tend to be too warm at 0000 UTC and too cool at 1200 UTC. Positive soil temperature errors in 0000 UTC Eta analyses likely stem in part from the documented excess of solar radiation during the daytime (Zamora et al. 2005), while the generally negative

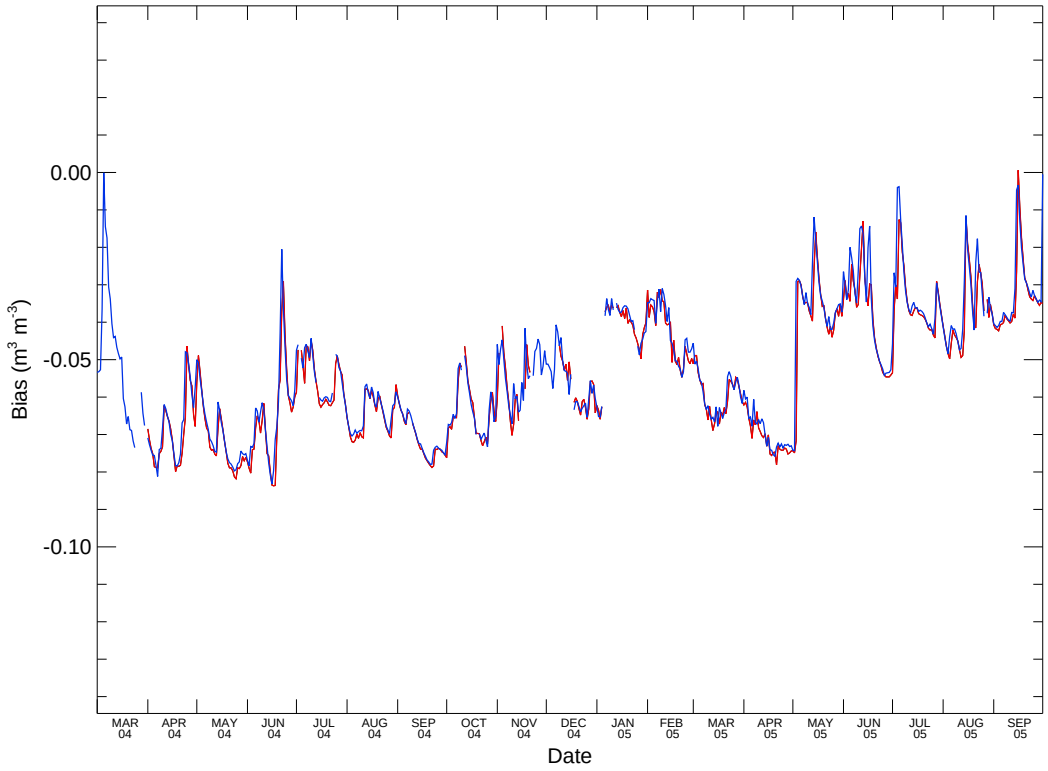

FIG. 4.6: Point calculations of daily soil moisture bias ($m^3\ m^{-3}$) averaged over all of Oklahoma in the 10–40 cm layer from 0000 UTC (red) and 1200 UTC (blue) Eta analyses compared with 25-cm soil moisture observations from the Oklahoma Mesonet.

soil temperature biases in 1200 UTC Eta analyses result from underestimated downward longwave radiative fluxes during nighttime hours (Stensrud et al. 2006). Modifications to the land surface physics on 3 May 2005 did not mitigate these errors; soil temperatures in the top soil layer remain too high in the 0000 UTC Eta analyses and dry soil moisture biases continue in each of the top three soil layers. Tests indicate that these systematic biases in both soil temperature and moisture do not appear to be strongly dependent upon soil or vegetation types defined in Eta model grid cells.

At the Eufaula Oklahoma Mesonet site, the EDAS soil moisture errors in the top two model layers result from both an inappropriate response to rainfall events and accelerated desiccation of the soil compared with observations, particularly in the 10–40 cm layer (Fig. 4.8). The response to precipitation in the 40–100 cm layer appears limited except after several consecutive days of heavy precipitation. The new precipitation assimilation

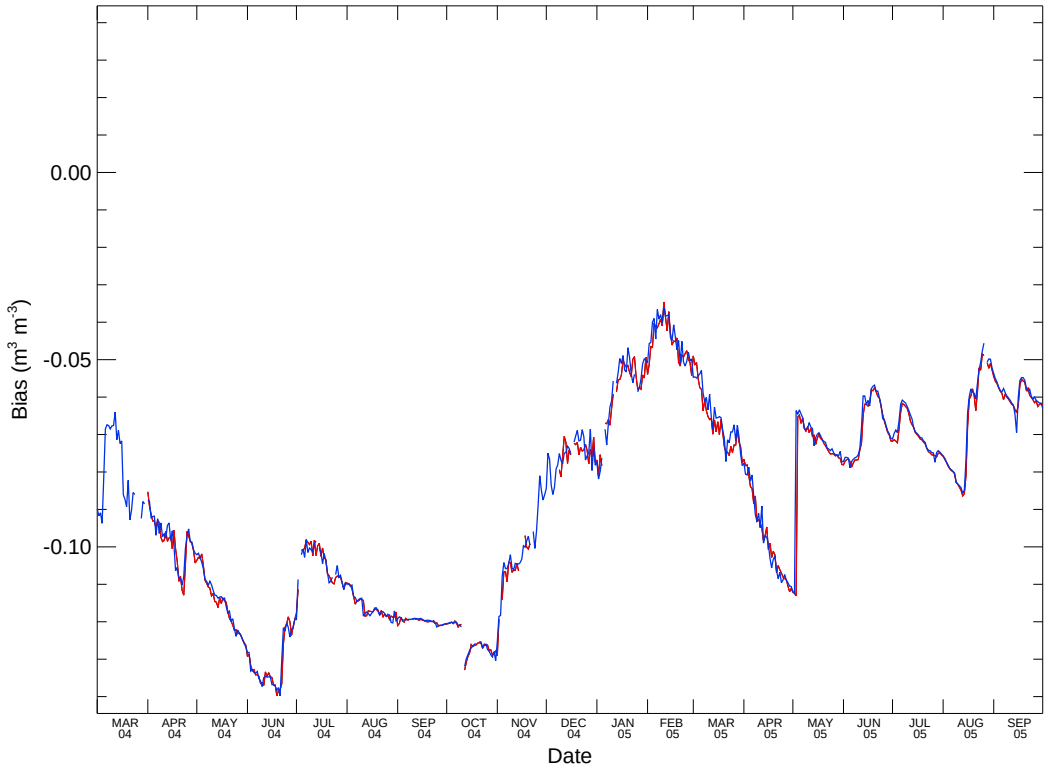

FIG. 4.7: Point calculations of daily soil moisture bias ($m^3\ m^{-3}$) averaged over all of Oklahoma in the 40–100 cm layer from 0000 UTC (red) and 1200 UTC (blue) Eta analyses compared with 60-cm soil moisture observations from the Oklahoma Mesonet.

procedure implemented on 3 May 2005 somewhat improved soil moisture estimates at some Mesonet sites, though systematic dry biases remain in the Eta analyses.

An exploration of the influence of soil heat capacity can help to address the effect of such a dry bias on soil temperatures. Soil heat capacity is a function of soil moisture and directly affects the diagnosis of soil temperature. Underestimates of soil moisture such as those in Eta model analyses could therefore result in poorly estimated soil temperatures. A simple, one-layer slab soil model driven by Oklahoma Mesonet observations allows approximate calculations of the influence of errors in soil moisture alone on soil temperature. The composite soil volumetric heat capacity employed in the slab model is

$$C_g = \Theta C_{\text{water}} + (1 - \Theta_s) C_{\text{soil}} + (\Theta_s - \Theta) C_{\text{air}}, \tag{4.1}$$

where Θ is the soil volumetric water content, $C_{\text{water}} = 4.2 \times 10^6\ \text{J m}^{-3}\ \text{K}^{-1}$, $C_{\text{soil}} = 1.26 \times$

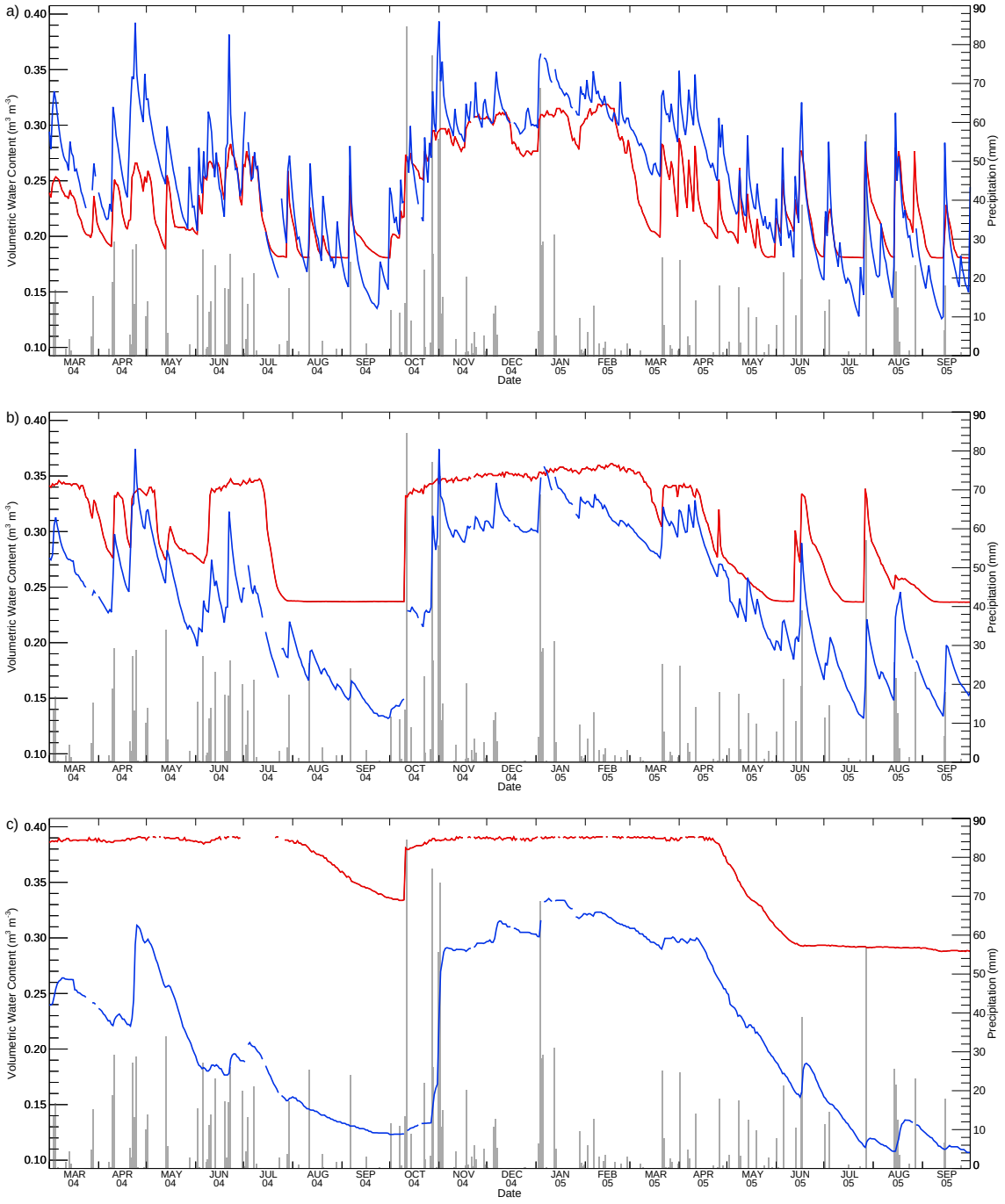

FIG. 4.8: Observed soil moisture at 1200 UTC at Eufaula (red) at depths of a) 5 cm, b) 25 cm, and c) 60 cm compared with 1200 UTC Eta analyses (blue) in the 0–10, 10–40, and 40–100 cm soil layers, respectively, and observed daily (0000 UTC–0000 UTC) precipitation totals (bars).

10^6 J m^{-3} K^{-1}, and $C_{\text{air}} = 1004$ J m^{-3} K^{-1} are the volumetric heat capacities of water, soil, and air, respectively, and Θ_s is the soil porosity (Chen and Dudhia 2001). The soil porosity depends upon the soil texture (Cosby et al. 1984) determined from soil cores at each observation site. The slab model predicts the soil temperature T at a depth of 5 cm using

$$C_g d_s \frac{\partial T}{\partial t} = Q_{G_S}, \tag{4.2}$$

where Q_{G_S} is the storage ground heat flux and $d_s = 10$ cm is the depth of the slab. Since the observation frequency for soil temperature is 15 minutes and that for soil moisture is 30 minutes, the slab model linearly interpolates the soil moisture observations to obtain a complete time series of data at 15-minute intervals. Unfortunately, the Oklahoma Mesonet sensors do not directly measure the storage ground heat flux, and instead obtain the best possible estimate based on soil temperature, soil moisture, and average soil properties at selected Mesonet sites. When using Eq. (4.2) with estimated Q_{G_S}, the observed volumetric water content value, and an initial soil temperature equal to the observed value at 5 cm, the slab model produces soil temperatures that slowly diverge from observations. For this reason, an improved estimate of Q_{G_S} is calculated by determining the value of Q_{G_S} needed to produce the observed 5-cm soil temperature, given the observed volumetric water content.

The sensitivity of the slab model to errors in the volumetric water content is explored using the improved estimates of Q_{G_S} for each 15-minute period. Ground temperatures from model simulations produced for equal positive and negative volumetric water content biases are compared with observations. While this simple model does not account for the influence of differing soil moisture on the storage ground heat flux or the surface energy balance, it represents an idealized approach to determine the effect of soil moisture errors on soil temperature forecasts.

Given observations and soil characteristics at the Watonga Oklahoma Mesonet site for 72 hours beginning at both 0000 UTC and 1200 UTC 20 July 2004, this simple one-layer slab soil model estimates the 5-cm soil temperatures that would develop if the observed 5-

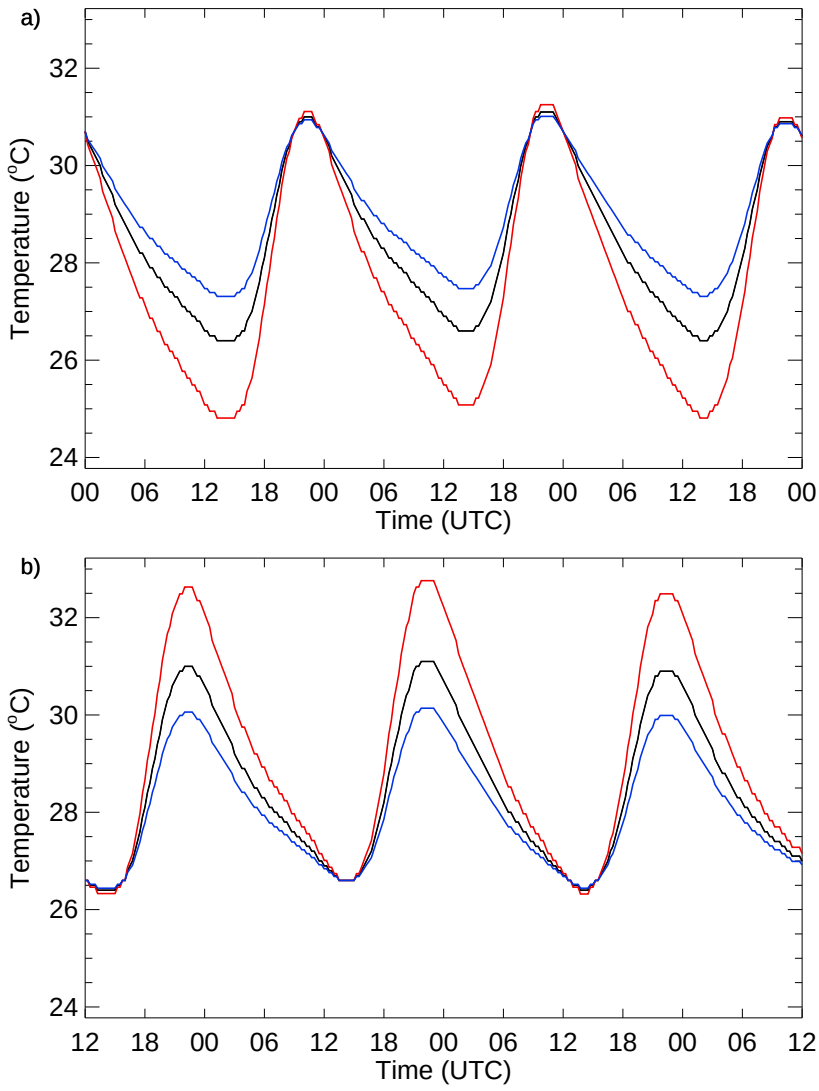

FIG. 4.9: Slab soil model temperatures (°C) initialized by a) 0000 UTC and b) 1200 UTC 5-cm soil temperature observations at Watonga on 20 July 2004. Soil moisture errors of $+0.1$ m^3 m^{-3} (blue) and -0.1 m^3 m^{-3} (red) yield temperatures that differ from observed soil temperatures (black).

cm soil moisture error were equal to ± 0.1 m^3 m^{-3} (Fig. 4.9), or twice the soil moisture error seen in the Eta analyses. Different initialization times show the effect of a soil moisture bias on each part of the diurnal cycle. Results reveal that negative soil moisture biases alone may account for more than 1.6°C increases (decreases) in maximum (minimum) daily soil temperatures. Positive soil moisture biases account for a more modest reduction of about 0.9°C in the amplitude of the diurnal soil temperature cycle. While underestimates of soil moisture may contribute to the sign of the soil temperature errors shown in Figure 4.3, soil

moisture alone apparently cannot account for the magnitude of the soil temperature errors in Eta analyses.

As previous studies have shown, soil temperature and soil moisture estimates strongly impact forecasts by numerical weather prediction models that implement sophisticated land surface parameterizations. Problems with soil fields in Eta analyses, which provide initial conditions for a variety of research and operational modeling applications, may negatively impact the resulting model forecasts. These existing biases suggest the strong need for an extensive network of soil observations, in addition to atmospheric surface observations, and the necessity for assimilating those observations into land surface initializations. Model simulations in the following chapter test the impact of proper initial soil conditions in mesoscale model forecasts.

Chapter 5

Model simulations

5.1 Model description

The Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) fifth-generation Mesoscale Model (MM5) version 3.6 (Dudhia 1993; Grell et al. 1995; Dudhia 2003) is a nonhydrostatic mesoscale modeling system with advanced model physics jointly developed by The Pennsylvania State University and NCAR. MM5 is a grid-point model with variables distributed on an Arakawa B grid (Arakawa and Lamb 1977). The model solves differential equations with finite differences centered in space and time. Second-order finite differences apply to advection terms and an Asselin time filter smoothes all prognostic variables. In the late 1990s, MM5 was the most popular mesoscale modeling system for local and regional modeling efforts (Mass and Kuo 1998) and today continues as a tool for numerical weather prediction, air quality studies, and hydrological studies (Chen and Dudhia 2001). As computing power increased, and grid spacing decreased, model developers recognized the need for a computationally inexpensive way to correctly treat land surface processes in order to capture the effect of small-scale land surface variations on surface energy fluxes and the atmosphere. In response, they implemented within MM5 the Oregon State University land surface model (OSULSM), an advanced LSM that improved the simulation of boundary layer and precipitation processes and surface heat fluxes (Chen

and Dudhia 2001). The Noah LSM is the newest version implemented in MM5 version 3.6 and represents a community effort between several agencies over many years. The Noah LSM contains four soil layers depicting soil temperature and soil moisture and accounts for vegetation categories, monthly vegetation fraction, and soil texture, and includes parameterizations for evaporation, soil drainage, runoff, the root zone, and canopy moisture (Skamarock et al. 2005). It also features frozen-soil physics, snow cover prediction, and the ability to ingest albedo data (Dudhia 2003).

Benefits of MM5 over similar numerical weather prediction models for use over the Great Plains include the implementation of the sigma vertical coordinate system, excellent documentation and technical support, parallelization, and the available land surface physics options. The Noah LSM functions as the primary land surface model in MM5 and contains nearly identical code to the land surface schemes found in both the NCEP operational Eta model and the Weather Research and Forecasting (WRF) modeling system. This allows for direct compatibility between the time-dependent soil variables and surface fluxes in MM5 simulations and the Eta model analyses that initialize the simulations.

The Eta model consistently provides analyses and forecast output every three hours from each of its four runs per day, making it an ideal source for initial and boundary conditions for any potential case study. The operational Eta model and its companion EDAS both have a relatively high resolution of approximately 12 km, making its initial soil moisture useful for initializing MM5 coupled with the Noah LSM (Chen and Dudhia 2001), though the approximately 40-km output grid initializes the model in the following simulations. Despite the higher resolution available from the native 12-km forecast grid (Eta grid 218), the 40-km output grid (Eta grid 212) contains substantially more information on atmospheric variables, is considerably smaller, and consumes less bandwidth during daily downloads.

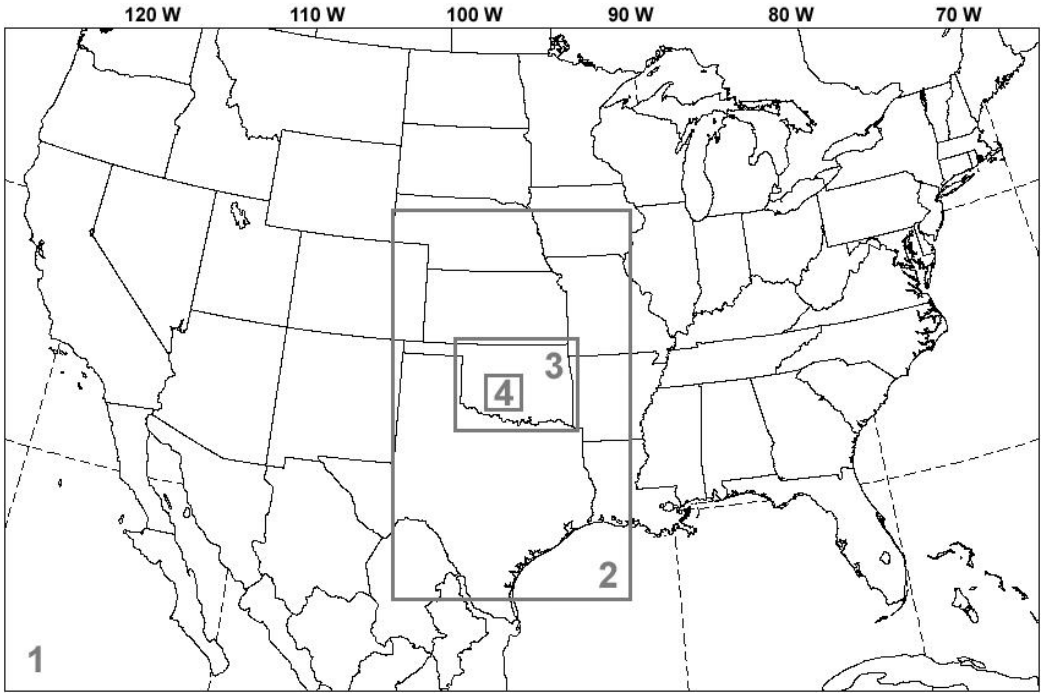

FIG. 5.1: Location of the four nested MM5 domains with 27-, 9-, 3-, and 1-km grid resolution.

5.2 Parameter selections

The primary study area focuses on Oklahoma due to the availability of Oklahoma Mesonet observations for soil measurements and model verification. MM5 is used to produce 48-hour forecasts on four nested model domains with 27-, 9-, 3-, and 1-km grid resolution (Fig. 5.1) and 23 vertical half-sigma levels. NCEP Eta model analyses initialize the MM5 forecasts and Eta model forecasts provide boundary conditions every six hours. Specific user-defined options include the Kain and Fritsch (1993) cumulus parameterization on domains one and two only, no shallow convection on any domain, the Medium-Range Forecast model (MRF) planetary boundary layer (PBL) parameterization (Hong and Pan 1996) with moist vertical diffusion in clouds, simple ice microphysics (Dudhia 1989), and the Rapid Radiative Transfer Model (RRTM) longwave radiation scheme (Mlawer et al. 1997). The Dudhia (1989) solar radiation parameterization determines the surface downward shortwave radiation. As mentioned previously, the Noah LSM provides the multi-

layer soil physics and vegetation package. The coarse domain time step is 60 seconds and each domain receives initial conditions from the Eta analyses rather than from values interpolated from the coarse domain. Other options and parameters remain set to their default values.

5.3 Atmospheric variables on standard observing levels

Valid comparisons between model forecasts and observations of typical atmospheric variables require interpolation between the surface and the lowest sigma level such that forecasts are valid at World Meteorological Organization (WMO) standard observing levels. Standard heights for air temperature measurements are 1.25–2 m above ground level (AGL) and the standard height for wind measurements is 10 m AGL (WMO 1996). In the MRF PBL scheme, MM5 calculates forecast fields for air temperature and mixing ratio at 2 m AGL and for wind speed and direction at 10 m AGL. The MRF PBL scheme follows Blackadar (1976) and assumes that the atmospheric layer between the surface and the lowest half-sigma level satisfies Monin-Obukhov similarity theory. The 2-m temperature is an interpolated value found by converting the 2-m potential temperature, θ_2, to temperature using the value of surface pressure. The 2-m potential temperature is

$$\theta_2 = \theta_g + \Delta\theta \frac{\psi_T(2\mathrm{m})}{\psi_T(\sigma)}, \tag{5.1}$$

where θ_g is the potential temperature of the ground surface and $\Delta\theta$ is the difference between the surface potential temperature and the potential temperature at the lowest half-sigma level, which is scaled by the ratio of empirical stability correction functions at 2 m, $\psi_T(2\mathrm{m})$, and at the height of the lowest half-sigma level, $\psi_T(\sigma)$ (Paulson 1970; Dyer 1974). Similarly, the 2-m mixing ratio is

$$w_2 = w_{s_g} + (w_\sigma - w_{s_g}) \frac{\psi_w(2\mathrm{m})}{\psi_w(\sigma)}, \tag{5.2}$$

where w_{s_g} is the saturation mixing ratio at the ground surface temperature, w_σ is the mixing ratio at the lowest half-sigma level, and $\psi_w(2\text{m})$ and $\psi_w(\sigma)$ are again empirical stability correction functions. The zonal and meridional wind components at the lowest half-sigma level are scaled by stability correction functions to give the components of horizontal wind speed at 10 m AGL. The stability correction functions for temperature, mixing ratio, and wind speed differ depending upon the stability class of the planetary boundary layer. Each of these four classes is determined by the value of the bulk Richardson number, R_B, where $R_B \geq 0.2$ is stable, $0.0 < R_B < 0.2$ is damped mechanical turbulence, $R_B = 0.0$ is forced convection, and $R_B < 0.0$ is free convection. For damped mechanical turbulence, the stability correction functions are strongly dependent upon the Obukhov length.

5.4 Model verification

Verification of MM5 model forecasts using Oklahoma Mesonet data is a necessary step in a proper evaluation of the impact of improved land surface conditions on model forecasts. An objective analysis of the observations to the model grids would result in verification statistics that are dependent upon the chosen analysis procedure. Exceptions may include a verification procedure requiring smoothed fields to eliminate small-scale variability or computations of difference fields between observations and model forecasts. A more accurate method, therefore, involves interpolating forecasts from the model grids to the Oklahoma Mesonet site locations (e.g., Crawford et al. 2001; Kurkowski et al. 2003; Marshall et al. 2003; Brotzge 2004). Verification statistics derive from corresponding forecast and observation pairs.

The MM5 model grid points are created using a Lambert conformal map projection. A traditional and accepted method for interpolating forecasts to observation points is to convert the latitude and longitude site locations to a point in Lambert conformal map coordinates and then bilinearly interpolate the forecasts to this point. This procedure requires valid forecasts at each of the four grid points forming a grid square surrounding the obser-

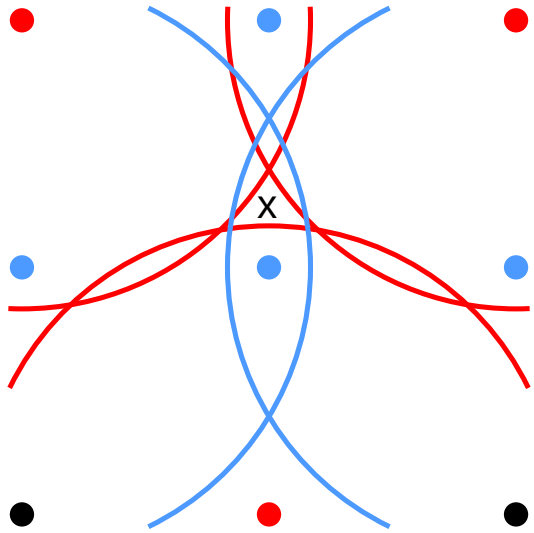

FIG. 5.2: The four closest grid points (blue) to the observation site ('x'), more distant points (red), and very distant points (black). Arcs with constant radii indicate relative distance from the observation site to several grid points.

vation point.

There are two problems with a bilinear interpolation approach to model verification. First, the four grid points defining the grid cell in which the observation point lies are not necessarily the four geographically closest grid points. Figure 5.2 illustrates this argument. The blue dots indicate the four closest grid points to an observation site at the location marked by the 'x'. Blue semi-circles of equal radii surround the easternmost and westernmost blue points showing that the observation site lies within this particular radius of influence. The observation site lies outside this same radius for the red points, as indicated by the red curves.

The second and most critical problem arises from the treatment of soil temperature and moisture in the Noah LSM chosen for MM5. Regardless of the initial conditions, the model replaces existing soil temperatures with a constant value over water bodies. The model ignores the influence of such an anomaly in its calculations by performing detailed LSM calculations only for grid points over land. However, soil temperature fields in the model output retain the constant values over water. Naturally, the model also replaces volumetric water content with a constant 1.0 m^3 m^{-3} on grid points over water bodies.

Bilinear interpolation for the situation depicted in Figure 5.3 would result in inaccurate soil temperature and moisture values interpolated to the observation site indicated by the 'x'. A comparison would match unrealistically low soil temperature forecasts and high soil moisture forecasts with observations in such an arrangement. This situation is a problem at the Bee, Butler, and Clayton Oklahoma Mesonet sites for MM5 grid resolutions of both 3-km and 1-km. The land use category is water beneath one (Butler and Clayton) or two (Bee) of the four grid points defining the grid cells surrounding each station.

A simple solution to each of these problems is to calculate a distance-weighted average forecast value at each observation point with the requirement that the grid point is over land. That is,

$$X = \frac{\sum_{i=1}^{N} G_i \left(1 - \frac{d_i}{\Sigma_{i=1}^{N} d_i}\right)}{N-1}, \tag{5.3}$$

where X is the interpolated value at the observation site, G_i is the forecast value at grid point i, d_i is the distance from the i^{th} grid point to the observation site, and $N > 1$ is the total number of grid points used in the weighted average. If one of the four closest points lies over a water body, the interpolation instead uses the next closest grid point over land.

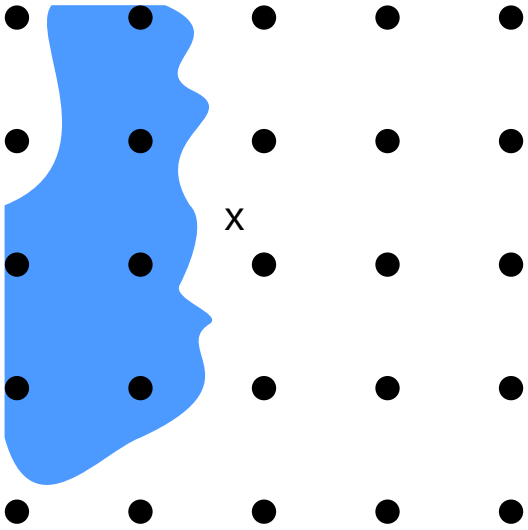

FIG. 5.3: Grid points (black dots) surrounding an observation site ('x') near a lake (blue).

The four closest grid points over land determine the interpolated value and the result does not necessarily require an average of the forecasts at the four grid points defining a grid cell. Comparisons for all forecast variables use the same four closest grid points over land.

Distances between grid points and observing sites are very small on high spatial-resolution model grids, such as domains three and four. Because this method exists as an improvement over bilinear interpolation, it is extremely important to measure such distances accurately. Assuming a spherical Earth with a constant radius would introduce large errors in distance calculations. To avoid such errors, the World Geodetic System 1984 (WGS 84) Ellipsoid provides the geographic datum upon which to calculate distances. Complete details of the WGS 84 Coordinate System appear in National Imagery and Mapping Agency (2000). The most accurate inverse geodetic formulae presented in Sodano (1965) yield the distance between grid points and the observation site. Available latitude and longitude coordinates for each Mesonet site have a precision that corresponds to about 0.01 km at the latitude of Oklahoma. This precision error is much larger than the distance errors from the Sodano (1965) inverse geodetic formulae.

A distance-weighted average interpolation scheme should provide reasonable and smooth interpolated forecast values at observation points. To verify that this is true, define an analytic soil temperature field by

$$\begin{aligned} F(x,y) = T_{\max} - \frac{T_{\max} - T_{\min}}{x_n y_n} xy + 1.5\cos^2\left(\frac{2\pi x}{L_x}\right)\sin^2\left(\frac{2\pi y}{L_y}\right) \\ - 0.5\sin\left(\frac{4\pi(x-y)}{L_x + L_y}\right) + 0.5\cos\left(\frac{\pi y}{2L_y}\right), \end{aligned} \tag{5.4}$$

where $T_{\max}$ and $T_{\min}$ are the maximum and minimum soil temperatures in the field, x_n and y_n are the maximum grid dimensions, and L_x and L_y are the zonal and meridional wavelengths, respectively. Figure 5.4 shows this field on a rectangular 178 $\times$ 133 point grid superimposed on a map of Oklahoma with $T_{\max} = 22$°C, $T_{\min} = 16$°C, $L_x = 178$ grid units and $L_y = 133$ grid units. This analytic field represents a typical magnitude, range, and

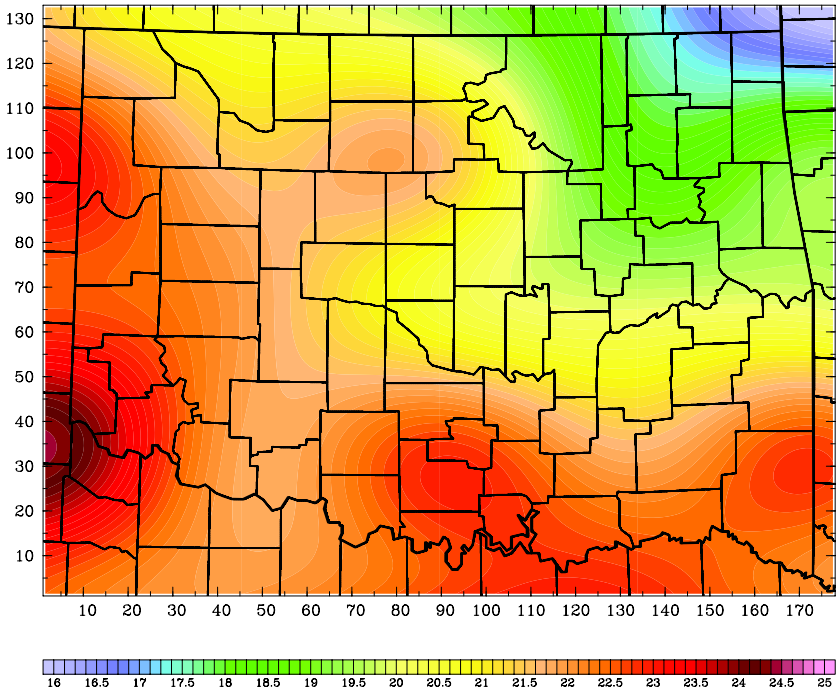

FIG. 5.4: Analytic soil temperature (°C) defined by Eq. (5.4).

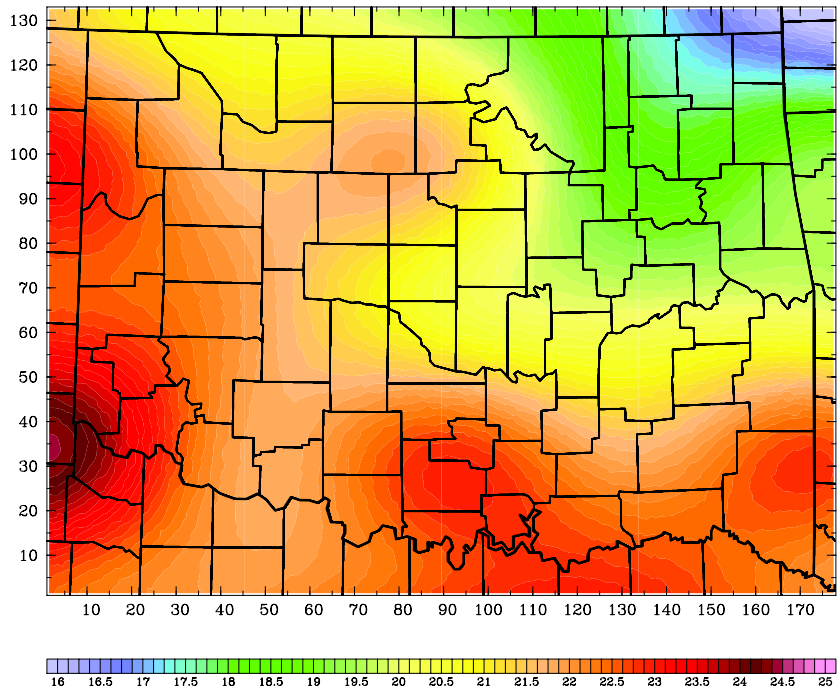

FIG. 5.5: Soil temperatures (°C) defined at random points by Eq. (5.4) and interpolated back to a regular grid using a distance-weighted average of the four random points closest to each grid point.

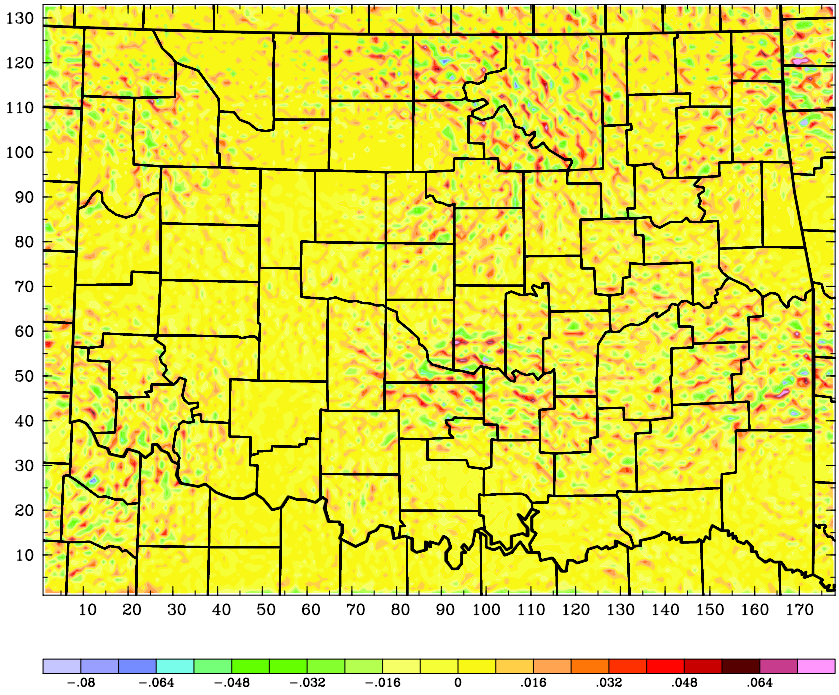

FIG. 5.6: Difference field (°C) showing the temperatures in Figure 5.5 minus the temperatures in Figure 5.4.

distribution of warm season 5-cm soil temperatures observed by the Oklahoma Mesonet across Oklahoma.

Eq. (5.4) determines the soil temperature field at $x_n \times y_n$ random points (to ensure adequate coverage) assigned by a uniform random number generator (Bratley et al. 1987, chapter 6). Figure 5.5 shows this field interpolated back to a regular grid using the distance-weighted average interpolation scheme given by Eq. (5.3). The analytic and interpolated fields (Figs. 5.4 and 5.5) are nearly identical. A difference field defined as the interpolated field minus the analytic field shows larger errors in regions with sharp gradients, though the largest percentage error is less than 0.006% (Fig. 5.6). The scheme clearly produces reasonable interpolated values. Note that this test of the distance-weighted average interpolation scheme represents the worst possible scenario. In practice, the fourth closest regular grid point to a Mesonet observation site must lie within 1.18 grid intervals. The closest four random points to a regular grid point may not necessarily fall anywhere close to the grid point. Indeed, the distance between a regular grid point and its fourth closest random

point in this test exceeded 1.18 grid intervals for nearly 37% of the regular grid points.

The benefits of this interpolation scheme outweigh its unconventional nature. Alternatively, a bilinear interpolation scheme for an MM5 forecast verification necessitates ignoring comparisons with soil variables from observation sites near water bodies. This reduces the data available for the verification and is a particular problem on high-resolution model domain four, which lies over only a small portion of Oklahoma. In contrast, the distance-weighted average interpolation scheme retains these observations while maintaining the integrity of the forecast variables.

5.5 Comparative tests with differing initial conditions

5.5.1 Selection of case studies

To aid in a selection of several case studies, MM5 is used to compute daily 48-hour forecasts on the larger three domains during the period 1 April–30 September 2004. For these preliminary forecasts, 1200 UTC Eta analyses provide initial conditions only for domain one, from which each nested domain receives interpolated initial and boundary conditions. Results from 9-, 24-, and 33-hour forecasts of 2-m air temperature and mixing ratio on domain three for each model run, when compared with corresponding Mesonet observations, highlight several good and bad forecasts for further study. Based on these comparisons, case studies selected for further study include 48-hour model forecasts initialized at 1200 UTC on 3 May, 20 July, 1 August, and 3 September 2004. There are strong similarities between observations and modeled 2-meter air temperature and mixing ratio fields at all three forecast times for forecasts initialized on 3 May and 20 July 2004. Clear conditions prevailed over most of Oklahoma for 3–4 May, with partial cloudiness entering the northern portion of the state during the morning of 4 May and dissipating by late afternoon. Clear skies dominated for 20–21 July, though an upper-level low initiated thunderstorms near 2000 UTC 20 July 2004 over the Texas Panhandle. Cirrus anvils from these storms over-

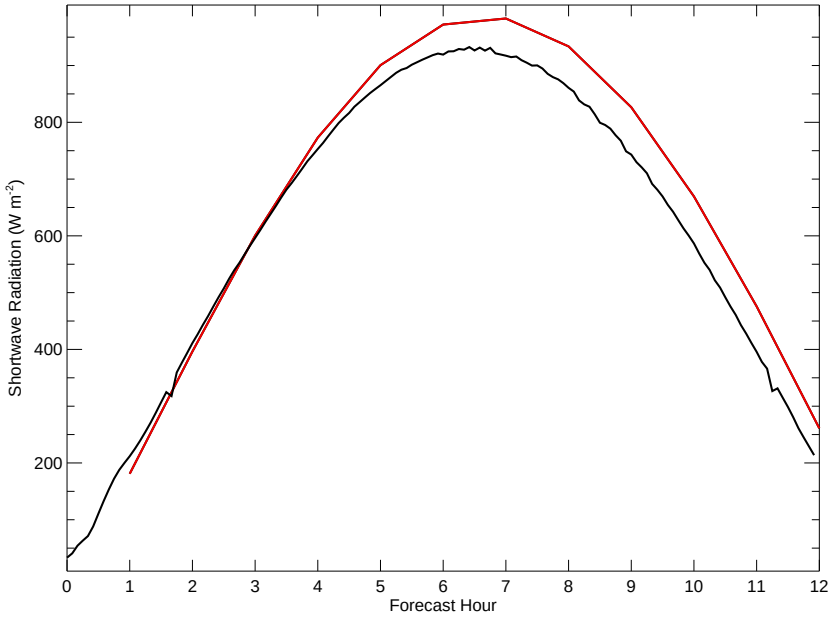

FIG. 5.7: Modeled (red) and observed (black) downward shortwave radiation at Foraker, Oklahoma for the forecast initialized at 1200 UTC 1 August 2004 using the default value of $\sigma_x = 1.0$ and the default radiation calculation frequency of 30 min.

spread the western third of Oklahoma overnight. A warm bias of several degrees Celsius exists in afternoon forecasts of 2-meter air temperature initialized on 1 August 2004, with an early morning cool bias. A dry bias also is present for the forecast period. A ridge of high pressure centered over Oklahoma prevented cloud development for 1–3 August. Results of the 3 September 2004 preliminary MM5 run exemplify a bad forecast with a warm bias for 2-meter air temperatures and a poor representation of the moisture field compared with observations. Shallow cumulus fields developed over much of Oklahoma during the afternoons of 3 and 4 September. In all cases, no strong synoptic features passed over the body of Oklahoma. This assortment of fair and poor preliminary forecasts under synoptically quiescent conditions provides several ideal cases for studying the impact of improved initial conditions on forecasts of radiative fluxes and maximizes the potential for isolating the effect of changes to the land surface model on near-surface atmospheric variables.

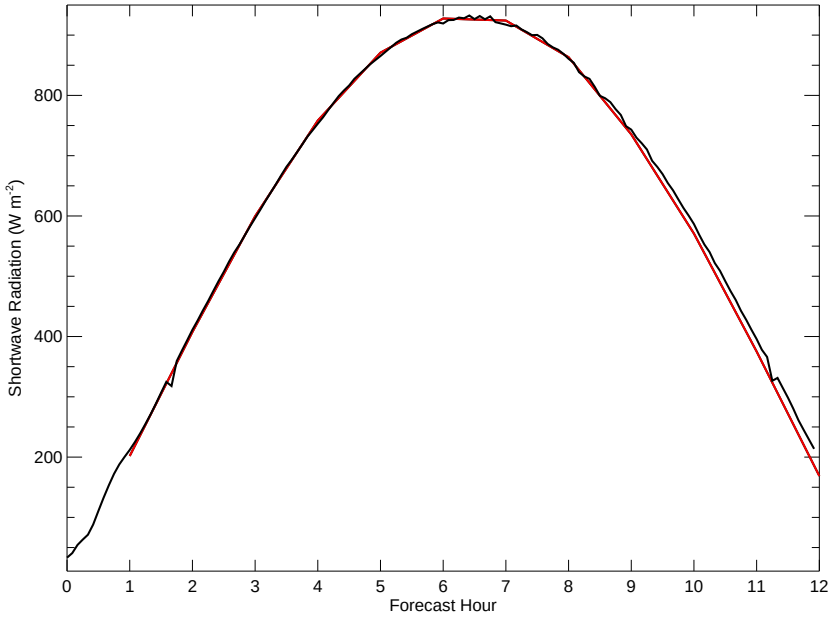

FIG. 5.8: Modeled (red) and observed (black) downward shortwave radiation at Foraker, Oklahoma for the forecast initialized at 1200 UTC 1 August 2004 with $\sigma_x = 1.6$ and a radiation calculation frequency of 5 min.

5.5.2 Solar radiation tuning

Model simulations on a primary domain over the body of Oklahoma reveal a systematic overestimation in surface downward shortwave radiation from the Dudhia (1989) solar radiation parameterization within MM5. Other studies report similar findings and note the strong dependence of the scheme on aerosol optical depth (e.g., Zamora et al. 2003, 2005). Marshall et al. (2003) discuss similar problems with Eta model simulations. The large positive bias in incoming solar radiation may result from the parameterization's neglect of stratospheric ozone and the treatment of scattering and absorption. Rather than explicitly accounting for ozone absorption, Rayleigh scattering, aerosol absorption, and upward aerosol scattering, the Dudhia shortwave parameterization combines these physical attenuation processes into a single scattering parameter, scaled by a constant, σ_x. Forecasts of incoming solar radiation using the default value of $\sigma_x = 1.0$ result in solar radiation overestimates exceeding 50 W m^{-2} under cloudless skies (Fig. 5.7). The excess radiation from the solar parameterization alone overwhelms the effect of improvements to initial land-surface

conditions. Additionally, the default frequency of 30 minutes for calls to the atmospheric radiation scheme causes the phase difference apparent in Figure 5.7.

A modification to the atmospheric radiation calculation frequency to five minutes properly shifts the downward shortwave radiation curve such that it becomes in-phase with the observations, though the magnitude of the radiation remains unchanged. In lieu of adding explicit formulations for ozone and aerosol absorption and Rayleigh and upward aerosol scattering, as suggested by Zamora et al. (2003), tuned σ_x values for each case study provide the best match to solar radiation observations from all nine OASIS super sites within the body of Oklahoma (Fig. 5.8). The best overall σ_x values are 1.4, 1.7, 1.6, and 1.4 for forecasts initialized at 1200 UTC on 3 May, 20 July, 1 August, and 3 September 2004, respectively. These tuned σ_x values remain constant in each call to the solar radiation subroutine through the entire forecast period.

5.5.3 Initial conditions

To explore the importance of the land surface on the model forecasts, four different sets of initial conditions for the soil and land surface initialize MM5. The control MM5 (CTRL) uses a 0.15° × 0.15° climatological σ_f, assumed valid in the middle of each of the 12 months of the year, produced from a five-year climatology of NDVI observations (Gutman and Ignatov 1998). Values of σ_f at each grid point are interpolated temporally according to the day of the month for each model run. The model also assumes a constant LAI (the default is set to 4.0) based on categorical USGS land usage data, regardless of the season or location. Eta model analyses provide initial soil temperature and soil moisture conditions.

The second MM5 (MM5VEG) initial condition includes the 1-km resolution σ_f and LAI observations derived from a 15- or 16-day NDVI composite. Composite windows span the periods 16–30 April, 1–15 July, 16–31 July, and 16–31 August 2004 for the 3 May, 20 July, 1 August, and 3 September 2004 case studies, respectively. All four nested

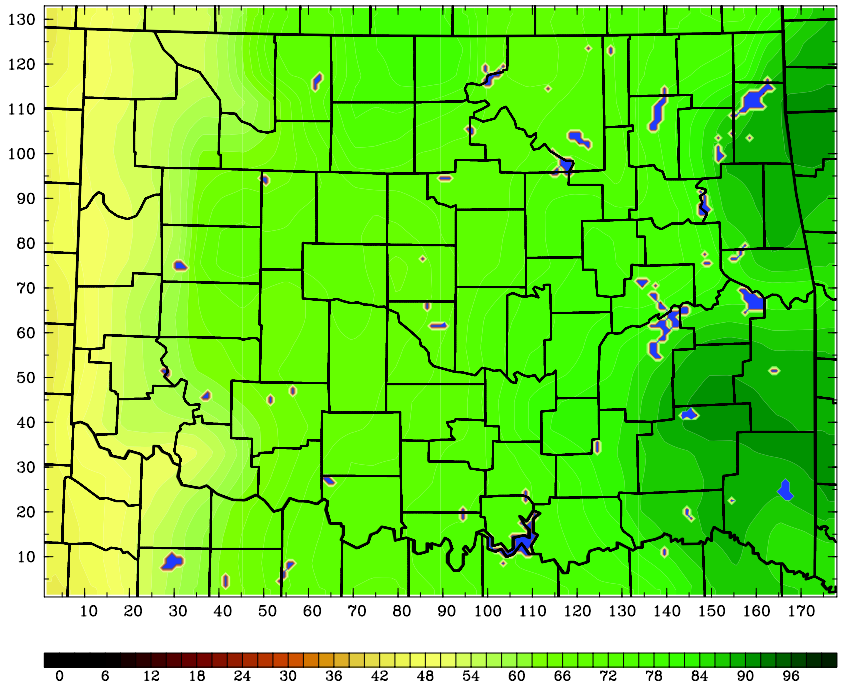

FIG. 5.9: 3 May 2004 fractional vegetation coverage (percentage) for domain three based on a 5-year climatology. Blue areas indicate water bodies.

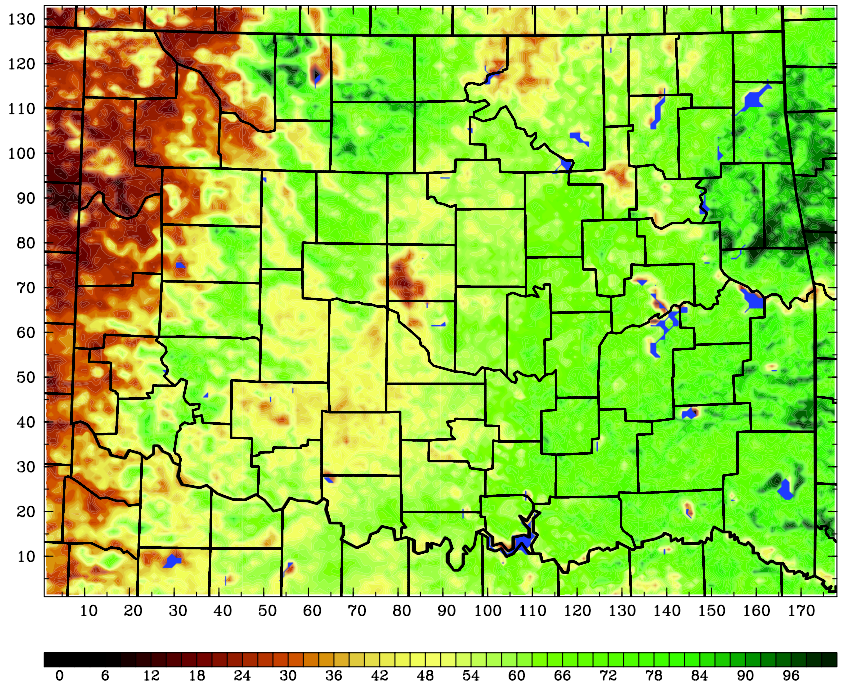

FIG. 5.10: Fractional vegetation coverage (percentage) for domain three calculated from a maximum NDVI composite over the period 16–30 April 2004. Blue areas indicate water bodies.

model domains receive interpolated vegetation data. A final vegetation index value at the center of each grid square derives from an arithmetic mean of all pixels within each grid square. Modifications to MM5 allow the model to accept an array of LAI values for all domains. Satellite-derived σ_f and LAI data cover a swath similar to the area of domain two. LAI and σ_f values for points outside the area of the satellite pass in domains one and two remain set to a constant 4.0 and climatology, respectively.

Since NDVI is a function of both σ_f and LAI, Gutman and Ignatov (1998) conclude that both σ_f and LAI cannot be regarded as two independent pieces of information and should not be used together in the same land-surface parameterization. However, inadequacies inherent in a modern numerical weather prediction model vastly outweigh the importance of errors introduced by the dual specification of vegetation parameters from a single NDVI observation. Additionally, it is advantageous to provide the model with as much information as possible by adding a second variable rather than using a constant LAI. Adding spatial variability to the LAI field reduces errors in surface heat fluxes (e.g., Li and Avissar 1994). Moreover, these two vegetation indexes are relatively independent within the land surface model selected for this study. In the Noah LSM, canopy resistance is a function of LAI, while the partitioning of latent heat between bare soil and vegetation relies on σ_f. Together, LAI and σ_f specify the total canopy transpiration.

The third MM5 (MM5SOIL) initial condition uses Mesonet soil data, but climatology for the vegetation. A two-pass Barnes analysis (Barnes 1973) generates gridded fields of observed soil moisture at 5, 25, and 60 cm depth and soil temperature at depths of 5, 10, and 30 cm beneath native vegetation on MM5 model domains three and four. Optimized analyses produce a large response for mesoscale waves, but damp unrealistic high-frequency waves across Oklahoma (Maddox 1980). Soil temperature and moisture observations replace Eta model analyses of soil fields on nested domains three and four in two of the four initial soil temperature layers and three of the four initial soil moisture layers in MM5. The Mesonet observations of soil temperature at a depth of 5 cm replace the initial model soil temperature in the 0–10 cm layer. The 10–40 cm model layer is the second layer in the

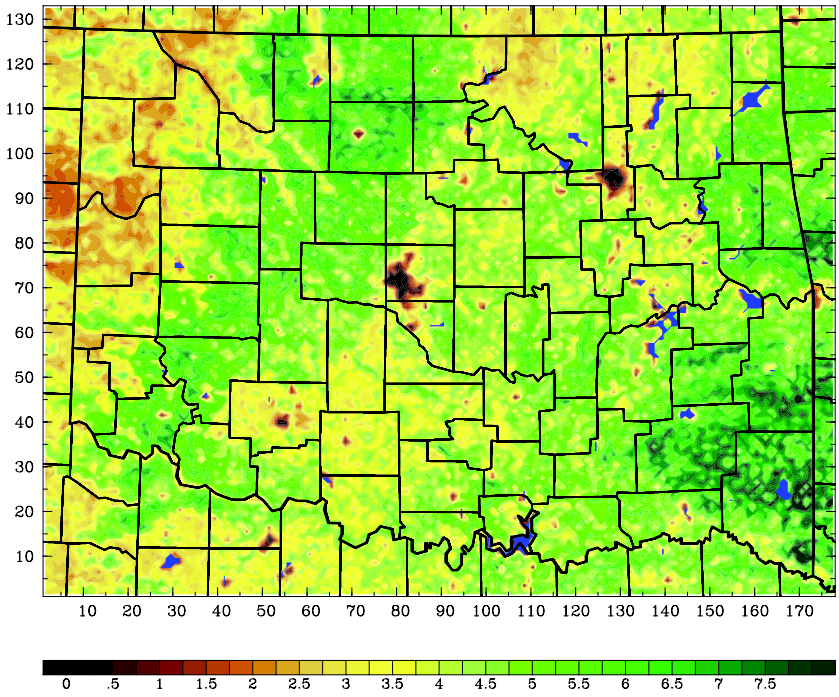

FIG. 5.11: Leaf area index (dimensionless) for domain three calculated from a maximum NDVI composite over the period 16–30 April 2004. Blue areas indicate water bodies.

soil model. To maintain consistency with soil temperatures in the deeper model layers, a cubic spline interpolation supplies a fit between all three observed soil temperatures and the initial model soil temperature in the 40–100 cm layer, with the assumption that the 40–100 cm layer temperature is valid at a depth of 70 cm. The interpolated value at a depth of 25 cm replaces the initial MM5 soil temperature in the 10–40 cm layer. The observed volumetric water content at a depth of 5 cm from the Mesonet replaces the initial soil moisture field in the 0–10 cm model layer. The 25 cm volumetric water content measurements replace the initial soil moisture field in the 10–40 cm model layer, and the 60 cm volumetric water content measurements replace the initial soil moisture field in the 40–100 cm model layer. The initial soil temperature field in the 40–100 cm layer and both the soil temperature and moisture fields in the 100–200 cm layer remain unchanged from the interpolated Eta analyses for domains three and four. All soil fields for domains one and two also remain unchanged from the interpolated Eta analyses.

The fourth MM5 (MM5VEGSOIL) initial condition uses the 1-km AVHRR-derived σ_f

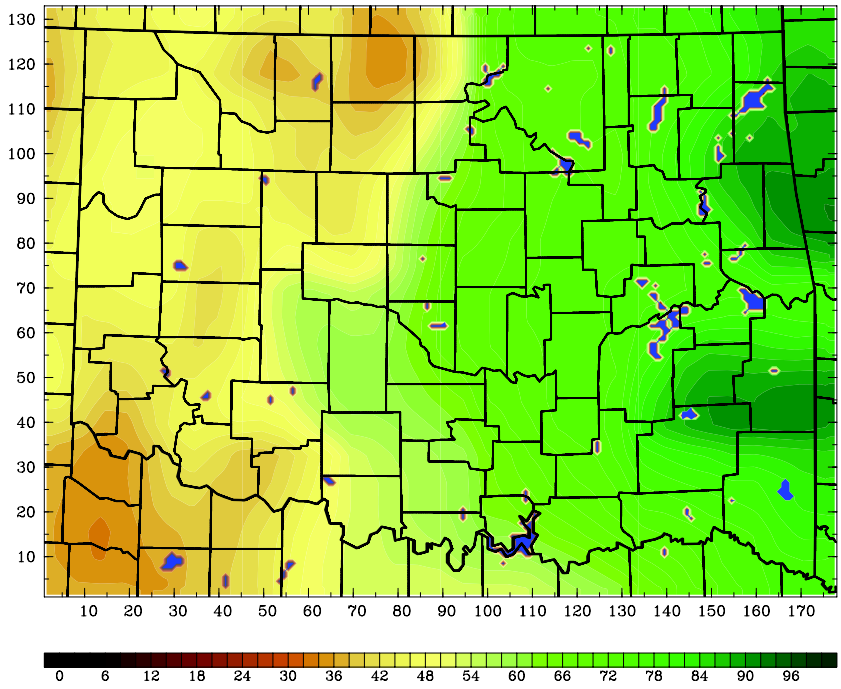

FIG. 5.12: 3 September 2004 fractional vegetation coverage (percentage) for domain three based on a 5-year climatology. Blue areas indicate water bodies.

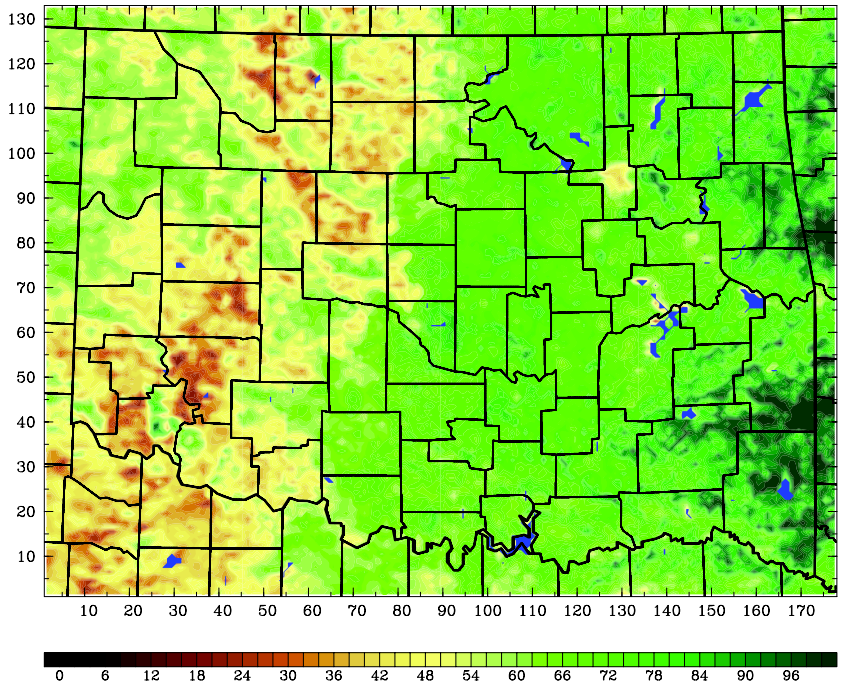

FIG. 5.13: Fractional vegetation coverage (percentage) for domain three calculated from a maximum NDVI composite over the period 16–31 August 2004. Blue areas indicate water bodies.

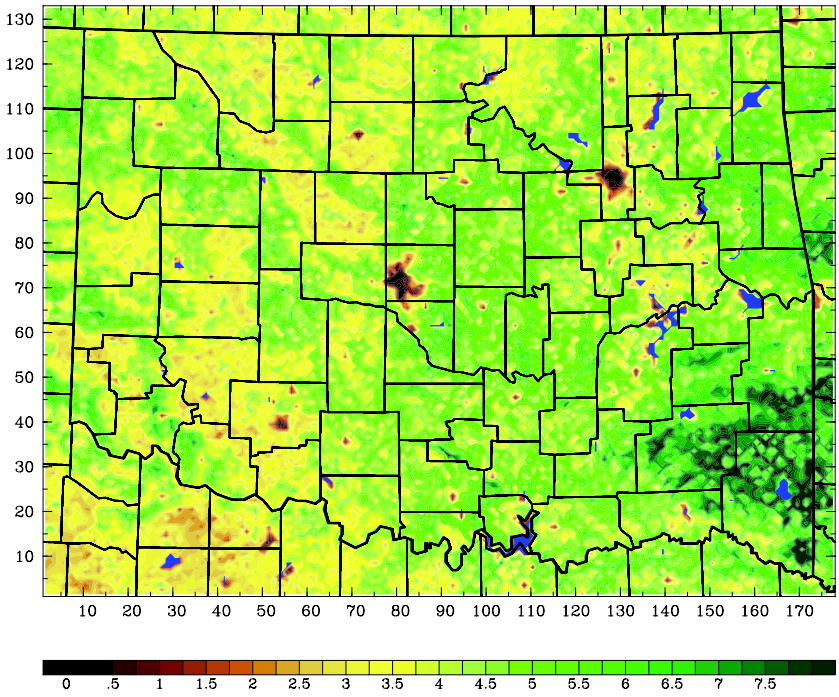

FIG. 5.14: Leaf area index (dimensionless) for domain three calculated from a maximum NDVI composite over the period 16–31 August 2004. Blue areas indicate water bodies.

and LAI values along with the soil data from the Oklahoma Mesonet. This initial condition provides the most accurate specification of the land surface and soil conditions for the model.

A comparison between initial σ_f for the 3 May 2004 CTRL and MM5VEG forecasts illustrates the stark contrast between climatological (Fig. 5.9) and observed σ_f derived from a 15-day maximum NDVI composite over the period 16–30 April 2004 (Fig. 5.10). The LAI for the same period ranges from just above 0.0 in urban areas to greater than 8.0 in the forests of southeastern Oklahoma (Fig. 5.11). The fractional vegetation coverage climatology on 3 September 2004 (Fig. 5.12) more closely resembles the σ_f observations over the period 16–31 August 2004 that initialize the 3 September MM5VEG simulation (Fig. 5.13). However, the LAI from the same observing period deviates substantially from a constant 4.0 (Fig. 5.14). When MM5 employs a constant LAI in control forecasts, it clearly loses information about the vegetation biomass covering the actual land surface.

Initial soil temperature conditions in the 0–10 cm layer in the 1200 UTC 3 May 2004

SOILMM5 simulation compared with initial conditions in the CTRL simulation differ by several Kelvin in some locations and show that the Eta analysis for this soil layer is too cool (Figs. 5.15 and 5.16). Errors are generally smaller in the 10–40 cm soil temperature layer (Figs. 5.17 and 5.18). Soil moisture initial conditions in the CTRL simulations suffer from the dry bias in Eta analyses. For 1200 UTC 3 May 2004, initial soil moisture conditions in the 0–10 cm layer of the CTRL simulation contain a dry bias compared with the MM5SOIL simulation (Figs. 5.19 and 5.20). Maximum differences are on the order of 50% of the range of observed values across the domain. This dry bias persists into deeper soil layers (Figs. 5.21 and 5.22). Initial conditions for other case studies exhibit comparable differences.

The detailed spatial resolution in CTRL initial soil conditions is not present within the Eta analyses used for the initial conditions, but results from the way the interpolation scheme within MM5 accounts for topography, land use, and soil type. Since the insertion of observed soil fields into MM5 initial conditions occurs after the interpolation of Eta analyses onto the model domains, this variability does not appear in MM5SOIL and MM5VEGSOIL initial soil conditions. Lastly, rather than interpolating initial conditions from the mother domain to each nested domain, Eta analyses provide initial conditions directly on the four MM5 domains for each case study. The model, with each of the four MM5 initial land surface and soil conditions, uses the same atmospheric initial and boundary conditions to produce 48-hour forecasts. When compared with Mesonet observations, results from all four types of forecasts reveal the impact of improved land surface initial conditions.

5.5.4 Results

At Oklahoma Mesonet sites, thermistors and relative humidity probes are installed at 1.5 m AGL with anemometers at 10 m AGL. Verification of MM5 model results compares the 1.5-m observations of air temperature and mixing ratio with the 2-m model output.

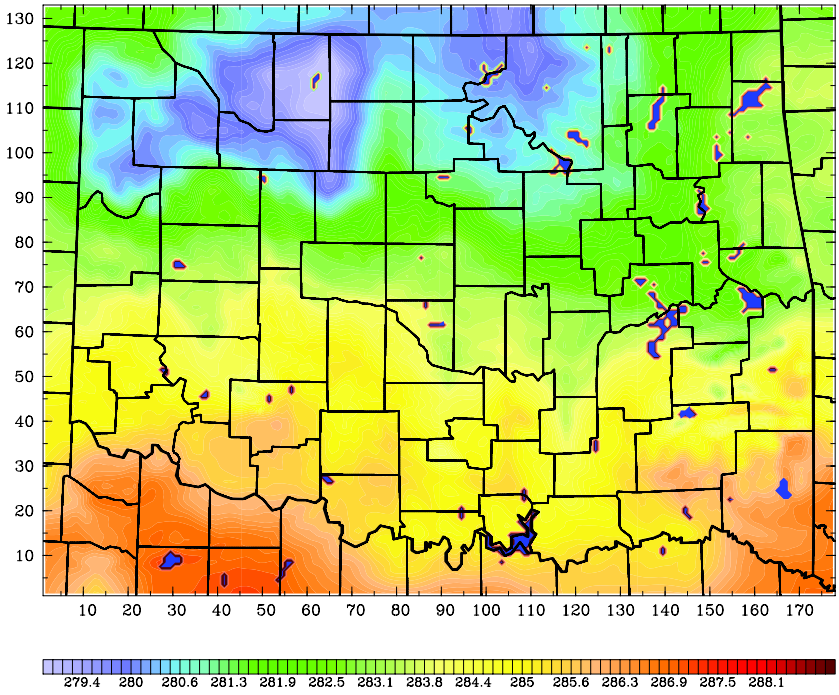

FIG. 5.15: Initial soil temperature (K) in the 0–10 cm layer for domain three of the 1200 UTC 3 May 2004 CTRL simulation.

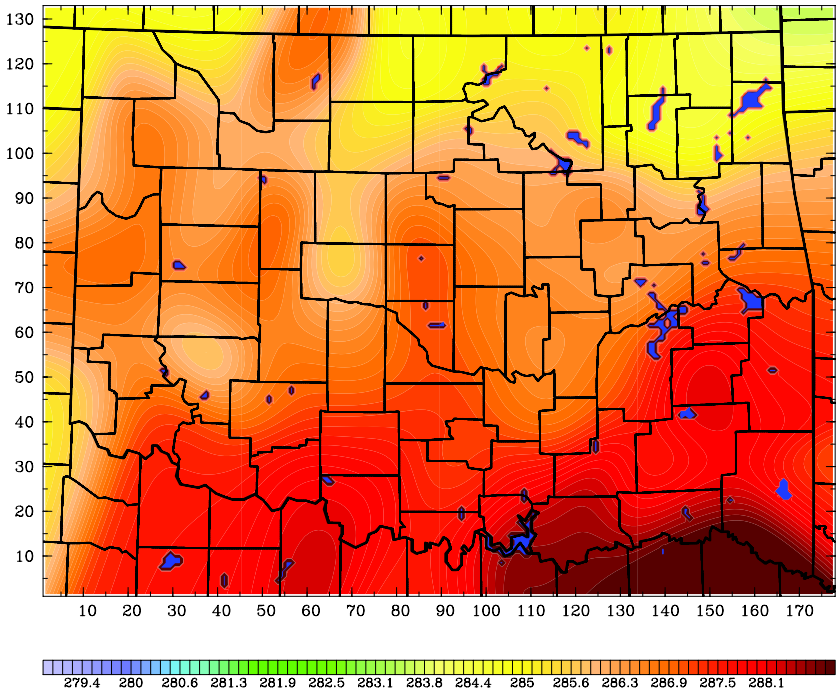

FIG. 5.16: Initial soil temperature (K) in the 0–10 cm layer analyzed from Oklahoma Mesonet observations for domain three of the 1200 UTC 3 May 2004 MM5SOIL simulation.

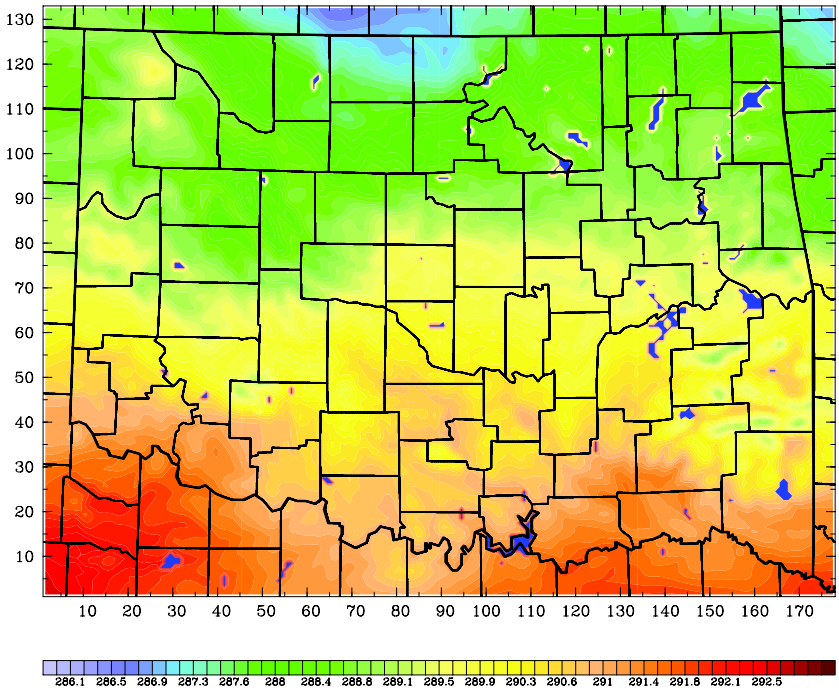

FIG. 5.17: Initial soil temperature (K) in the 10–40 cm layer for domain three of the 1200 UTC 3 May 2004 CTRL simulation.

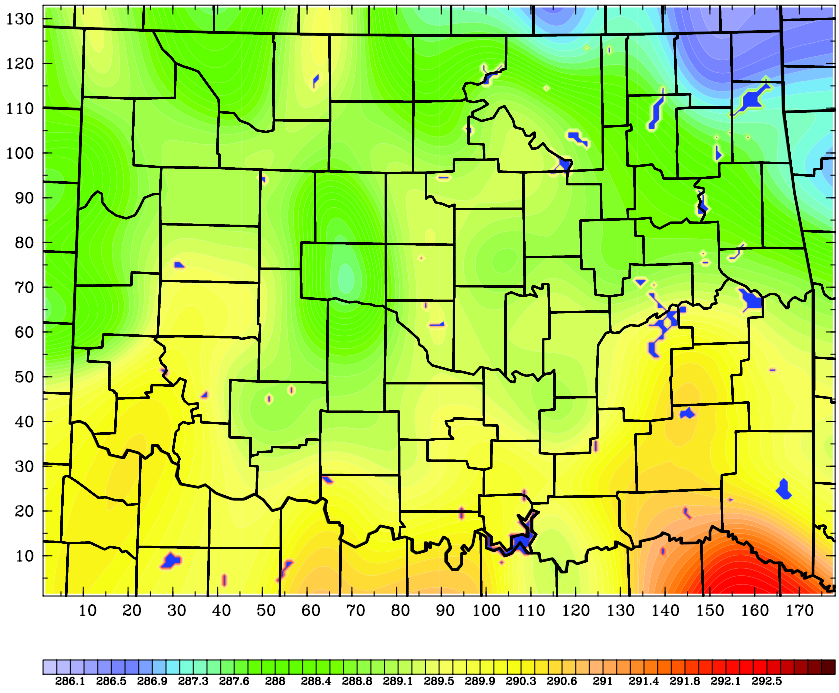

FIG. 5.18: Initial soil temperature (K) in the 10–40 cm layer analyzed from Oklahoma Mesonet observations for domain three of the 1200 UTC 3 May 2004 MM5SOIL simulation.

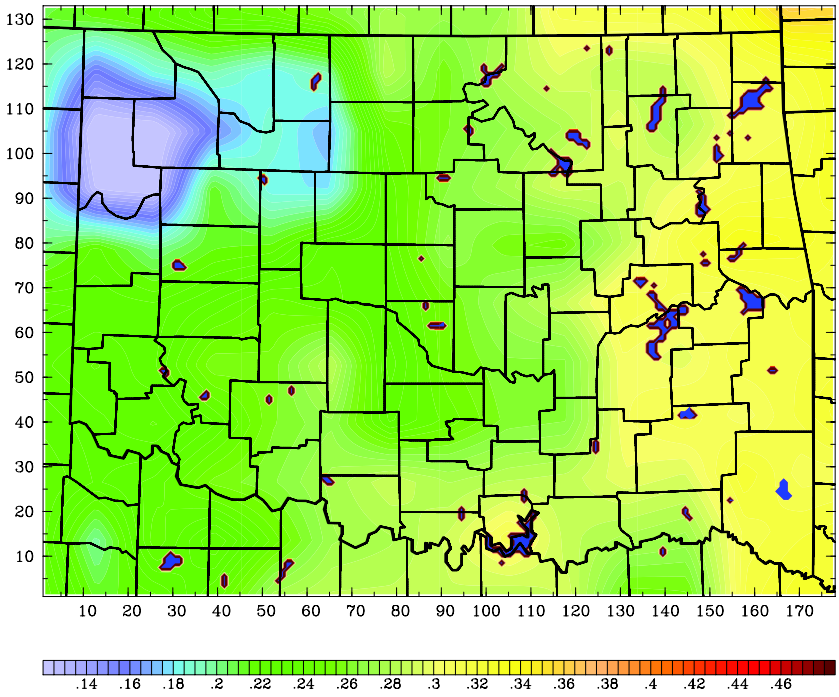

FIG. 5.19: Initial soil moisture (m^3 m^{-3}) in the 0–10 cm layer for domain three of the 1200 UTC 3 May 2004 CTRL simulation.

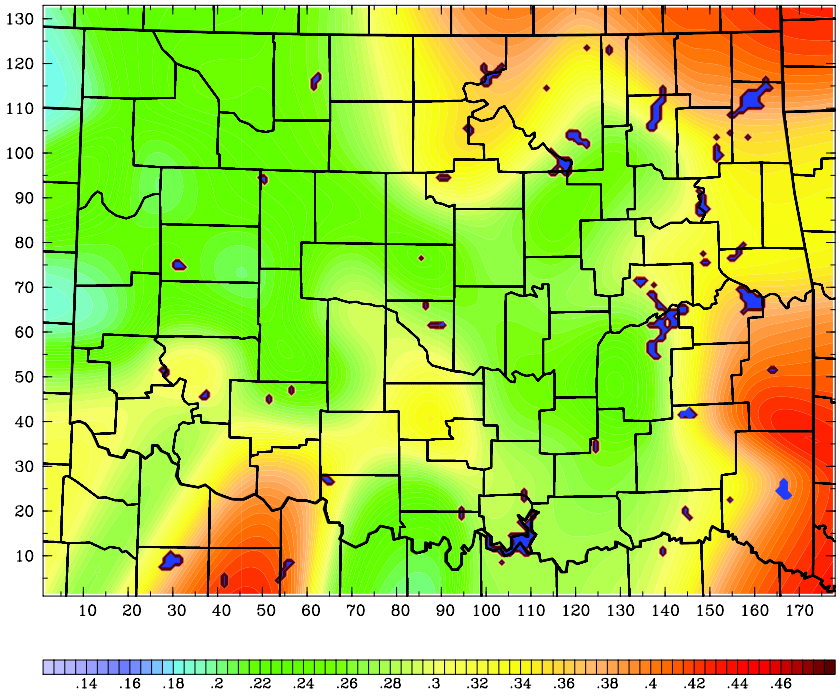

FIG. 5.20: Initial soil moisture (m^3 m^{-3}) in the 0–10 cm layer analyzed from Oklahoma Mesonet observations for domain three of the 1200 UTC 3 May 2004 MM5SOIL simulation.

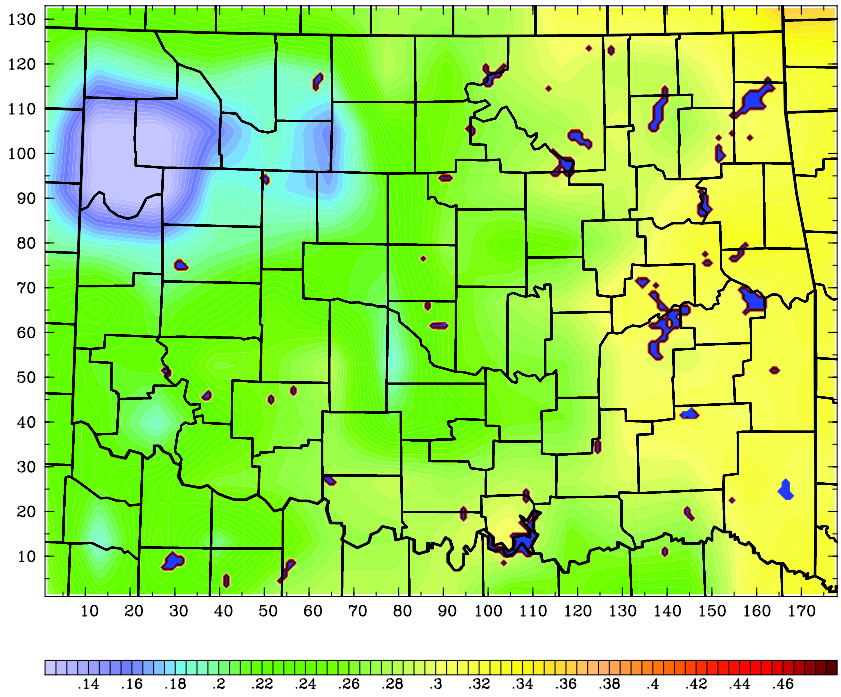

FIG. 5.21: Initial soil moisture (m^3 m^{-3}) in the 10–40 cm layer for domain three of the 1200 UTC 3 May 2004 CTRL simulation.

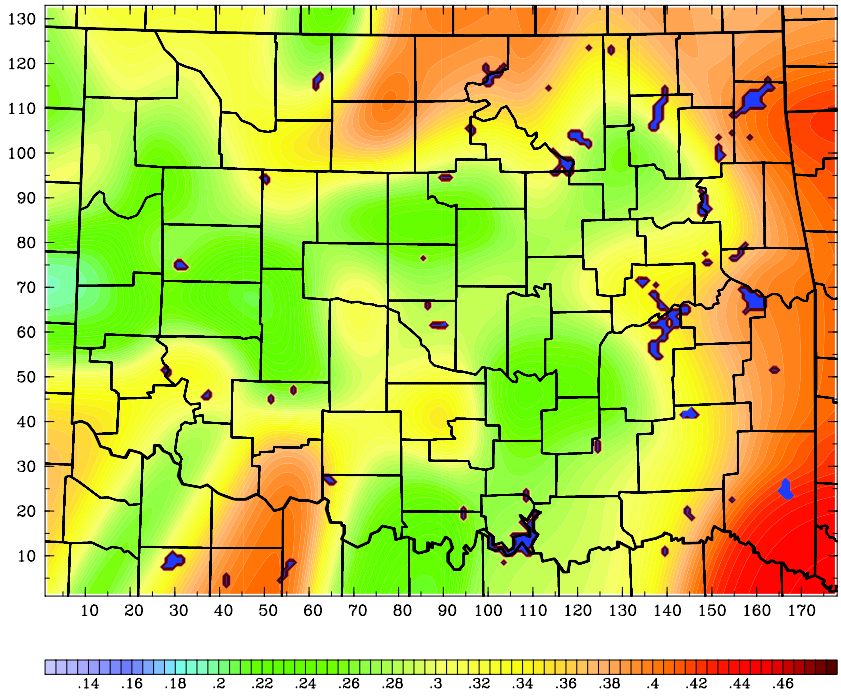

FIG. 5.22: Initial soil moisture (m^3 m^{-3}) in the 10–40 cm layer analyzed from Oklahoma Mesonet observations for domain three of the 1200 UTC 3 May 2004 MM5SOIL simulation.

Root-mean squared error, mean absolute error, and bias (Wilks 2006, 278–280) for forecasts of 2-m air temperature and mixing ratio, 10-meter wind magnitude, soil temperature and moisture, and surface energy fluxes for all of domain three illustrate the performance of each forecast. Errors for each forecast rely on observations from no fewer than the number of Mesonet sites listed in Table 5.1 for each output variable. Forecast errors for domain three are very similar to and representative of the forecast errors for domain four.

MM5SOIL generally produced the best results, reducing bias errors for temperature forecasts during the day (Fig. 5.23) and for mixing ratio forecasts at nearly all forecast hours (Fig. 5.24). The unrealistic spike in mixing ratio errors at about 0000 UTC (forecast hours 12 and 36) results from the interpolation procedure in the MRF PBL scheme for determining the 2-m mixing ratio. As discussed in section 5.3, the 2-m mixing ratio is strongly dependent upon the Obukhov length. The upward portion of the mixing ratio spike is a consequence of this sensitivity. During the PBL regime transition from free convection ($R_B < 0$) to stable conditions ($R_B \geq 0.2$) near sunset, the Obukhov length drops sharply from positive to negative values through a period of damped mechanical turbulence ($0.0 < R_B < 0.2$). The decrease in Obukhov length reduces the ratio of $\psi_w(2\text{m})$ to $\psi_w(\sigma)$ in Eq. (5.2), giving more weight to w_{s_g}. Since the saturation mixing ratio at the ground temperature is much larger than the mixing ratio slightly above the ground, the extra weight

TABLE 5.1: Minimum number of observations from the Oklahoma Mesonet at any forecast hour used in calculating error statistics for MM5 simulations for each case study.

	3 May 2004	20 Jul 2004	1 Aug 2004	3 Sep 2004
2-m temperature	108	109	106	111
2-m mixing ratio	107	109	106	111
10-m wind speed	109	110	109	111
5-cm soil temperature	95	98	98	97
5-cm soil moisture	79	81	76	82
25-cm soil moisture	69	85	78	83
60-cm soil moisture	42	53	52	58
Sensible heat flux	9	9	8	9
Surface downward shortwave	9	9	8	9
Surface downward longwave	9	9	8	9
Ground heat flux	72	64	65	65
Precipitation	107	110	108	112
NR-Lite net radiation	76	72	72	72

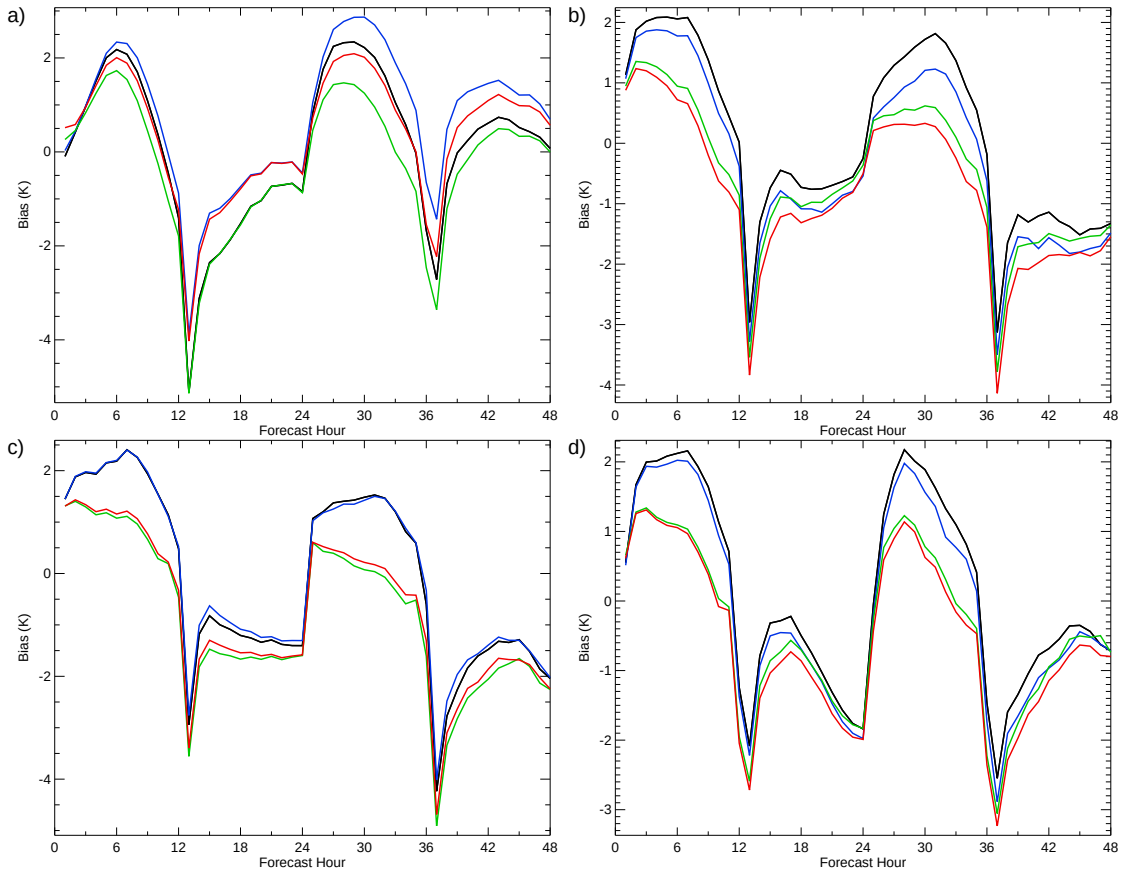

FIG. 5.23: 2-m air temperature bias (K) after comparison with Oklahoma Mesonet observations for CTRL (black), MM5SOIL (green), MM5VEG (blue), and MM5VEGSOIL (red) domain three simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004.

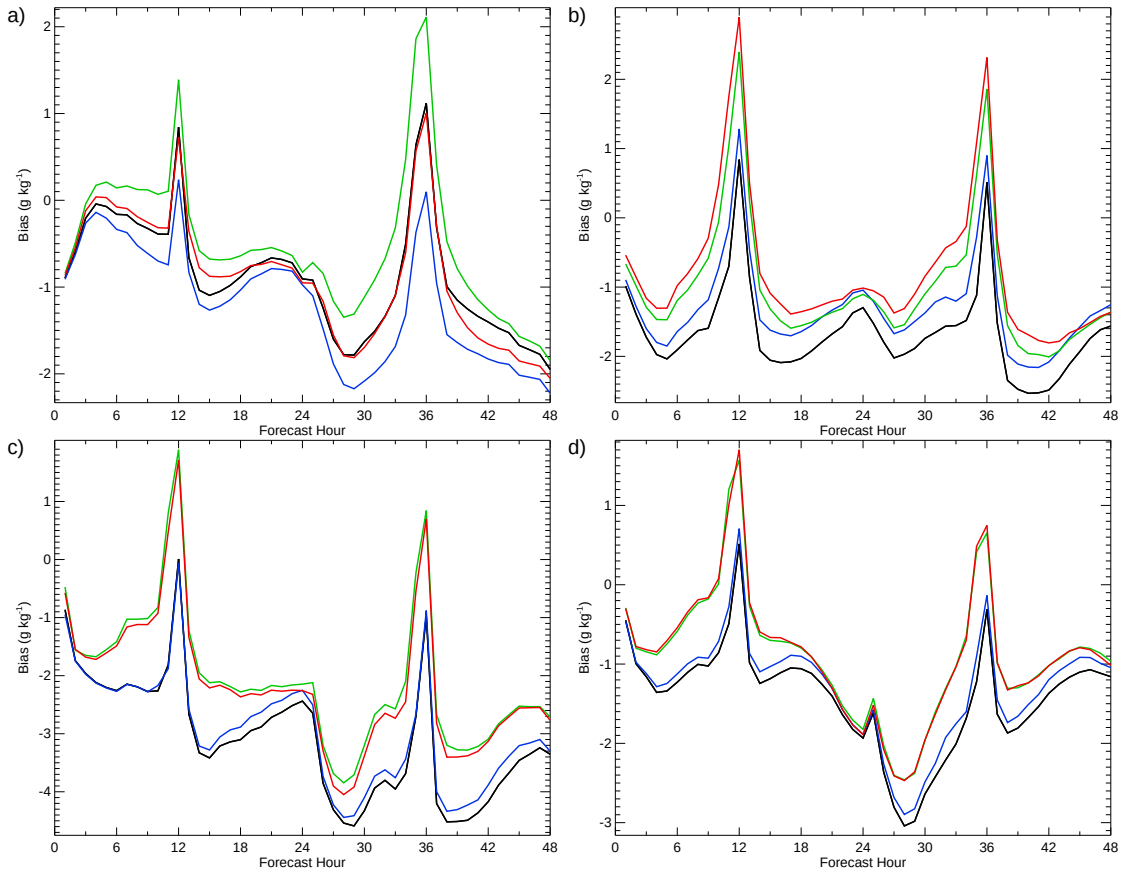

FIG. 5.24: 2-m mixing ratio bias (g kg^{-1}) after comparison with Oklahoma Mesonet observations for CTRL (black), MM5SOIL (green), MM5VEG (blue), and MM5VEGSOIL (red) domain three simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004.

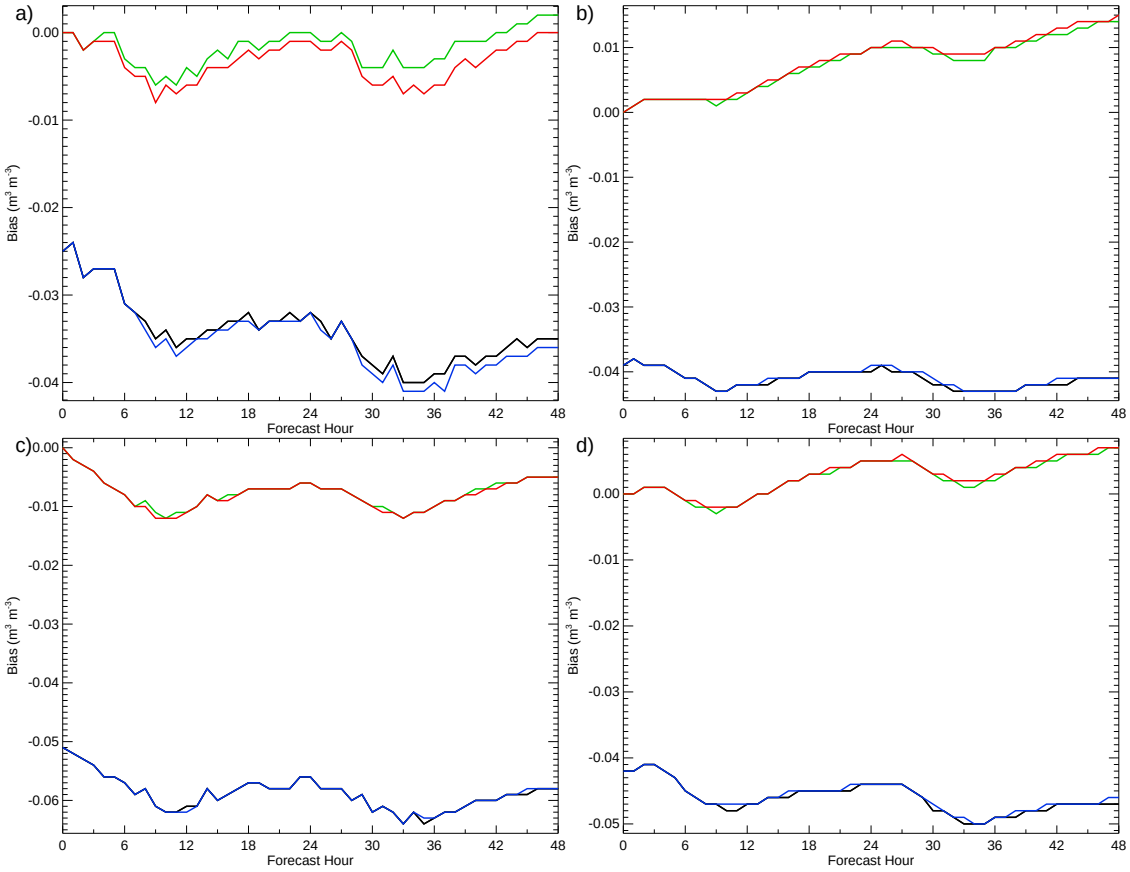

FIG. 5.25: Soil moisture bias ($m^3\ m^{-3}$) in the 0–10 cm model layer after comparison with Oklahoma Mesonet observations at 5-cm depth for CTRL (black), MM5SOIL (green), MM5VEG (blue), and MM5VEGSOIL (red) domain three simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004.

given to w_{s_g} causes the interpolated 2-m mixing ratio to increase. At the peak in the mixing ratio spike, $\psi_w(\sigma)$ reaches an artificially imposed minimum negative value, discontinuing a reduction in the ratio of $\psi_w(2\text{m})$ to $\psi_w(\sigma)$. At this point, the drop in the 2-m mixing ratio is driven solely by the reduction in w_{s_g} as the ground temperature decreases due to the reversal of the sign of the sensible heat flux. A stronger spike occurs in the MM5SOIL and MM5VEGSOIL simulations because mixing ratios are higher throughout these simulations, primarily due to the availability of more soil moisture for evaporation (Fig. 5.25). A similar reduction of the ratio of $\psi_T(2\text{m})$ to $\psi_T(\sigma)$ in Eq. (5.1) causes air temperatures to drop too quickly during the PBL transition to stable conditions. Errors for all 10-m wind magnitude forecasts differed only slightly, though each simulation failed to capture the diurnal cycle of increased wind speeds during the daytime and reduced wind speeds at

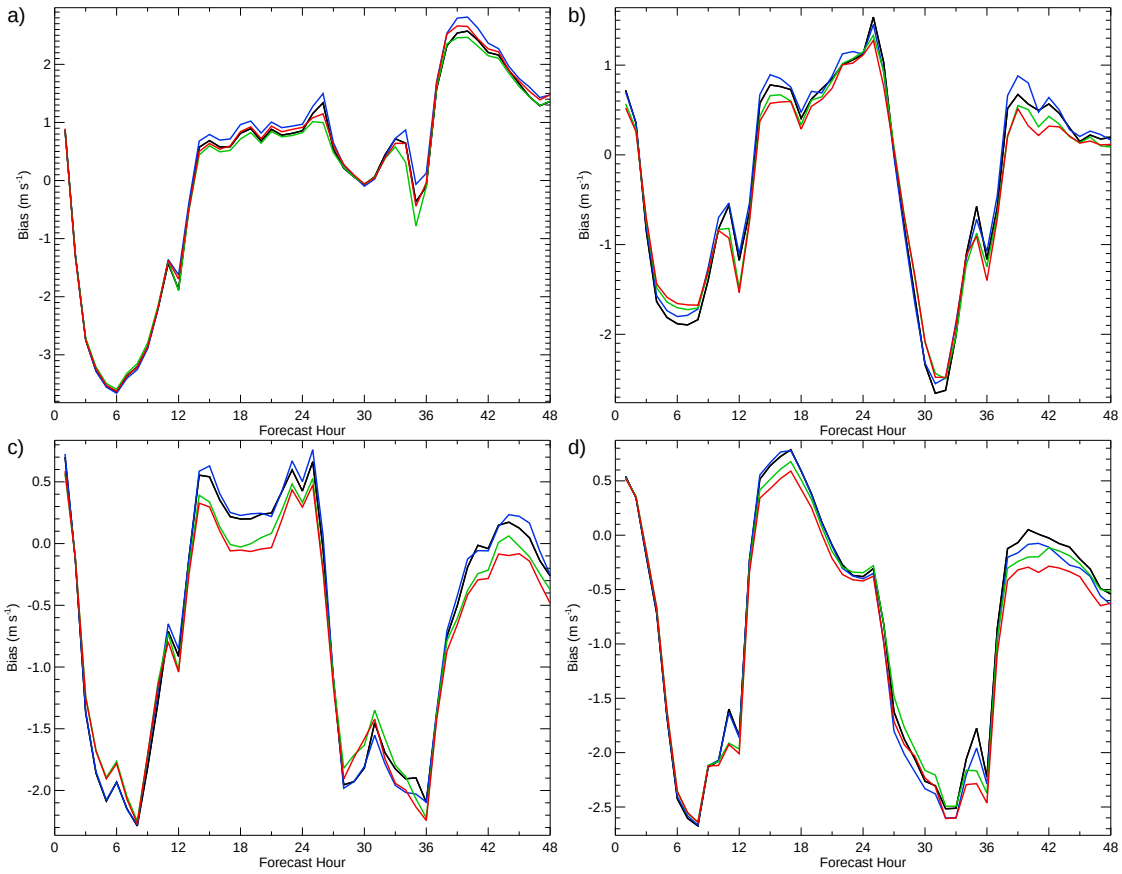

FIG. 5.26: 10-m wind speed bias (m s^{-1}) after comparison with Oklahoma Mesonet observations for CTRL (black), MM5SOIL (green), MM5VEG (blue), and MM5VEGSOIL (red) domain three simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004.

night (Fig. 5.26) as also seen in Zhang and Zheng (2004). These errors may also arise from parameterization choices within the MRF PBL scheme.

Soil temperature and soil moisture values were clearly more accurate in the MM5SOIL and MMVEGSOIL simulations. In most cases, however, the model strongly overestimates daytime soil temperatures and underestimates soil temperatures at night (Fig. 5.27), despite driving the surface energy budget with solar radiation tuned to match observations. The overall MM5SOIL forecast and MM5VEGSOIL forecast errors for soil temperature and moisture do not drift appreciably back to control forecast errors during the forecast period.

Due to the marked discrepancy between observed and climatological σ_f, the observed departure from a constant LAI in MM5VEG, and increased initial soil moisture in MM5SOIL, large differences exist between latent heat fluxes for each forecast type (Fig.

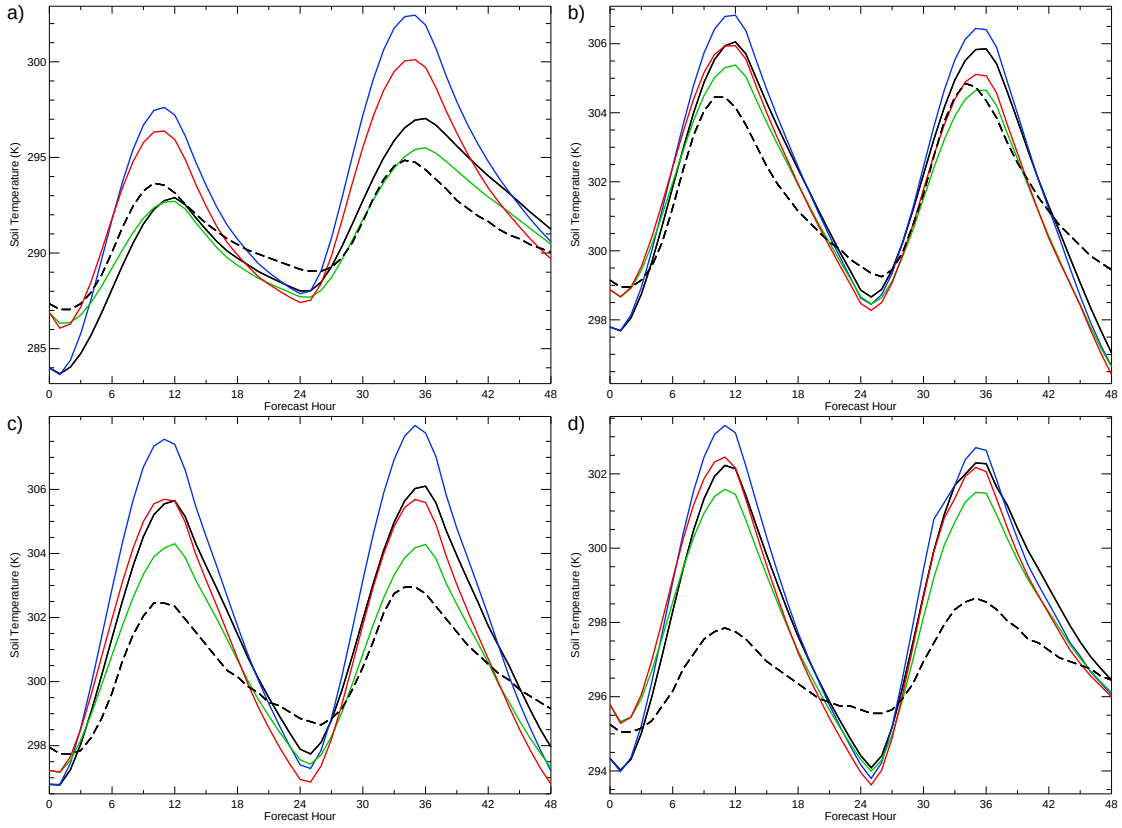

FIG. 5.27: Soil temperature (K) at Norman, Oklahoma in the 0–10 cm model layer for CTRL (black), MM5SOIL (green), MM5VEG (blue), and MM5VEGSOIL (red) domain four simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004 compared with Oklahoma Mesonet observations at 5-cm depth (dashed).

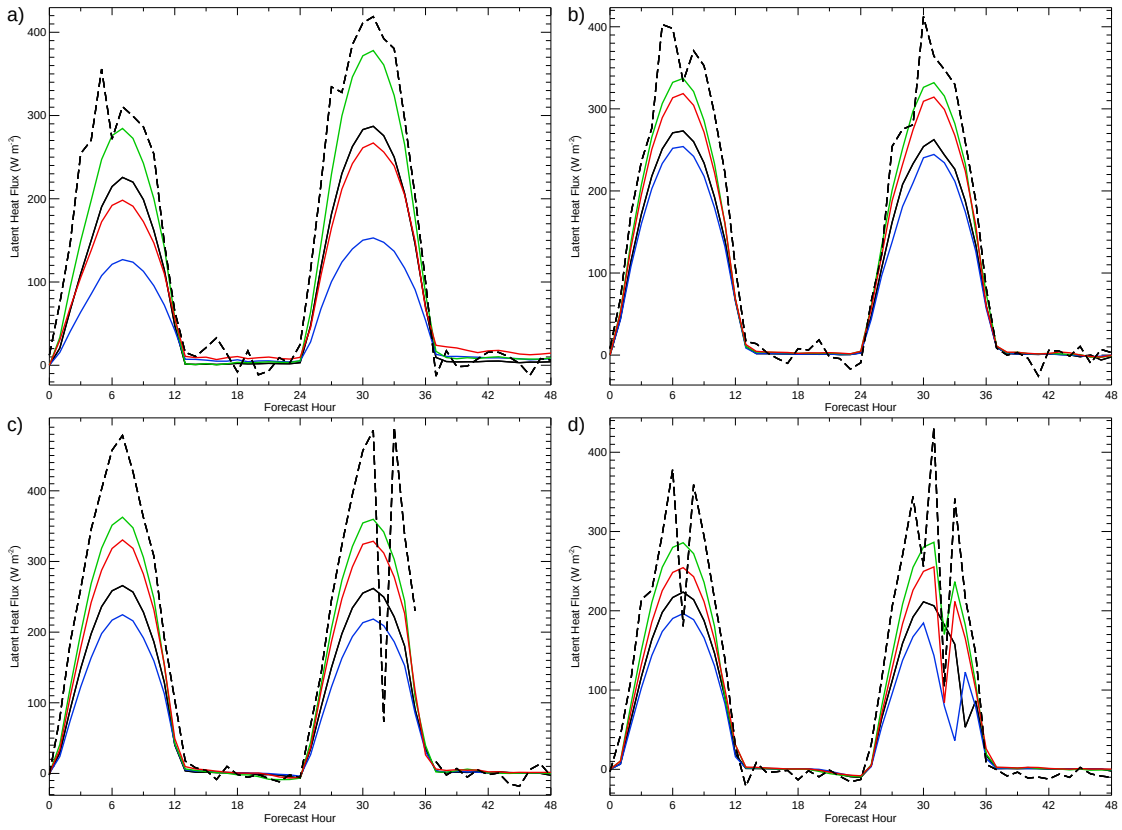

FIG. 5.28: Latent heat flux (W m^{-2}) at Norman, Oklahoma for CTRL (black), MM5SOIL (green), MM5VEG (blue), and MM5VEGSOIL (red) domain four simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004 compared with the residual of the surface energy balance computed from Oklahoma Mesonet observations (dashed).

5.28). The maximum difference between the domain four MM5VEG and MM5SOIL forecasts of latent heat flux at Norman, Oklahoma exceeds 225 W m^{-2} in the forecast initialized on 3 May 2004. These flux changes are consistent with the increased soil moisture in all MM5SOIL simulations and the substantially reduced fractional vegetation coverage compared with climatology in the 3 May 2004 MM5VEG simulation. Additionally, the maximum difference in sensible heat flux exceeds 145 W m^{-2} (Fig. 5.29). Daytime sensible heat flux bias errors for MM5VEG across domain three exceed 90 W m^{-2} on 3 May and 1 August, though MM5SOIL shows remarkable improvements over the CTRL forecast (Fig. 5.30). A severe underestimation of sensible heat flux only at the Burneyville Mesonet site (Fig. 5.31) and an unrealistic spike in observations at the Grandfield Mesonet site both contribute to the uncharacteristically negative daytime sensible heat flux errors for the

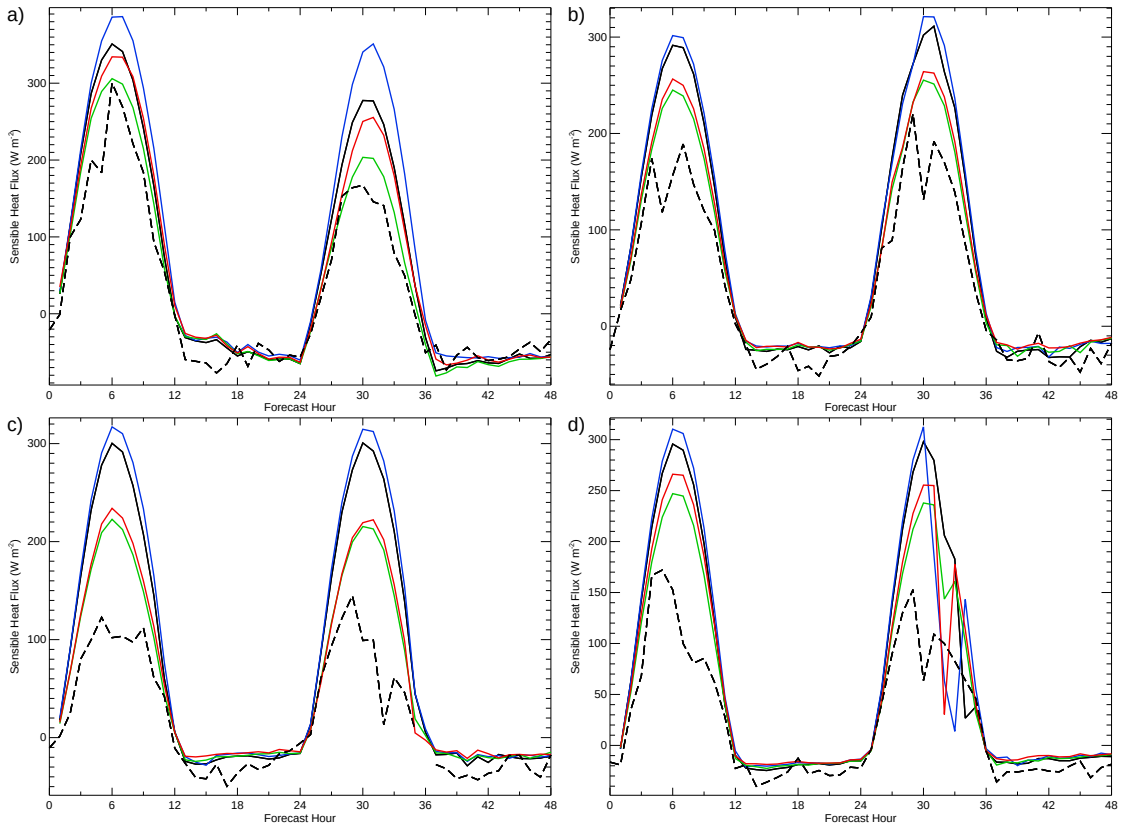

FIG. 5.29: Sensible heat flux (W m^{-2}) at Norman, Oklahoma for CTRL (black), MM5-SOIL (green), MM5VEG (blue), and MM5VEGSOIL (red) domain four simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004 compared with Oklahoma Mesonet observations (dashed).

20 July and 3 September 2004 simulations, respectively. Sensible heat flux comparisons between the model and observations at all other OASIS sites for these simulations exhibit results that resemble those for the other two case studies. These strange anomalies require more investigation. Ground heat fluxes are consistently too high at night (Fig. 5.32), though the magnitude of the ground heat flux errors throughout each simulation is substantially less than the errors in sensible and latent heat fluxes. Consistent with results from Marshall et al. (2003), the differences in partitioning between sensible and latent heat fluxes contribute to smaller daytime temperature and mixing ratio errors for MM5SOIL forecasts and larger errors for MM5VEG forecasts compared with the control forecast. At Norman, these factors result in differences in 2-meter air temperatures between the domain four MM5SOIL forecasts and the MM5VEG forecasts that may exceed 4°C at a given hour, though in each

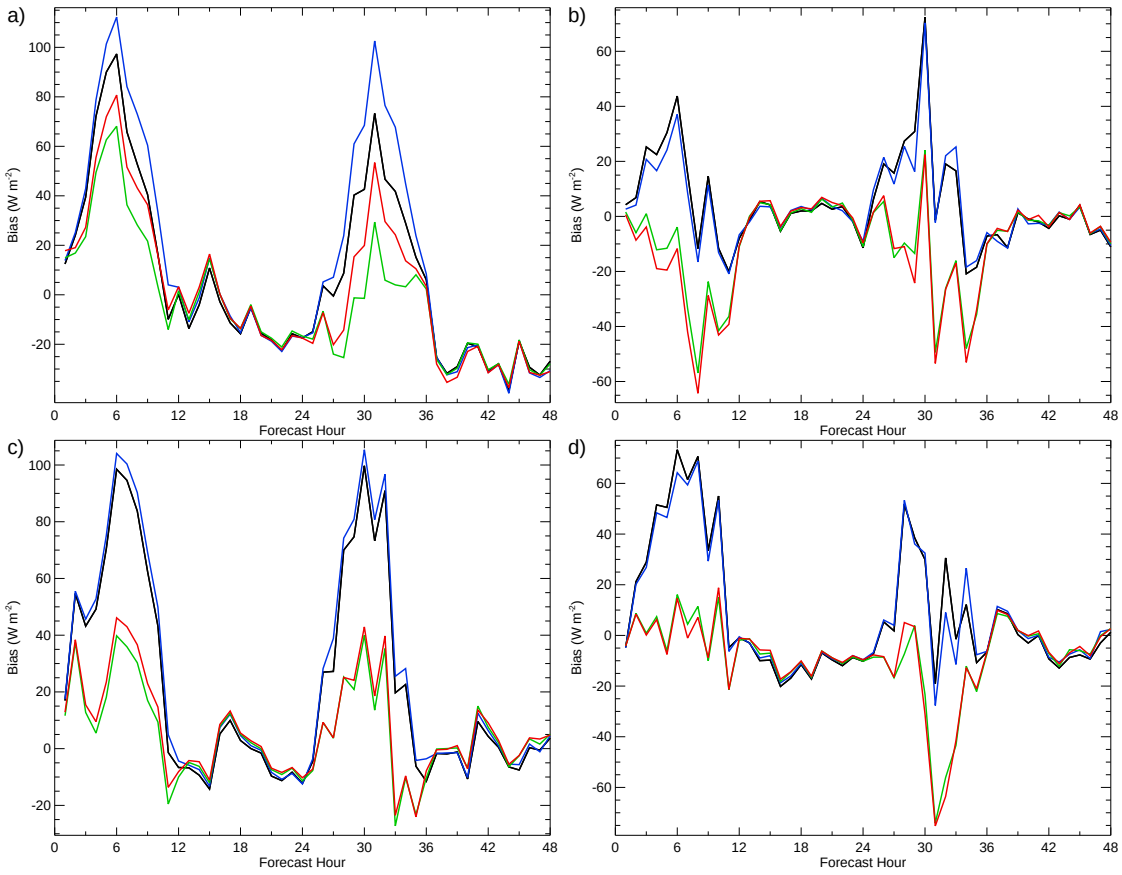

FIG. 5.30: Sensible heat flux bias (W m^{-2}) after comparison with Oklahoma Mesonet observations for CTRL (black), MM5SOIL (green), MM5VEG (blue), and MM5VEG-SOIL (red) domain three simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004.

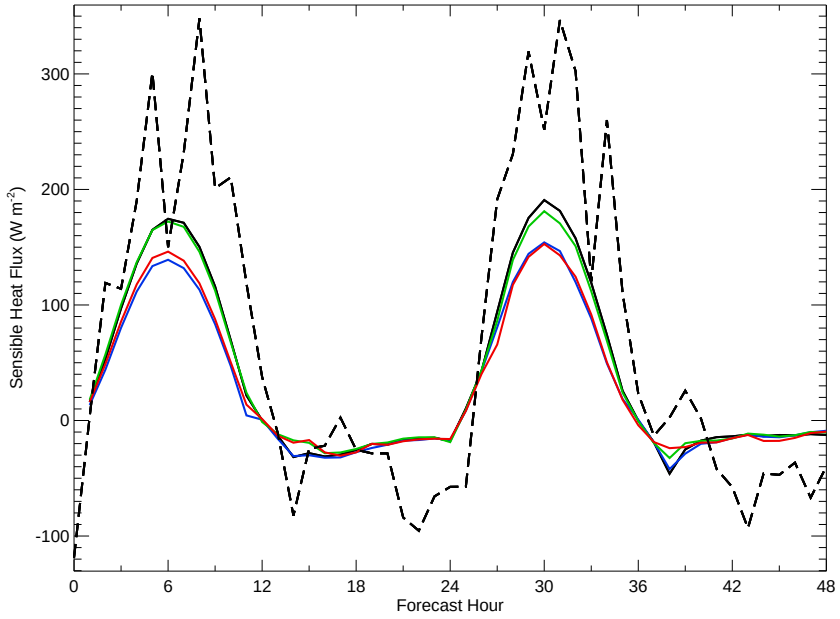

FIG. 5.31: Sensible heat flux (W m^{-2}) at Burneyville, Oklahoma for CTRL (black), MM5SOIL (green), MM5VEG (blue), and MM5VEGSOIL (red) domain three simulations initialized at 1200 UTC on 20 July 2004 compared with Oklahoma Mesonet observations (dashed).

case study daily maximum temperatures still exceed observations for all four forecast types (Fig. 5.33).

Positive temperature biases and negative mixing ratio biases during the day suggest that MM5 erroneously partitions sensible and latent heat fluxes. Summing the sensible and latent heat fluxes from the model and comparing the result with observations reveals a remarkable agreement between simulations and observations, particularly for the CTRL and MM5SOIL simulations (Fig. 5.34). This evidence indicates that partitioning errors between sensible and latent heat fluxes are responsible for a significant portion of the model errors. However, the magnitude of these heat fluxes tends to fall short of the observed values at night and in simulations that include observed vegetation indexes, indicating that other mechanisms also contribute to forecast errors.

Only the MM5SOIL forecast showed considerable improvement over the control forecast; verification with observations shows reduced forecast errors for most variables, despite using only climatological vegetation conditions. Forecasts incorporating satellite-derived vegetation indexes into the initial conditions, however, generally produced worse

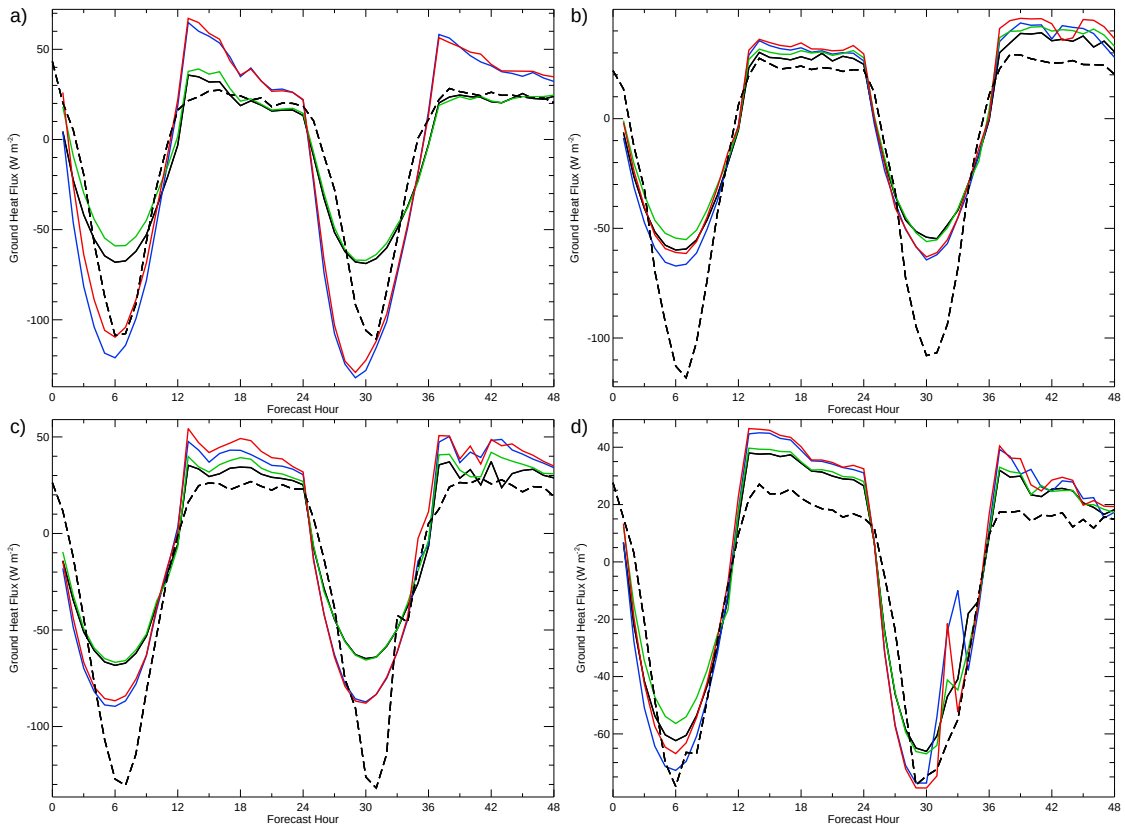

FIG. 5.32: Ground heat flux (W m^{-2}) at Norman, Oklahoma for CTRL (black), MM5-SOIL (green), MM5VEG (blue), and MM5VEGSOIL (red) domain four simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004 compared with Oklahoma Mesonet observations (dashed).

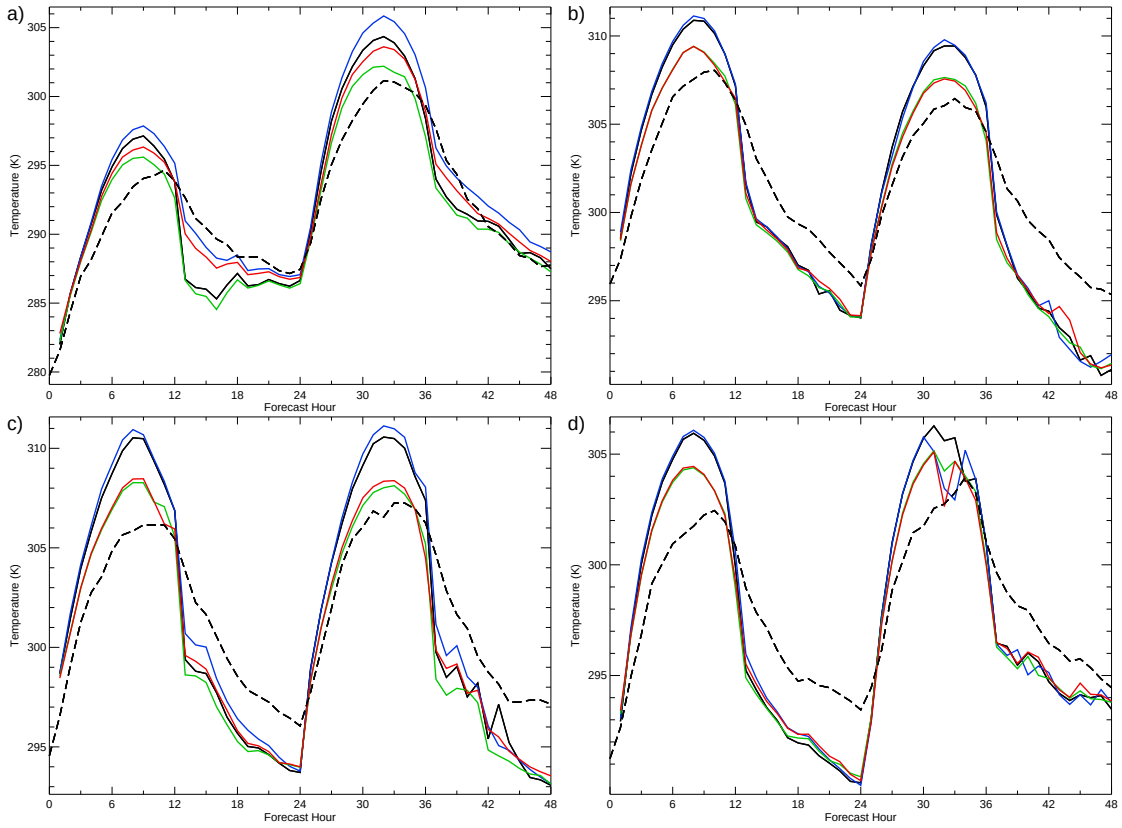

FIG. 5.33: 2-m air temperature (K) at Norman, Oklahoma for CTRL (black), MM5SOIL (green), MM5VEG (blue), and MM5VEGSOIL (red) domain four simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004 compared with Oklahoma Mesonet observations (dashed).

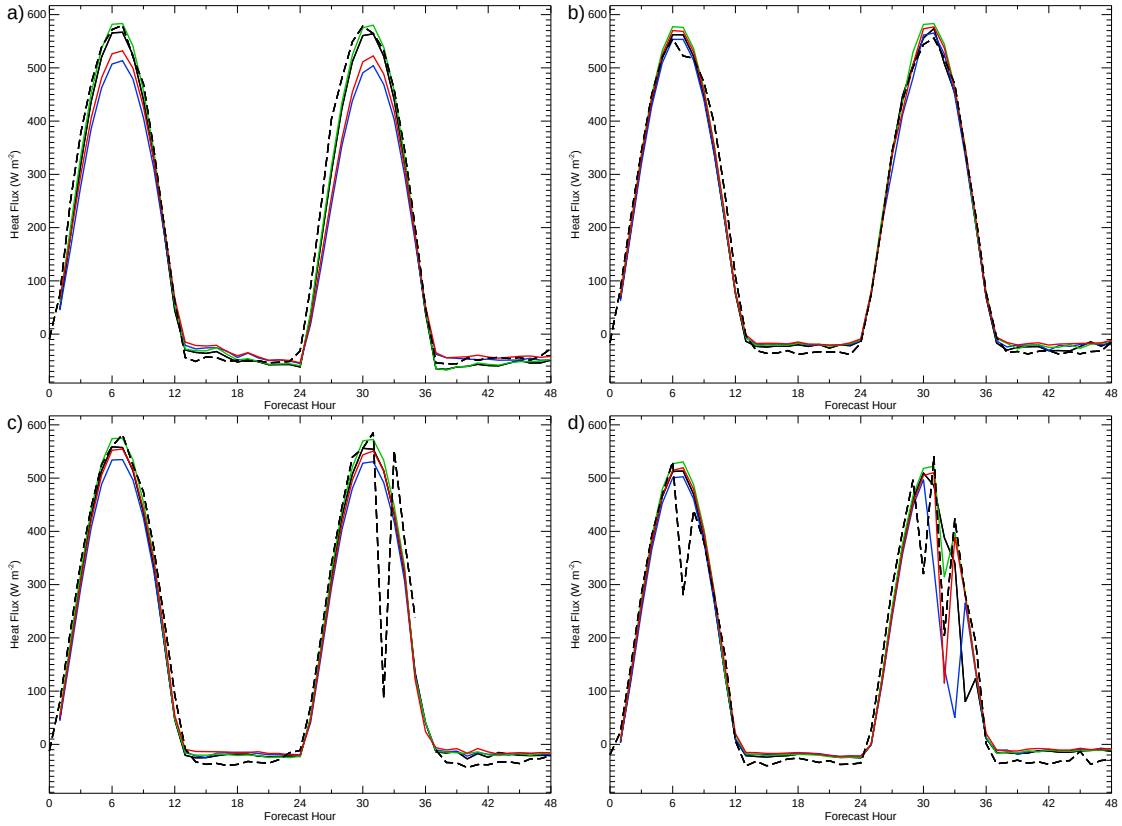

FIG. 5.34: Sum of sensible and latent heat fluxes (W m^{-2}) at Norman, Oklahoma for CTRL (black), MM5SOIL (green), MM5VEG (blue), and MM5VEGSOIL (red) domain four simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004 compared with Oklahoma Mesonet observations (dashed).

results than the control forecast. Testing the possibility that tuned parameterizations within the Noah LSM may provide the best results for constant LAI, forecasts incorporating observed σ_f and a constant LAI produce nearly identical results to those of the MM5VEG simulations. The drier than normal spring in 2004, and to some extent the dry summer, may exacerbate existing model errors. The MM5VEGSOIL forecast partially compensated for surface energy flux errors in the MM5VEG forecast by improving the initial soil temperature and moisture fields. It is important to note, however, that the greatly improved initial conditions in the MM5VEGSOIL simulation did not provide much improvement over the CTRL simulation. This result echoes the message from Robock et al. (2003), who stress that initial conditions with greater accuracy do not necessarily guarantee an improvement in model performance.

When given the best possible characterization of the initial land surface, these simulations highlight the errors present in the physical parameterizations within the Noah LSM. Such errors specifically suggest the need for the improved latent heat flux parameterization developed in chapter 6.

Chapter 6

Surface fluxes in the Noah LSM

The Noah LSM is the primary driver for land surface processes in the Eta, WRF, and MM5 forecast models. As shown in chapter 5, forecasts by the Noah LSM consistently underestimate midday latent heat fluxes by 20–40% compared with observations on clear days with weak synoptic forcing, even when given the best possible characterization of the initial land surface conditions. These errors primarily result from errors in the partitioning between the fluxes of sensible and latent heat. Partitioning errors can arise from a number of problems, including incorrect estimates of moisture availability, skin temperature, and resistance to heat flux, which is a function of air temperature and the vertical separation of atmospheric model layers. Improving surface fluxes may lead to better surface and boundary layer temperature and moisture forecasts, which will increase predictability (e.g., Crook 1996). The following discussion details the development of a new latent heat flux parameterization that derives from several months of Oklahoma Mesonet observations and provides improved predictions of sensible and latent heat fluxes. First, a summary of the present methods by which the Noah LSM calculates surface energy fluxes sets the stage for the development of this new scheme.

6.1 Latent heat flux

More complicated factors influence latent heat fluxes compared with sensible heat fluxes, exacerbating the difficulty of modeling evaporation near the land surface. However, surface observations under a variety of atmospheric conditions may aid in appropriately tuning the latent heat flux. In the current formulation within the Noah LSM, the latent heat flux E is the sum of the contribution from each of three types of evaporation: direct evaporation from bare soil (E_{dir}), transpiration from the vegetation canopy and roots (E_t), and evaporation of precipitation intercepted by the vegetation canopy (E_c). All three of these terms depend directly on the calculation of potential evaporation.

6.1.1 Potential evaporation

The potential evaporation E_p is the maximum possible evaporation that could occur over an open water surface under existing atmospheric conditions. The Noah LSM calculation for potential evaporation involves an energy balance approach based on the Penman relationship (Penman 1948) and includes a stability-dependent aerodynamic resistance term (Mahrt and Ek 1984). Since calculation of the net radiation in the model requires knowledge of the surface temperature, the actual set of equations in the model differs slightly from the usual Penman relationship and the equation for potential evaporation appearing in Mahrt and Ek (1984). This results in

$$E_p = \frac{\rho c_p C_h}{L_v}\left(\frac{\Delta\left[\frac{R_n}{\rho c_p C_h}+\left(\theta_0 - T_0\right)\right]+A(r+1)}{\Delta + r + 1}\right), \tag{6.1}$$

where ρ is the air density, c_p is the specific heat at constant pressure, and θ_0 and T_0 are the potential and actual temperatures at the lowest model level, respectively,

$$R_n = (1-\alpha)R_g + L_d - \sigma T_0^4 - G, \tag{6.2}$$

is the net radiation (W m^{-2}), where α is the surface albedo, R_g is the incoming solar radiation, L_d is the downward longwave radiation, G is the ground heat flux, and σ is the Stefan-Boltzmann constant,

$$\Delta = \frac{dq_s}{dT}\frac{L_v}{c_p}, \tag{6.3}$$

where dq_s/dT is the slope of the saturation specific humidity curve with respect to temperature and L_v is the latent heat of vaporization,

$$r = \frac{4\sigma T_0^4 R_d}{p_{\text{sfc}} c_p C_h}, \tag{6.4}$$

where R_d is the dry gas constant and p_{sfc} is the surface pressure (Pa), and

$$A = \frac{L_v}{c_p}\left(q_s\left(T_0\right) - q_0\right), \tag{6.5}$$

where q_0 and $q_s(T_0)$ are the actual and saturation specific humidities at the model level closest to the ground surface, respectively (Ek and Mahrt 1991). Here C_h is the surface exchange coefficient for heat and momentum, the definition of which varies depending upon the stability of the lower atmosphere, and is a function of the wind speed and height above the surface at the first model level, the roughness lengths for momentum and heat, and the bulk Richardson number for the surface layer. For details of the calculation of C_h, see Mahrt and Ek (1984) and Ek and Mahrt (1991). Note that in the Noah LSM, the model replaces the actual and saturation specific humidities with their nearly equivalent mixing ratio counterparts.

6.1.2 Direct evaporation from bare soil

The direct evaporation term is a simple linear relationship based on the work of Mahfouf and Noilhan (1991), who use a moisture availability parameter β to scale the evaporation from the soil. The Noah LSM employs a similar approach based on the results from a

sensitivity test for evaporation in the NCEP Eta model (Betts et al. 1997) in which

$$E_{\text{dir}} = (1 - \sigma_f)\beta E_p, \tag{6.6}$$

where σ_f is the fractional vegetation coverage for a model grid cell and

$$\beta = \left(\frac{\Theta_1 - \Theta_w}{\Theta_{\text{ref}} - \Theta_w} \right)^f \tag{6.7}$$

represents a normalized soil moisture availability term where Θ_w is the wilting point and Θ_{ref} is the field capacity, both of which depend on soil texture, and Θ_1 is the volumetric water content of the top soil layer (Chen and Dudhia 2001). Some studies set $f = 1$ (e.g., Betts et al. 1997; Chen and Dudhia 2001), though in the version of the Noah LSM used here, $f = 2$ as suggested by Ek et al. (2003).

6.1.3 Canopy transpiration

The canopy transpiration from the vegetated portion of a model grid cell is

$$E_t = \sigma_f E_p P_c \left[1 - \left(\frac{W_c}{S} \right)^{0.5} \right], \tag{6.8}$$

where W_c is the intercepted canopy water content and S is a constant but tunable maximum canopy water capacity. The plant coefficient P_c includes the influence of stomatal control and is expressed as

$$P_c = \frac{r + \Delta}{r(1 + C_h R_c) + \Delta}, \tag{6.9}$$

where r, Δ, and C_h were defined in section 6.1.1 and

$$R_c = \frac{R_{\text{cmin}}}{\text{LAI} F_1 F_2 F_3 F_4} \tag{6.10}$$

is the canopy resistance following the formulation of Jacquemin and Noilhan (1990) where R_{cmin} is the minimum stomatal resistance for each vegetation type and LAI is the leaf area index. The canopy resistance factors F_1, F_2, F_3, and F_4 are each bounded between 0 and 1 and represent the effects of solar radiation, vapor pressure deficit, air temperature, and soil moisture, respectively (Chen and Dudhia 2001). These factors are defined by

$$\begin{aligned}
F_1 &= \frac{R_{\text{cmin}}/R_{\text{cmax}} + f}{1+f} \quad \text{where } f = 0.55\frac{R_g}{R_{\text{gl}}}\frac{2}{\text{LAI}}, \\
F_2 &= \frac{1}{1+h_s\left[w_s\left(T_0\right) - w_0\right]}, \\
F_3 &= 1 - 0.0016\left(T_{\text{ref}} - T_0\right)^2, \text{and} \\
F_4 &= \sum_{i=1}^{n} \frac{(\Theta_i - \Theta_w)d_{z_i}}{(\Theta_{\text{ref}} - \Theta_w)\left(\sum_{j=1}^{n} d_{z_j}\right)},
\end{aligned} \tag{6.11}$$

where R_{cmax} is a constant maximum canopy resistance set to 5000 s m^{-1}, R_g is the incoming solar radiation, R_{gl} and h_s are species-dependent radiation stress and empirical parameters, respectively, w_s is the saturation mixing ratio at air temperature T_0, w_0 is the mixing ratio, T_{ref} is 298 K as in Noilhan and Planton (1989), Θ is the volumetric water content of each soil layer, d_z is the depth of each individual soil layer, and n is the species-dependent number of root zone soil layers.

The canopy resistance is the most important factor contributing to canopy transpiration. Holtslag and Ek (1996) write "the [total] latent heat flux is mostly determined by the canopy resistance." Despite this physical importance, Eq. (6.10) that describes the canopy resistance is arguably the most questionable formulation in the Noah LSM, since it simply multiplies together four physically important atmospheric and land surface effects. Jarvis (1976) proposed a very similar formulation in an effort to forecast stomatal conductance (the inverse of which is resistance) based on the known independent influence of air temperature, leaf-air vapor pressure difference, carbon dioxide concentration, and water stress on the stomatal conductance of leaves illuminated by solar radiation. Without knowing the effect on stomatal conductance from each variable acting in concert, Jarvis (1976)

hypothesized that the final stomatal conductance "is the result of complete expression of the influence of all the variables without any synergistic interactions." The final stomatal conductance is thus the product of the percentages of the maximum stomatal conductance contributed by each variable. This formulation, which several authors have adopted and implemented in land surface models with some modification (e.g., Noilhan and Planton 1989; Jacquemin and Noilhan 1990; Chen and Dudhia 2001), leads to the canopy resistance factors that appear in Eq. (6.11).

6.1.4 Wet canopy evaporation

The evaporation of precipitation intercepted by the vegetation canopy is substantially smaller than the other evaporation terms and is

$$E_c = \sigma_f E_p \left(\frac{W_c}{S} \right)^{0.5}. \tag{6.12}$$

In the experiments that follow, Oklahoma Mesonet observations facilitate the development of an empirical parameterization for latent heat flux. At Oklahoma Mesonet sites, where the predominant vegetation cover is grass, it is assumed that the canopy water content is zero, thereby removing the contribution to evaporation by moisture in the vegetation canopy (cf., Betts et al. 1997). The total latent heat flux is therefore the sum of the direct evaporation and canopy transpiration terms. This is a reasonable assumption given the relative insignificance of E_c compared with E_{dir} and E_t.

6.2 Sensible heat flux

The Noah LSM calculates the sensible heat flux as

$$H = \rho c_p C_h (T_s - \theta_0), \tag{6.13}$$

where ρ is the air density, c_p is the specific heat of air at constant pressure, C_h is the surface exchange coefficient for heat and momentum discussed in section 6.1.1, T_s is the skin temperature, and θ_0 is the potential temperature at the first model level.

6.3 Ground heat flux

The ground heat flux in the Noah LSM is an integration of the three-dimensional heat conduction equation for soil evaluated at the land-atmosphere interface such that

$$G = \left(\kappa \frac{\partial T}{\partial z} \right) \Bigg|_{z=0}, \tag{6.14}$$

which is scaled to account for the effect of vegetation coverage on soil heat fluxes. As implemented in the model, the ground heat flux becomes

$$G = \frac{\kappa (T_s - T_{\text{soil}_1})}{0.5 d_{z_1}} e^{-2\sigma_f}, \tag{6.15}$$

where T_s is the skin temperature, T_{soil_1} is the temperature of the first soil layer, d_{z_1} is the depth of the first soil layer, and σ_f is the fractional vegetation coverage. The thermal conductivity

$$\kappa = K_e (\kappa_{\text{sat}} - \kappa_{\text{dry}}) + \kappa_{\text{dry}} \tag{6.16}$$

is weighted by the Kersten number (K_e), which is a normalized thermal conductivity and is a function of the volumetric water content in the first soil layer (Θ_1) and the soil porosity (Θ_s). The dry thermal conductivity, κ_{dry}, is a function of Θ_s and the saturated thermal conductivity, κ_{sat}, is a function of both Θ_s and quartz content. Details of the thermal conductivity calculations appear in Peters-Lidard et al. (1998).

6.4 Empirical latent heat flux parameterization

The latent heat flux formulation in the Noah LSM is much more complicated, and therefore prone to more errors, than the sensible and ground heat flux equations. The combination of the complexity of the physical processes leading to evapotranspiration and the assumptions inherent in the current canopy transpiration term make the latent heat flux a prime candidate for refinement. The unique set of soil, vegetation, atmospheric, and surface flux observations available in Oklahoma presents an ideal opportunity for improving model forecasts by examining the latent heat flux.

Given the physical importance of canopy resistance in the current approach to calculating the canopy transpiration term in the Noah LSM, one technique for improving short-term latent heat flux forecasts is to focus on the formulation for canopy resistance, while leaving the remainder of the Noah LSM untouched. Driven by observations, a reversed form of the Noah LSM provides values of plant coefficient, P_c, and thereby canopy resistance, R_c, that would be necessary for the original model to yield the observed latent heat flux. Unfortunately, many of these values are unphysical, including exceedingly large canopy resistances and unbounded plant coefficients. Results indicate that this occurs because 1) the E_{dir} term (Eq. 6.6) is greater than the observed latent heat flux or 2) the sum of E_{dir} and $\sigma_f E_p$ (Eq. 6.8) in the Noah LSM is less than the observed latent heat flux, even after adjusting for a maximum $\pm$ 20 W m^{-2} bias in the latent heat flux observations based on instrumentation error studies by Brotzge (2000). Thus, the direct evaporation from bare soil and canopy transpiration terms clearly yield inappropriate values when forced with observations. Any empirical scheme designed to forecast P_c or R_c based on these formulae would lead to poor model forecasts of latent heat flux. Improved forecasts for latent heat flux clearly require a different approach. Therefore, the popular canopy resistance approach to modeling canopy transpiration is abandoned and instead a completely new empirical latent heat flux scheme is developed from all available sets of observations. Tests indicate that least squares simple and multiple linear regression models with automatic and manual predictor selection have

limited potential, though a principal-component regression procedure holds more promise as a viable alternative for predictor selection. The following development of a new latent heat flux parameterization takes advantage of the properties of this principal-component regression technique.

6.4.1 Principal-component regression

Meteorological data generally exhibit large spatial and temporal correlations. Least squares multiple linear regression models trained on such highly correlated data are therefore unstable and may perform poorly on independent data (Wilks 2006, p. 505). These mutual correlations in the independent data can be removed through a principal component analysis, which transforms a time series of correlated variables into temporally uncorrelated, spatially and temporally orthogonal time series that remain linear functions of all the original variables (Haan 1977, p. 237; Richman 1986). These principal components become the set of predictor variables in a least squares multiple linear regression. One benefit of using principal component analysis is that it provides an objective method for eliminating variables that are not highly correlated with any of the principal components before using those variables in a principal-component regression. Secondly, because each principal-component predictor is temporally uncorrelated with the others, elimination of any principal component as an independent variable in a multiple regression analysis does not affect the contribution of any of the other components.

Principal-component regression techniques are not new in studies of the atmosphere. Predictions of tropical precipitation from marine surface observations (Tsonis 2002), mean winter temperatures from sea-surface temperatures and pressure-surface heights (Harnack 1979), and wheat yield from temperature and rainfall observations (Wigley and Qipu 1983) have all employed this method. Air quality studies have also exercised this technique to forecast surface ozone concentrations (Pryor et al. 1995) and to determine source regions for fine particulates and sulphate (Wolff et al. 1984). However, application of this technique

in an attempt to predict fluxes of latent heat from a wealth of surface observations represents a novel approach.

Since land surface models contain separate expressions for latent heat flux over bare soil and vegetated surfaces, employing separate principal-component regression analyses yields the best possible expressions for both E_{dir} and E_t to match the observed latent heat fluxes. The first step in predictor selection requires determining which original variables display a quasi-linear relationship with each contributor to the total latent heat flux. A locally weighted regression accomplishes this goal by smoothing the dependent variable as a function of the independent variable over a moving window, analogous to a moving average for a time series (Cleveland and Devlin 1988). Locally weighted regression curves that appear linear with non-zero slope suggest that a particular variable would perform well in a multivariate regression. This provides an objective method for determining which mathematical transformations of potential variables would improve the forecasts of both E_{dir} and E_t.

Another necessary step before analyzing the surface data in a principal component analysis is to examine the autocorrelation of the selected variables. Whether chronologically or randomly ordered, many of the variables exhibit strong serial correlations out to several tens of lags. Since the data are largely from Oklahoma Mesonet measurements at approximately half-hour intervals (some data are missing), these serial correlations likely represent diurnal, synoptic, and seasonal cycles. Sampling the data with a frequency low enough to reduce these serial correlations substantially would vastly reduce the number of available observations for use in determining a new empirical scheme for latent heat flux, thereby limiting the ability of a new scheme to accurately forecast latent heat fluxes for all times, seasons, and locations. Thus, there lies a massive serial correlation problem within the data available for analysis such that the degrees of freedom are too high and the resulting $p-$values are too small in the principal-component regression because the statistical test has too much power. Rather than implementing strictly objective methods for determining the appropriate terms in the principal-component regression, an alternative method relies

on variable selection through a trial-and-error approach (M. B. Richman 2006, personal communication).

Training data for both E_{dir} and E_t principal-component regressions derive from randomly selected sets of observations containing possible predictors and their respective predictands, which constitute approximately half of the available data. The remaining data are used for independent cross-validation. These independent data provide a measure of the strength of the multiple regression relationship through several measures, including the coefficient of determination R^2 and the residual standard error (Wilks 2006, 185–186). One negative characteristic of the coefficient of determination is that its value continually increases by simply adding more variables to a prediction equation. Thus, an adjusted R^2, such that

$$\overline{R}^2 = 1 - \left(1 - R^2\right)\left(\frac{n-1}{n-p-1}\right), \tag{6.17}$$

instead corrects for such a problem by the inclusion of a penalty term, where p is the number of predictors in the multiple regression model and n is the sample size (Yamane 1967, p. 765). The $\overline{R}^2$ value justifies the results of each principal-component regression in each independent cross-validation data set.

For the principal component analysis, let $\mathbf{X}$ denote the $n \times p$ matrix of n observations on p predictor variables and $\mathbf{R}$ denote the correlation matrix of these data. Standardizing the data through the mathematical transformation

$$\mathbf{X} = \frac{x_{ij} - \overline{x}_j}{\sigma_j}, \tag{6.18}$$

where σ_j is the population standard deviation of each column j, removes the problem of differing units with widely varying magnitudes in the data and the analysis computes the principal components from the resulting correlation matrix. This is standard practice when computing principal components from variables with differing units (Haan 1977, p. 243). If the units were similar in magnitude and the data were instead standardized using $\mathbf{X} = x_{ij} - \overline{x}_j$, then the covariance matrix could substitute for the correlation matrix $\mathbf{R}$ in the

following calculations.

The $n \times p$ matrix of n principal component scores for each of p principal components **U** is

$$\mathbf{U} = \mathbf{XE}, \tag{6.19}$$

where the matrix of eigenvectors **E** is a $p \times p$ linear transformation matrix whose j^{th} column is the eigenvector e_j determined from

$$\left(\mathbf{R} - \lambda_j \mathbf{I}\right) e_j = 0, \tag{6.20}$$

where **I** is the identity matrix and λ_j is the set of eigenvalues for the matrix **R**. Following the calculation of the principal components, a varimax rotation (Kaiser 1958; Richman 1986) simplifies the physical interpretation of each eigenvector by maximizing or minimizing its eigenvector elements. A varimax rotation is an orthogonal rotation that rotates the matrix of eigenvectors through a $p \times p$ transformation matrix **T** such that

$$\widetilde{\mathbf{E}} = \mathbf{ET}. \tag{6.21}$$

This transformation produces rotated orthogonal eigenvectors, but the corresponding rotated principal component time series are no longer uncorrelated (Wilks 2006, p. 498). For details of the varimax rotation mechanics, see Kaiser (1958), Richman (1986), or Wilks (2006, p. 494). Finally,

$$\widetilde{\mathbf{U}} = \mathbf{X}\widetilde{\mathbf{E}} = \mathbf{XET} \tag{6.22}$$

defines the $n \times p$ matrix of rotated principal components.

Any forecast equation derived through principal component analysis requires measurements of all of the original variables. Eliminating variables exhibiting low correlations with any principal component reduces the number of variables that appear in any resulting regression equation. A second principal component analysis on the remaining variables,

followed by removal of insignificant principal components containing rotated eigenvector elements smaller than ± 0.85 for all variables, yields a final set of rotated principal components $\widetilde{\mathbf{U}}$ that serve as predictors for a multiple linear regression with either E_{dir} or E_t as the predictands. A reverse transformation matrix

$$\mathbf{B} = \mathbf{E}\mathbf{D}_{\lambda}^{-1/2}\mathbf{T}\mathbf{D}_{\lambda}^{1/2} \tag{6.23}$$

converts the resulting principal-component regression equation into an equation in terms of the original variables for use in the Noah land surface model. Here, $\mathbf{D}_{\lambda}$ is a $p \times p$ diagonal matrix whose diagonal elements are the eigenvalues of $\mathbf{R}$ and $\mathbf{D}_{\lambda}^{-1/2}$ is a diagonal matrix formed by inverting $\mathbf{D}_{\lambda}$ and taking the square root of its elements.

6.4.2 Selection of observations

Practical and physical considerations limit the range of possible predictor variables in a principal-component regression. The simplest choices for possible predictors include variables that already exist within the Noah LSM. To calculate surface energy fluxes, the model manipulates several atmospheric and soil variables determined from either internal calculations or input from the parent atmospheric model. These quantities include the downward component of both longwave and shortwave radiation, surface pressure, precipitation rate, air temperature, mixing ratio, wind speed, potential temperature, fractional vegetation coverage, leaf area index, soil temperature and soil moisture for several layers, and skin temperature. Combinations of these variables define other necessary quantities, including the saturation mixing ratio and the slope of the saturation mixing ratio curve at the current air temperature. The Oklahoma Mesonet provides observations of the majority of these variables, while NOAA AVHRR satellites measure the fractional vegetation coverage and leaf area index. Since the satellite-derived quantities are only available as biweekly composites, these observations require temporal interpolation to match the observation time of the Oklahoma Mesonet observations. Refer to chapter 3 for details on the

surface and space-based instrumentation.

Data available for analysis span the period May 2004–June 2006 with satellite-derived vegetation data spanning only the period 15 April–15 September 2004. Within this time frame, there are several restrictions on the available observations from the Oklahoma Mesonet. Complete sets of quantities determining the surface energy balance are only available from nine OASIS super sites.[1] Since the latent heat flux represents the residual of the surface energy balance, the sensible heat flux, ground heat flux, and net radiation must be present in each observation record. Estimates of the storage ground heat flux require measurements of the volumetric water content at 5-cm, which is only sampled every 30 minutes. Latent heat flux observations are therefore only available at the top and bottom of each hour. Since precipitation is known to interfere with measurements from a sonic anemometer, which measures the sensible heat flux, periods of rainfall are removed from consideration beginning with the first non-zero daily precipitation total through local midnight on the day of the observation.

Sets of possible predictor variables are limited to periods with ample incoming shortwave radiation. To focus on this most important part of the day, and to limit the influence of very small nighttime latent heat fluxes in creating a new scheme for latent heat flux forecasts, the principal-component regression only considers sets of observations with incoming shortwave radiation that exceeds 10 W m^{-2}.

In addition to these quantities, the Noah LSM also contains several fixed quantities that could be useful directly or in transformations as possible predictors. These include albedo, roughness length, and soil type. The wilting point and field capacity at model grid points are functions of soil type. Soil cores from each OASIS super site provide the soil texture class at each of four different levels corresponding roughly with the four available soil layers in the Noah LSM. While observations of both incoming and outgoing shortwave radiation provide accurate albedo measurements at each OASIS super site, practical usage in the Noah LSM limits the utility of such a parameter in an improved formula for latent heat flux.

[1]The Burneyville, Oklahoma site did not measure ground heat flux during this period.

Similarly, observations of the wind speed at multiple levels may provide information about the roughness length. Simply assuming neutral stability in the boundary layer to determine the roughness length from observations leads to widely varying roughness lengths that would have little practical application in the Noah LSM. Therefore, the set of possible predictor variables does not describe variability due to changes in albedo and roughness length, leaving the model to determine these values from lookup tables based on the USGS 24-class vegetation type database categorization.

Several assumptions apply to the Oklahoma Mesonet data when considering potential predictor variables in a regression equation for use in the Noah LSM. A major branch occurs at the beginning of the Noah LSM code that accounts for the effect of snow cover. Since Oklahoma Mesonet data do not explicitly provide information on snow cover, accounting for this effect while searching for predictor variables in a new latent heat flux parameterization would prove difficult. Observations therefore fall under the assumption that there is no snow cover, that water in the soil is not frozen, and that precipitation is not falling in the form of snow or freezing rain, which a Mesonet tipping bucket gauge would not immediately register. These are reasonable assumptions for Oklahoma for the relevant spring and summer observation period.

6.4.3 Direct evaporation from bare soil

Since vegetated surfaces surround every observation site, direct measurements of evaporation from bare soil are unavailable from the Oklahoma Mesonet. However, a long time series of soil moisture observations from May 2004 through June 2006 contains several periods during which the vegetation in Oklahoma suffered under moderate to extreme drought conditions. The permanent wilting point where transpiration ceases for most vegetation is roughly where the matric potential $\psi = -1500$ kPa (Marshall et al. 1996, 37–38). At locations where the matric potential is larger in magnitude than the permanent wilting point, the only contribution to the total latent heat flux is from bare soil evaporation, assuming zero

canopy water content. By separating only those sets of observations at Oklahoma Mesonet sites where the soil has reached the permanent wilting point at the 5-cm level, the residual of the surface energy balance becomes a good approximation to the direct evaporation from bare soil.

Despite the purported accuracy of the empirical matric potential formula used previously (Eq. 3.2), this formula is insufficient for capturing the matric potential under very dry soil conditions, since the minimum possible matric potential is $\psi = -852.2$ kPa, well above the permanent wilting point of most vegetation. A different matric potential formula,

$$\psi = \frac{1}{a}\left(\frac{\Delta T_w - \Delta T_d}{\Delta T_{\text{ref}} - \Delta T_d} - 0.9\right)^{1/n}, \tag{6.24}$$

where the standard temperature differences for dry and saturated soil are $\Delta T_d = 4.0$°C and $\Delta T_w = 1.45$°C, respectively, and $a = -0.01$ kPa^{-1} and $n = 0.77$ are empirical coefficients (Basara and Crawford 2000), yields a better response and is a more physical representation of actual soil processes under very dry soil conditions. Figure 6.1 compares the response of Eq. (3.2) with Eq. (6.24) for the entire spectrum of allowable normalized reference temperature observations. Each formula responds similarly for intermediate normalized reference temperatures, but Eq. (6.24) adjusts the matric potential such that moist soil is more moist and dry soil is more dry compared with Eq. (3.2). The implication of the response of this equation to very dry soil conditions is that it yields more physically realistic matric potential values and simplifies the identification of extremely dry periods.

Since the precision of the normalized reference temperature observations is 0.1°C, the discrete matric potential value calculated from Eq. (6.24) that is nearest to the permanent wilting point is just under $\psi = -1490$ kPa. Therefore, a subset of the available Oklahoma Mesonet observations containing matric potentials less than $\psi = -1490$ kPa represents extremely dry soil conditions where presumably transpiration has ceased and the latent heat flux observations are equivalent to the direct soil evaporation. This subset of data comprises more than 6300 sets of observations between May 2004 and June 2006 for use

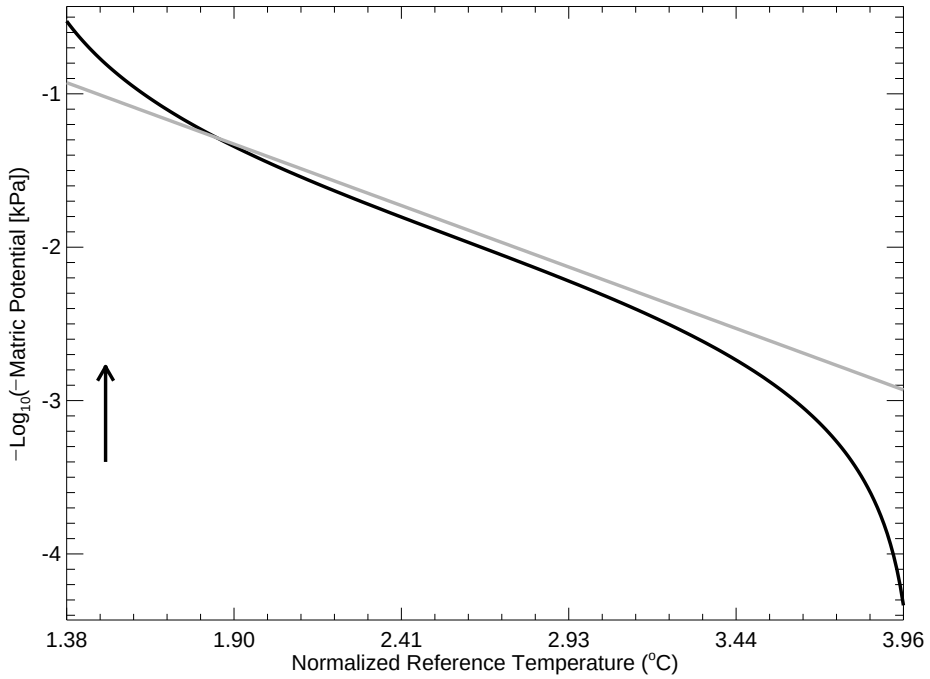

FIG. 6.1: Comparison of matric potential (kPa) obtained from Eq. (3.2) (gray) and Eq. (6.24) (black) over the range of allowable normalized reference temperature observations (°C) from the Oklahoma Mesonet. The arrow indicates the direction of increasing soil moisture.

in determining a new direct soil evaporation parameterization.

A substantial portion of the data during dry periods falls outside the summertime 2004 satellite measurement window. However, under the assumption that transpiration ceases when the magnitude of the matric potential exceeds the permanent wilting point, any vegetation coverage would not contribute to the total latent heat flux. The fractional vegetation coverage therefore is set to zero regardless of the availability of satellite observations.

Following the procedure outlined in section 6.4.1, a locally weighted regression with E_{dir} as the predictand and a host of observable variables and selected transformations as predictors leads to potentially useful variables in the final regression equation. From a wide selection of possible observable or transformed variables, multiple passes through a principal component analysis lead to a reduced pool of possible predictors for E_{dir}. In addition to the overarching goal of achieving the largest possible adjusted R^2 value in the cross-validation data, several other factors contribute to the decision to retain or eliminate variables from the principal-component regression. Among these factors is the ease of

implementation of the resulting flux equation in the Noah LSM. For example, matric potential holds promise as a possible predictor for latent heat flux and relates directly to water movement and plant-water uptake. Though several soil properties depend upon soil type, observations of matric potential from the Oklahoma Mesonet are independent of soil type. Including matric potential in the Noah LSM removes an exponential dependence upon crude estimates of soil type and presumably improves the specification of water in the soil. However, results from principal-component regression tests show that replacing volumetric water content with matric potential does not generate enough improvement in E_{dir} forecasts to justify the difficulty of adding matric potential as a prognostic variable. Other factors include the physical relevance of each variable to evaporative processes and the statistical significance of each variable when included in a multiple linear regression. Additionally, several combinations of variables possess strong mutual correlations and must not appear together in the final regression equation. For example, the correlation coefficient between the mixing ratio and the 2-m air temperature is 0.66. Correlations are also high between incoming longwave radiation, the 2-m and 9-m air temperature, mixing ratio, saturation mixing ratio, potential temperature, and the derivative of saturation mixing ratio with respect to temperature because of the strong relationship between the air temperature and atmospheric moisture content. The existence of such highly correlated variables justifies the use of the principal-component approach in variable selection, even if the final regression equation retains all of the principal components.

The principal-component regression procedure yields a regression equation for direct evaporation from bare soil assuming that no transpiring vegetation contributes to the total latent heat flux. In practice, the fractional vegetation coverage scales the direct evaporation from bare soil. Therefore, the final equation for direct evaporation from bare soil is

$$E_{\mathrm{dir}} = \left(22.33 + 0.0226 \left[R_g \left(1 - \alpha \right) \right]^{(3/2)} \left[\frac{\Theta_1 - \Theta_w}{\Theta_{\mathrm{ref}} - \Theta_w} \right]^f - 3.426V + 3650w \right) \left(1 - \sigma_f \right), \tag{6.25}$$

where R_g is the incoming solar radiation (W m^{-1}), α is the albedo based on the Noah LSM

land use category, Θ_1 is the volumetric water content ($m^3\ m^{-3}$) at 5-cm depth, Θ_w is the wilting point and Θ_{ref} is the field capacity, both of which depend on the 5-cm soil texture measured at each Oklahoma Mesonet site, f=1, V is the 10-m wind speed ($m\ s^{-1}$), w is the 2-m mixing ratio ($kg\ kg^{-1}$), and σ_f is the fractional vegetation coverage. As implemented in the Noah LSM, Θ_1 is the volumetric water content of the top soil layer, Θ_w and Θ_{ref} refer to the wilting point and field capacity of the relevant gridded soil type, and V and w are the wind speed and mixing ratio at the lowest model level. The adjusted R^2 for the independent cross-validation data is 0.61, giving a correlation coefficient between forecasts and observations of 0.78, and the residual standard error is 48.4 $W\ m^{-2}$. By comparison, the correlation coefficient between the same predictand and the direct soil evaporation from the original Noah LSM formulation is 0.52.

As indicated by locally weighted regressions prior to the principal-component regression, each of the variables in Eq. (6.25) exhibits a quasi-linear relationship with the observed latent heat flux during dry conditions. The second term on the right-hand side of Eq. (6.25) materializes by recognizing that the available soil moisture tempers the evaporative power of the sun. An excellent linear relationship with E_{dir} in a locally weighted regression arises by multiplying the effective incoming solar radiation (incoming solar radiation minus outgoing solar radiation) raised to the 3/2 power by the normalized soil moisture availability term β from Eq. (6.7).

With the exception of the vegetation fraction term, each term in Eq. (6.25) represents a single variable present in the principal component analysis. Since each component contains a very strong signal from one of each of the three variables, the final regression equation retains all three principal components. A multiple linear regression on these variables produces the same regression equation, but the large correlations between the available variables justifies using the principal-component regression approach both to ascertain the significance of the mutual correlations and as a robust variable-selection method.

Compared with the existing direct soil evaporation parameterization in the Noah LSM, the forecasts from the new empirical scheme more closely match the total latent heat flux

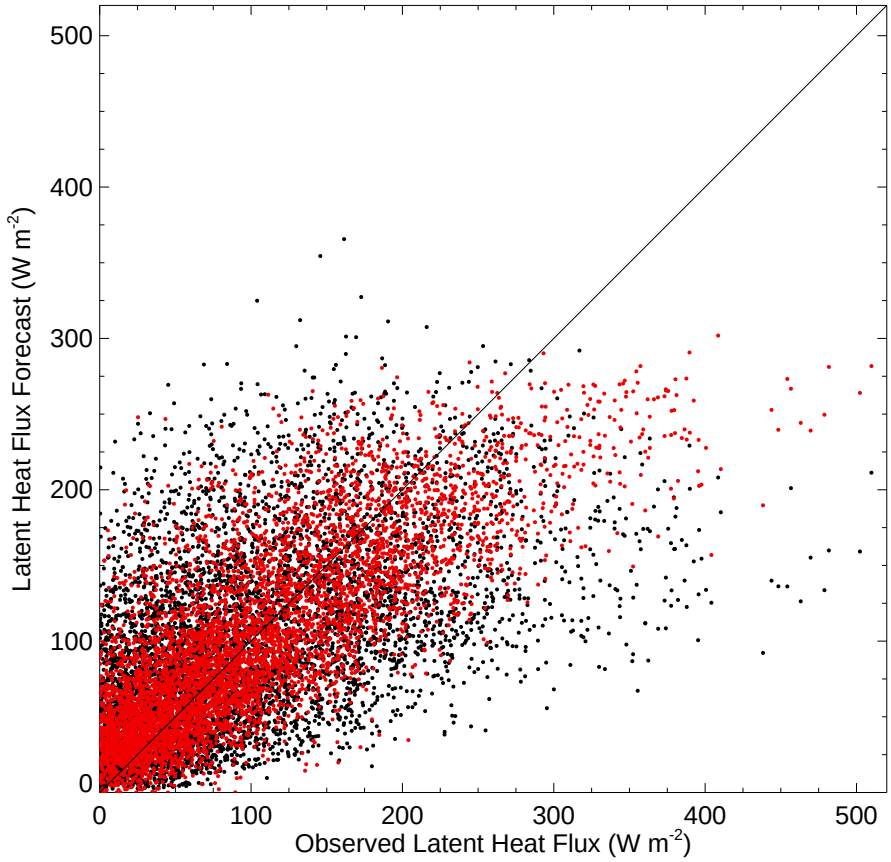

FIG. 6.2: Direct soil evaporation from the original Noah LSM formulation (black) and the empirical scheme (red) compared with the observed total latent heat flux under dry soil conditions.

observations when the soil is dry enough to assume senescent vegetation, particularly for increased direct soil evaporation (Fig. 6.2). When the soil contains sufficient moisture to support canopy transpiration, the individual contribution to evaporation from bare soil in the Noah LSM certainly should not exceed the observed total latent heat flux. Regardless of soil conditions, the new empirical scheme improves upon the original Noah LSM direct soil evaporation formulation when applied to all available sets of observations during the period 15 April–15 September 2004 by reducing substantially the frequency of unrealistic E_{dir} values that exceed the observed total latent heat flux.

6.4.4 Canopy transpiration

With a proper parameterization for the direct evaporation from bare soil in place, a similar principal-component regression procedure leads to a new empirical canopy transpiration scheme. The training and independent cross-validation data include fractional vegetation coverage and leaf area index observations spanning the period 15 April–15 September 2004. The canopy transpiration term defined by

$$E_t = \frac{E_{\text{obs}} - E_{\text{dir}}}{\sigma_f} \tag{6.26}$$

is the predictor in the multiple regression, where E_{obs} is the observed total latent heat flux and E_{dir} is the empirical direct soil evaporation term from Eq. (6.25) that already includes the vegetation fraction weighting.

Observed variables and those transformed based on physically plausible relationships and locally estimated regressions compose a diverse set of possible forecast variables. As with the direct soil evaporation parameterization, a principal component analysis combined with physical, statistical, and practical considerations leads to the final regression equation for canopy transpiration,

$$\begin{aligned} E_t = \Bigg(& -1392 + 0.9154 \left[R_g (1-\alpha) \left(\left[\frac{\Theta_3 - \Theta_w}{\Theta_{\text{ref}} - \Theta_w} \right]^{f/2} \right) \right] \\ & + 4.374 T_{\text{air}} + 60.59 \left[\frac{w}{w_s(T_{\text{air}})} \right] \Bigg) \sigma_f + 6.116\,\text{LAI}, \end{aligned} \tag{6.27}$$

where Θ_3 is the volumetric water content ($\text{m}^3\ \text{m}^{-3}$) at 60-cm depth, T_{air} is the 9-m air temperature (K), $w_s(T_{\text{air}})$ is the saturation mixing ratio at the 9-m air temperature (kg kg^{-1}), LAI is the leaf area index, and the remaining terms are the same as those defined for Eq. (6.25). The Θ_w and Θ_{ref} terms correspond with the measured soil textures at a depth of 60 cm at each Oklahoma Mesonet site. Observations from locations where measured soil textures are unavailable at this depth do not contribute to the training or independent cross-

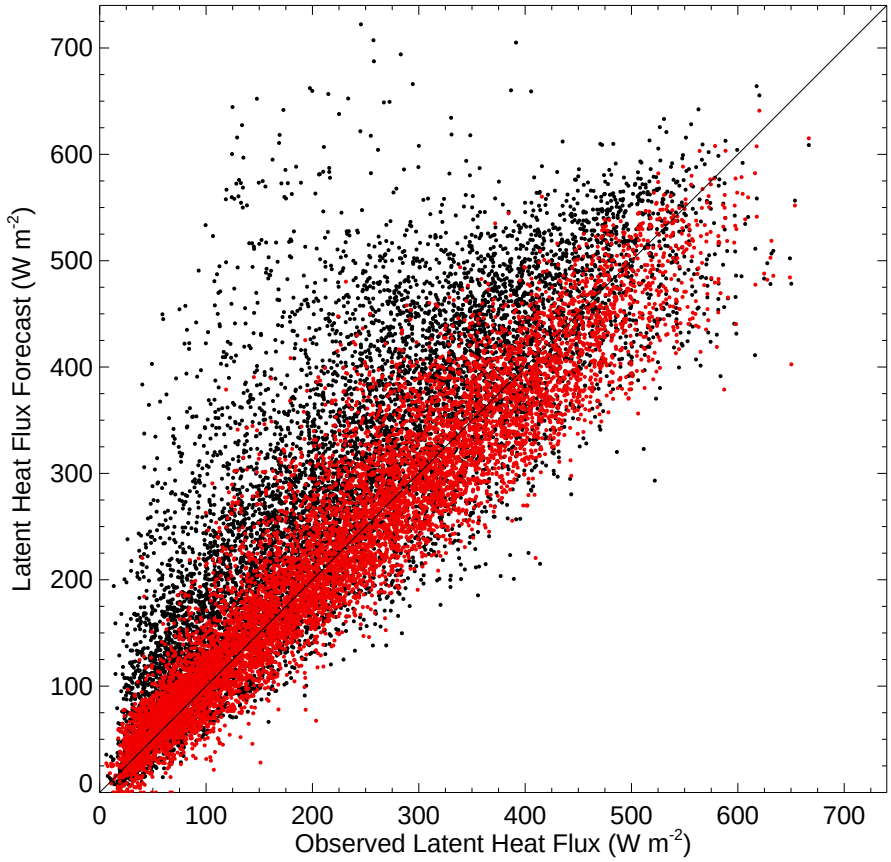

FIG. 6.3: Forecasts of total latent heat flux for 9239 forecast-observation pairs by the original Noah LSM formulation (black) and the new empirical direct soil evaporation and canopy transpiration schemes (red) compared with the observed total latent heat flux for the period 15 April–15 September 2004.

validation data. As implemented in the model, Θ_3 is the volumetric water content of the third soil layer and T_{air}, w, and w_s are the air temperature, mixing ratio, and saturation mixing ratio at the lowest model level. A large correlation for each variable corresponds with one of each of the four principal components. Therefore, the final regression equation again retains the contribution from all four principal components.

Each term in Eq. (6.27) represents a single variable in the principal component analysis. The leaf area index term arises by including LAI/σ_f as a variable. The first term describes how the root-zone soil moisture availability scales the evaporative power of the sun. This is by far the dominant term in the regression equation and its inclusion supports the results of an observational study showing a strong linear relationship between root-zone soil moisture

and both sensible and latent heat fluxes (Basara and Crawford 2002). The remaining air temperature, relative humidity, and leaf area index terms in the regression equation are less significant and may serve as tunable parameters for different locations. For this reason, the final regression equation retains these terms. Note, however, that Eq. (6.27) includes the effects of solar radiation, leaf area index, fractional vegetation coverage, vapor pressure deficit, air temperature, and soil moisture just as in the theoretical parameterization (i.e., Jacquemin and Noilhan 1990; Chen and Dudhia 2001) that appears in the original Noah LSM.

The adjusted R^2 for the independent cross-validation data is 0.72 and the residual standard error is 98.32 W m^{-2}, but recall that these numbers refer to the predictand from Eq. (6.26) and neglect the scaling by the fractional vegetation coverage. Using only the independent cross-validation data and summing the E_t forecasts from Eq. (6.27) with the E_{dir} forecasts from Eq. (6.25) to arrive at the total latent heat flux forecast, the correlation coefficient between the forecast and observed total latent heat flux is 0.94 with a residual standard error of 45.5 W m^{-2}. In contrast, the correlation coefficient between the original total latent heat flux forecasts from the Noah LSM and the observed latent heat flux for the same pool of observations is 0.83 with a residual standard error of 83.8 W m^{-2}. Combined into a single total latent heat flux term, the empirical direct soil evaporation and canopy transpiration parameterizations vastly improve the latent heat flux forecasts by the Noah LSM when driven by observations (Fig. 6.3). The original parameterization tends to overestimate latent heat fluxes, while the new parameterization corrects for this problem without introducing a negative bias.

6.5 Closure of the surface energy budget

As discussed in section 5.5.4, the Noah LSM adequately captures the sum of the latent and sensible heat fluxes when compared with observations, but fails to properly partition each. With a new parameterization for latent heat flux, the surface energy budget changes.

To force closure of the surface energy budget in the Noah LSM, one method calculates the sensible heat flux from the residual of the surface energy balance within the model. A second approach does not force closure of the surface energy budget and instead calculates the sensible heat flux from the original formula (Eq. 6.13). Tests using coupled MM5 simulations that implement the empirical direct soil evaporation and canopy transpiration schemes show that closing the surface energy budget does not significantly improve or degrade surface energy flux forecasts. Thus, the Noah LSM calculates each component of the surface energy balance individually as in the original model formulation.

Chapter 7

Results

Several MM5 simulations initialized with satellite-derived vegetation indexes and soil temperature and moisture observations from the Oklahoma Mesonet test the effectiveness of the new latent heat flux parameterizations for the direct evaporation from bare soil and the canopy transpiration. Each model simulation also allows contributions to the total latent heat flux from the original formula for wet canopy evaporation. When placed within the coupled MM5 model, the new latent heat flux parameterizations perform quite well for the same four spring and summer 2004 case studies introduced in section 5.5.1. While daytime latent heat flux forecasts improve compared with CTRL and MM5VEGSOIL forecasts, nighttime fluxes may exceed observations by nearly 50 W m^{-2}, especially shortly after sunset. Limiting the selection of predictor variables to those sets of observations measured when the incoming solar radiation exceeds 10 W m^{-2} in the principal-component regression constrains the resulting empirical formula. To overcome this limitation, when modeled downward shortwave radiation falls below 10 W m^{-2}, the latent heat flux parameterization reverts to the original canopy resistance approach.

Latent heat flux forecasts from simulations implementing the new empirical latent heat flux scheme and initialized with both satellite-derived vegetation indexes and soil temperature and moisture observations (MM5LATENT) show vast improvement over both the CTRL and MM5VEGSOIL simulations when compared with observations at Norman, Ok-

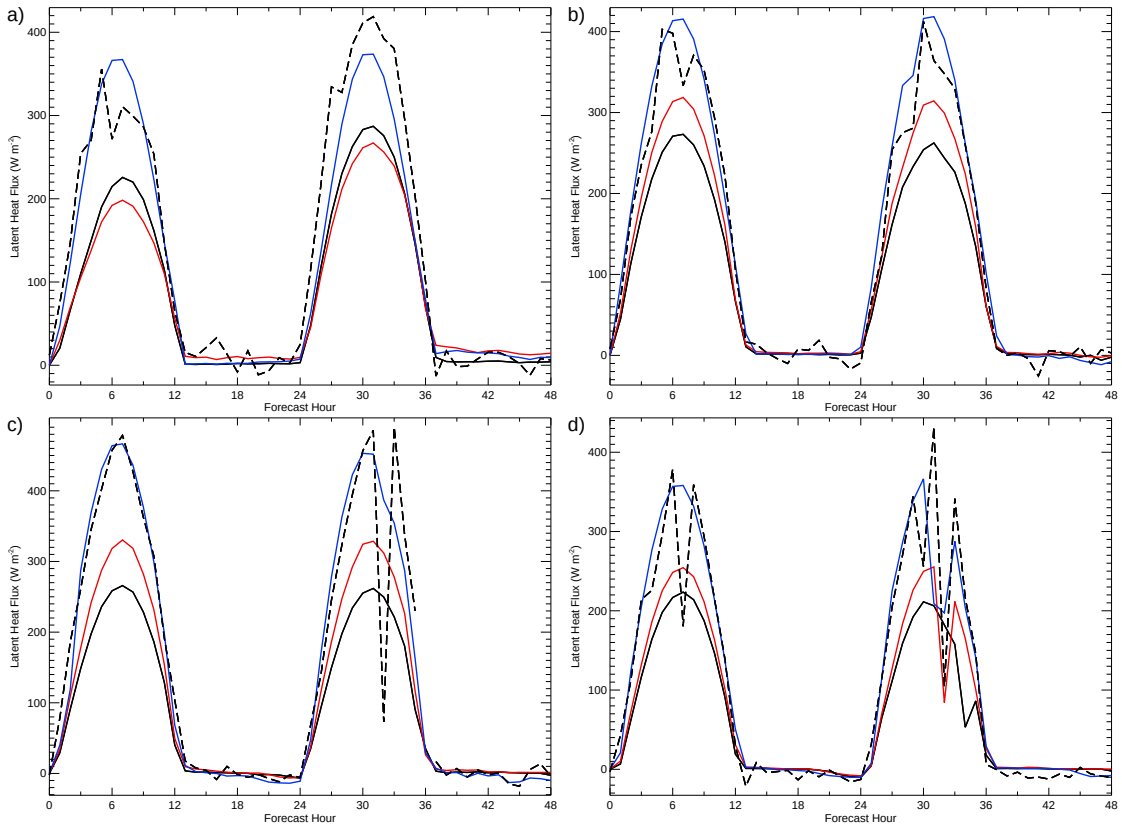

FIG. 7.1: Latent heat flux (W m^{-2}) at Norman, Oklahoma for CTRL (black), MM5VEG-SOIL (red), and MM5LATENT (blue) domain four simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004 compared with the residual of the surface energy balance computed from Oklahoma Mesonet observations (dashed).

lahoma (Fig. 7.1). In most cases, the model no longer severely underestimates daytime latent heat fluxes as in each of the other four MM5 simulations that use the original latent heat flux formulae with differing initial land surface and soil conditions. The MM5LATENT simulations consistently produce latent heat flux forecasts with domain-wide biases, root-mean squared errors, and mean absolute errors that are lower than or comparable to the error measures for the other forecasts.

With reasonable latent heat flux forecasts, the previously overestimated sensible heat flux forecasts more closely resemble the observations (Fig. 7.2), though in each case, the model still tends to underestimate the magnitude of the observed downward sensible heat flux at night. MM5LATENT ground heat flux forecasts show little to no overall improvement over CTRL forecasts. Whether the model overestimates or underestimates the ground

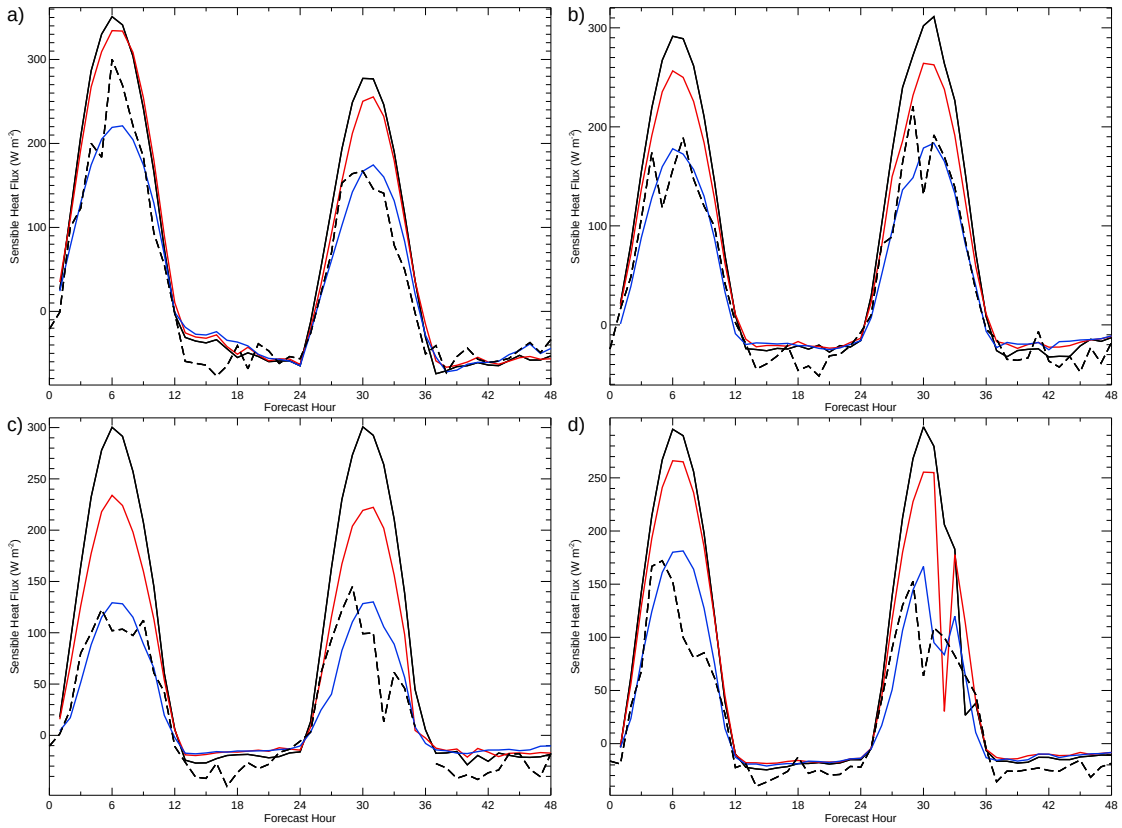

FIG. 7.2: Sensible heat flux (W m^{-2}) at Norman, Oklahoma for CTRL (black), MM5-VEGSOIL (red), and MM5LATENT (blue) domain four simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004 compared with Oklahoma Mesonet observations (dashed).

heat flux during the day varies by location, but the model typically overestimates the magnitude of the ground heat flux at night. Tests show that reducing the soil heat capacity in the MM5LATENT forecasts has a negligible effect on all forecast fields, so factors other than soil heat capacity errors are likely responsible for the poor ground heat flux estimates. The remaining errors in the partitioning between latent, sensible, and ground heat flux result in errors in the air temperature forecasts. While the empirical latent heat flux scheme improves the accuracy of temperature forecasts during the early morning, cumulative errors in the surface energy balance likely cause the air temperature to decrease too early in the diurnal cycle (Fig. 7.3). This problem appears in all forecast types. As discussed in section 5.5.4, the sharp drop in 2-m air temperature near sunset is a consequence of the extrapolation errors during planetary boundary layer regime transitions and not from surface

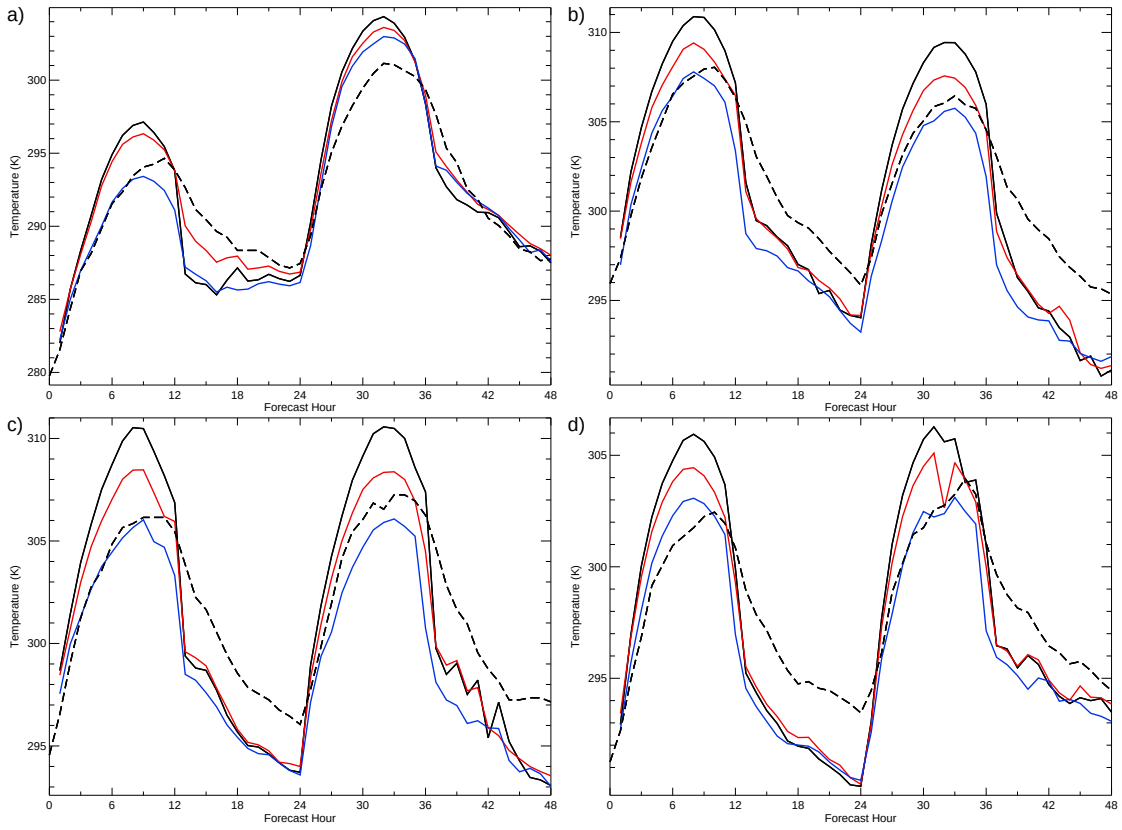

FIG. 7.3: 2-m air temperature (K) at Norman, Oklahoma for CTRL (black), MM5VEGSOIL (red), and MM5LATENT (blue) domain four simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004 compared with Oklahoma Mesonet observations (dashed).

energy flux errors.

As expected, mixing ratio forecasts improve with better latent heat flux forecasts. For these four case studies, MM5 consistently underestimates the 2-m mixing ratio, regardless of the latent heat flux parameterization or initial conditions. However, with the exception of the unrealistic spike in mixing ratio values during planetary boundary layer regime transitions, mixing ratio forecast errors decrease for the MM5LATENT simulations compared with all of the other simulations (Fig. 7.4).

Comparisons between the model and observations across the main body of Oklahoma show similar results. However, observations from the Oklahoma Mesonet in 2004 serve as the training data for the empirical latent heat flux parameterization in the Noah LSM. Two locations outside this region provide further evidence of the ability of the new la-

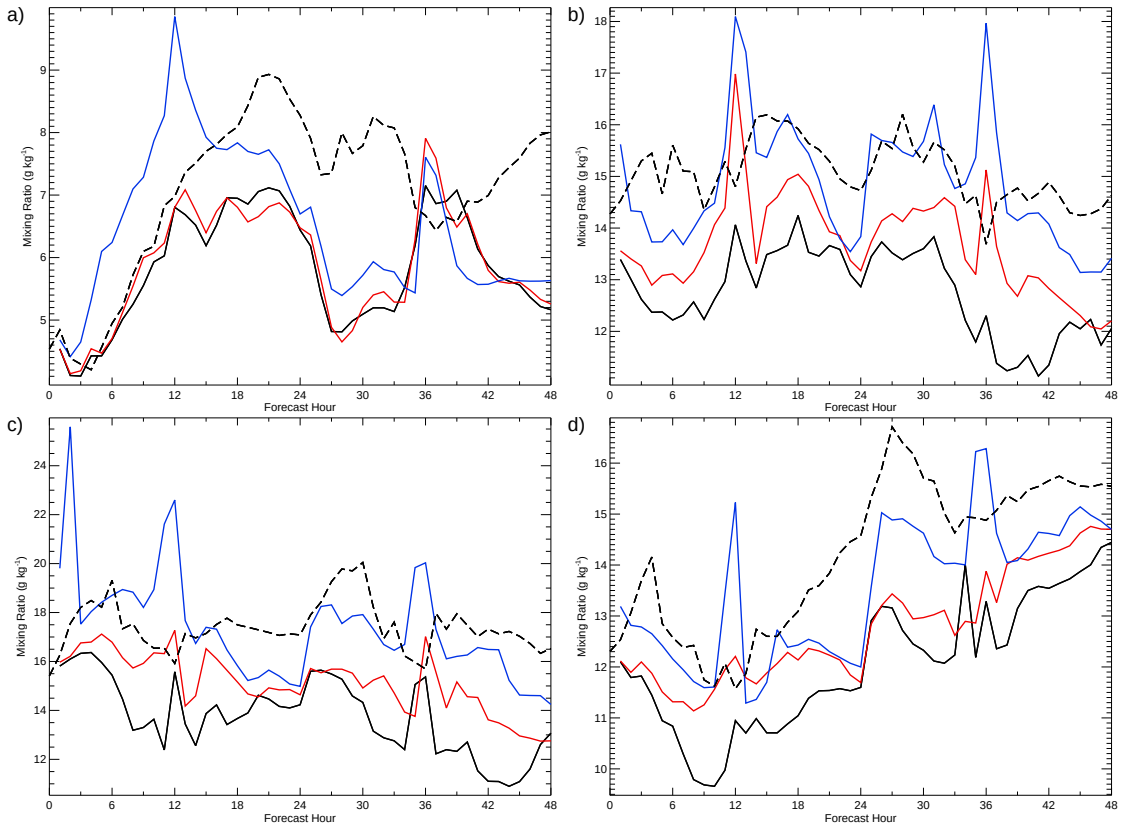

FIG. 7.4: 2-m mixing ratio (g kg^{-1}) at Norman, Oklahoma for CTRL (black), MM5-VEGSOIL (red), and MM5LATENT (blue) domain four simulations initialized at 1200 UTC on a) 3 May, b) 20 July, c) 1 August, and d) 3 September 2004 compared with Oklahoma Mesonet observations (dashed).

tent heat flux scheme to more accurately predict latent heat fluxes in short-term forecasts. Maintained by the United States Department of Agriculture (USDA) Agricultural Research Service (ARS) National Soil Tilth Laboratory (NSTL), these sites directly measure the four components of the surface energy balance using two meteorological-flux towers near Ames, Iowa. One tower stands over a soybean field and the other tower resides over a corn field. Roughly 2 m above the vegetation canopy at each location, Campbell Scientific CSAT3 sonic anemometers equipped with Campbell Scientific KH20 krypton hygrometers directly measure the sensible and latent heat flux using the eddy covariance method. REBS net radiometers measure the net radiation and REBS soil heat flow transducers measure the conductive ground heat flux at a depth of 6 cm with soil temperature probes buried at 2 and 4 cm to estimate the storage ground heat flux. Details of the instrumentation and

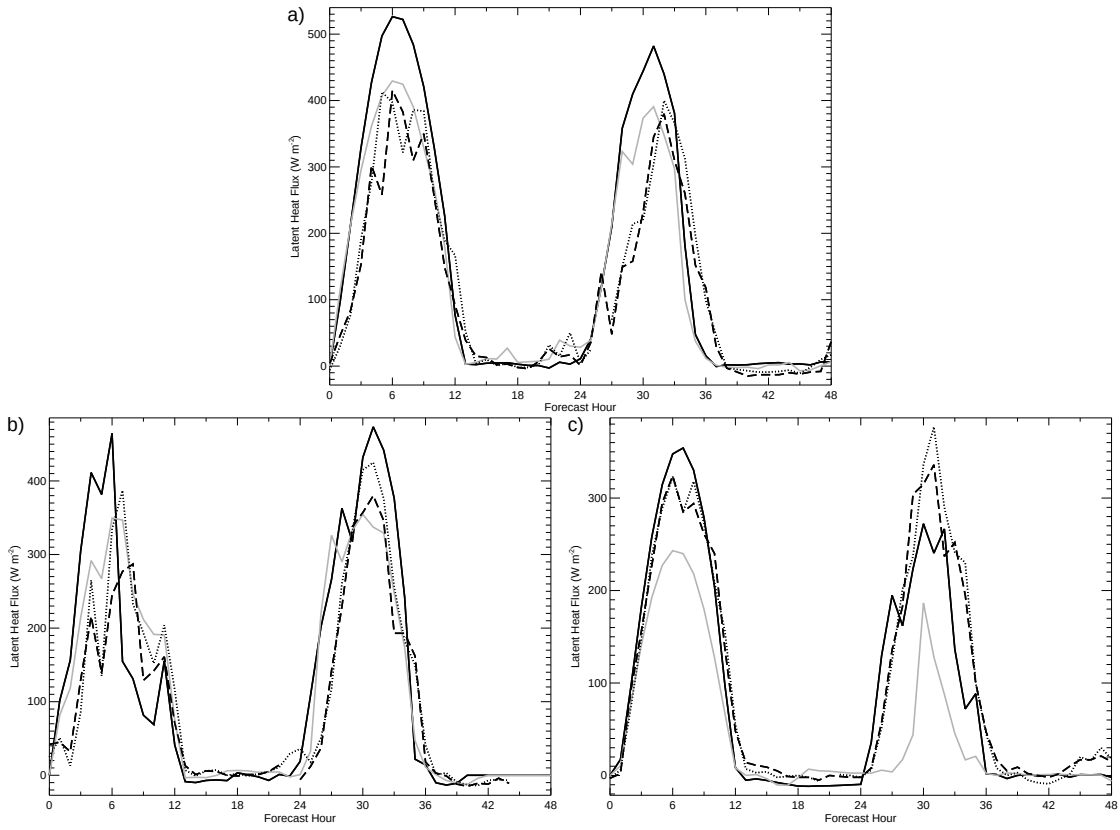

FIG. 7.5: Latent heat flux (W m^{-2}) near Ames, Iowa for CTRL (black) and MM5LATENT (gray) simulations initialized at 1200 UTC on a) 20 July, b) 1 August, and c) 3 September 2004 compared with observations of latent heat flux over a soybean field (dotted) and over a corn field (dashed).

site characteristics appear in Kustas et al. (2005). Data for the corn and soybean sites are available for the 20 July, 1 August, and 3 September 2004 case studies.

Since soil temperature and moisture observations are only available from the Oklahoma Mesonet, special initial conditions in the MM5LATENT forecasts only include satellite-derived vegetation indexes. As with the CTRL forecasts, the remaining initial conditions derive from Eta analyses. Despite lacking accurate initial soil temperature and moisture conditions, the 20 July and 3 August 2004 MM5LATENT simulations perform remarkably well compared with the latent heat fluxes measured over both corn and soybeans and reduce errors in the CTRL forecast by as much as 100 W m^{-2} (Fig. 7.5). For these three cases, the forecasts in the CTRL simulation *overestimate* rather than underestimate the observed latent heat flux as in Oklahoma, perhaps due to cloud cover, but the new empirical latent

heat flux scheme still realistically captures the total evapotranspiration at these sites. Since the gridded model results are interpolated to each flux site from a 9-km grid, the modeled fluxes over the nearly collocated corn and soybean fields are nearly identical. That fluxes measured simultaneously over the corn and soybean fields may differ by more than 100 W m^{-2} highlights the variability of surface fluxes over small spatial scales as well as the difficulty of comparing gridded model output with point measurements of atmospheric fluxes. The MM5LATENT forecast underestimates the observed latent heat fluxes at each Iowa site in the forecast initialized on 3 September 2004. The corn and soybeans were not harvested until 24 and 29 September, respectively, but irrigation practices near the time of harvest could increase the available soil moisture over the fields. This increase would not appear in the soil moisture initialization from the Eta model. Additionally, the satellite-derived fractional vegetation coverage averaged over a 9 km $\times$ 9 km forecast grid includes vegetation conditions typical for early September in Iowa and may not represent the relatively small region of photosynthetically active corn and soybean fields.

Chapter 8

Conclusions

While recent advances in numerical weather prediction models have led to improved short-term forecasts, land surface models still inaccurately portray near-surface conditions such as air temperature, mixing ratio, soil temperature and moisture, and surface energy fluxes. Assessing and reducing these model errors remains a difficult task because of both the wide variety of errors within the model and the lack of sufficient data for an accurate specification of the land surface. As others have suggested (e.g., Matsui et al. 2005), calibration of transpiration schemes within land surface models requires reliable soil and vegetation data. The availability of Oklahoma Mesonet observations of soil temperature and moisture, as well as vegetation conditions based on satellite observations, provides a unique opportunity to begin the process of improving land surface model parameterizations by initializing the model with the best possible characterization of the land surface. Indeed, for the case studies discussed here, soil moisture and vegetation conditions strongly impact model forecasts.

When observations replace climatological vegetation conditions, MM5 and its companion Noah LSM in their current state produce degraded surface energy flux forecasts when compared with control forecasts and corresponding surface observations. In one case, the peak in the diurnal cycle of latent heat flux is more than 135 W m^{-2} lower than the control forecast, which is another 130 W m^{-2} below the observed latent heat flux. Including only

soil temperature and moisture in the model initial conditions can improve sensible and latent heat flux estimates by as much as 95 W m^{-2} over control forecasts when compared with observations. Both the MM5SOIL and MM5VEG simulations show that a realistic specification of land surface variables clearly affects forecast accuracy substantially. However, the problems apparent in the simulations initialized with vegetation observations offset the improvements from the initial soil specification. Despite providing the Noah LSM with the best possible initial conditions, the model forecasts still fail to capture realistically the surface energy fluxes that drive the evolution of the planetary boundary layer. For the cases described in this study, the difference between the observed and MM5VEGSOIL latent heat fluxes may exceed 150 W m^{-2} and sensible heat flux errors may exceed 110 W m^{-2} in a 48-hour forecast period. This leads to temperature errors in excess of 2°C and mixing ratio errors that exceed 3 g kg^{-1}. That the MM5SOIL initial conditions lead to more accurate forecasts than the MM5VEGSOIL initial conditions indicates that there are problems with the physical parameterizations within the Noah LSM.

These results emphasize the significance of minimizing errors in surface initial conditions, while illustrating the profound difficulty in evaluating individual model components when all of the schemes are interdependent. Because the model physics determine the partitioning of the surface energy budget, forecast improvements for simulations with excellent soil and vegetation initial conditions require a careful calibration of many of these interdependent parameterization schemes within the Noah LSM. A new empirical parameterization determined from a wealth of unique surface, soil, and vegetation observations dramatically improves the physical representation of latent heat flux in the Noah LSM. Applying a completely new approach, this scheme replaces the usual theoretical formulations that appear in several numerical weather prediction models. For one case study, the error for the maximum daily latent heat flux falls from close to 150 W m^{-2} for the MM5VEGSOIL simulation to approximately 12 W m^{-2} using the new empirical parameterization for latent heat flux starting with the same set of initial conditions.

Despite the dramatic improvement in latent heat flux forecasts using this empirical pa-

rameterization, nighttime latent heat fluxes require yet another approach. Here, the Noah LSM reverts to the old canopy resistance formula for determining latent heat flux at night. A different empirical parameterization based on nighttime surface energy flux observations could replace the canopy resistance approach altogether.

Model simulations that combine the improved latent heat flux parameterization with the best possible characterization of the land surface show improvements in daily maximum air temperature and mixing ratio forecasts of greater than 4°C and 2 g kg^{-1} over control simulations with the same initial conditions and model formulations present in the operational version of the Noah LSM. The dominant term both in the direct evaporation from bare soil and in the transpiration equations in the empirical latent heat flux parameterization requires a measure of soil moisture. Including soil moisture alone in model initializations therefore has the potential to improve maximum daily air temperature forecasts by 2–4°C. This underscores the importance of deploying a widespread soil moisture monitoring network that, when combined with real-time satellite-derived vegetation indexes and incorporated into high-spatiotemporal resolution numerical weather prediction models, will improve short-term near-surface air temperature and moisture forecasts. To further maximize the accuracy of initial surface conditions, this soil monitoring network should measure soil conditions at multiple levels from shallow depths down through the root zone. At the very least, the new empirical latent heat flux scheme requires two soil moisture measurements at shallow and root-zone soil depths. Observations at a minimum of five soil depths would match the levels of the existing Noah LSM and would provide enhanced accuracy for ground heat flux and subsurface runoff calculations. To improve the vegetation specification, a reduced time window for calculating maximum NDVI composites spanning fewer than 14 days and a moving time window ending within 24 hours of the model initialization time would produce more representative initial σ_f and LAI values.

Unfortunately, problems remain in the predicted sensible and ground heat fluxes that are needed for a realistic representation of the surface energy balance and more accurate air temperature and moisture forecasts. With the new latent heat flux parameterization, the

Noah LSM underestimates the magnitude of ground heat fluxes by up to 70 W m^{-2} and consistently performs more poorly than even control forecasts when compared with observations. In many locations for each of the case studies, inadequate sensible and ground heat fluxes indicate persistent problems in the Noah LSM physics. Several possible explanations for this problem, if addressed, could allow for continued improvements to model forecasts. For example, the five soil layers currently in the Noah LSM may fall short of the number of soil levels required to accurately represent soil processes. In particular, rapid soil-surface drying under certain conditions affects surface air temperatures and sensible heat fluxes, but soil models require many layers near the surface with depths on the order of a few centimeters in order to capture the phenomenon (Santanello and Carlson 2001).

With more detailed observations, including soil temperature and moisture at more frequent and deeper soil depths, particularly in the root zone, and direct observations of latent heat flux, an even more robust parameterization for latent heat flux could emerge. However, the new empirical scheme improves midday latent heat flux forecasts by nearly 100 W m^{-2} for some cases in a location far from the region where the training data were collected, even with no soil data to initialize the forecast model. This likely follows from the wide range of observations in the predictor data for the multiple linear regression. The observations from Oklahoma comprise 9-m air temperatures ranging from -10.9°C to 37.7°C, relative humidities ranging from 4% to 99%, 10-m wind speeds up to 20.3 m s^{-1}, mixing ratios between 0.7 and 23.6 g kg^{-1}, 5-cm soil volumetric water contents ranging from 0.19 to 0.42 m^3 m^{-3}, and 60-cm soil volumetric water contents ranging from 0.20 to 0.38 m^3 m^{-3}. This large span of temperature, moisture, wind, and soil conditions further indicates the applicability of the new latent heat flux parameterization to new locations across the continental United States, Canada, and Mexico. However, the behavior of the new scheme remains unclear during precipitation and when the ground lies under snow cover.

This research suggests that the scientists at NCEP should consider the following issues with regard to the future development of operational forecast models. Continued improvement of the characterization of the land surface would allow further upgrades to the Noah

LSM thermodynamics and soil hydrology parameterizations and would produce even more accurate forecasts of near-surface atmospheric and soil variables. Inclusion of soil temperature and moisture observations in the NAM data assimilation system should be a primary focus, with a secondary emphasis upon daily updates of fractional vegetation coverage and leaf area index at a high spatiotemporal resolution that matches the grid resolution of the forecast model. This would require deployment of a large-scale soil monitoring network. Increasing the number of model soil layers and implementing the empirical latent heat flux scheme developed during this study would also lead to operational forecast model improvements.

Improving short-term forecasts of surface energy fluxes, which directly affects more tangible temperature and moisture variables, has many implications for agriculture, mining, transportation, finance, insurance, real estate, and other industrial operations. One subjective estimate of the impact of weather on sensitive industries indicates that nearly 40% of the United States gross domestic product is sensitive to weather and climate (Dutton 2002). While other studies calculate a substantially smaller impact on the economy (e.g., Lazo 2007), these forecasts are undoubtedly extremely important. In addition, improved surface energy flux forecasts may also influence the results of long-range climate models. Since land surface characteristics such as soil moisture influence surface weather over long time scales, and climate models typically employ the same land-surface physics as the Noah LSM, short-term forecast improvements will ultimately enhance seasonal and long-term climate predictions.

REFERENCES

Anthes, R. A., 1984: Enhancement of convective precipitation by mesoscale variations in vegetative covering in semiarid regions. *J. Climate Appl. Meteor.,* **23,** 541–554.

Arakawa, A., and V. R. Lamb, 1977: Computational design of the basic dynamical process of the UCLA general circulation model. *Methods Comput. Phys.*, **17,** 173–265.

Arya, L. M., and J. F. Paris, 1981: A physioempirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data. *Soil Sci. Soc. Amer. J.,* **45,** 1023–1030.

Avissar, R., and R. A. Pielke, 1989: A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology. *Mon. Wea. Rev.,* **117,** 2113–2136.

Barnes, S. L., 1973: Mesoscale objective analysis using weighted time-series observations. NOAA Tech. Memo. ERL NSSL-62, National Severe Storms Laboratory, Norman, OK 73069, 60 pp. [NTIS COM-73-10781.]

Basara, J. B., and T. M. Crawford, 2000: Improved installation procedures for deep-layer soil moisture measurements. *J. Atmos. Oceanic Technol.,* **17,** 879–884.

——, and K. C. Crawford, 2002: Linear relationships between root-zone soil moisture and atmospheric processes in the planetary boundary layer. *J. Geophys. Res.,* **107,** 4274, doi: 10.1029/2001JD000633.

Beard, J. S., 1949: *The Natural Vegetation of the Windward and Leeward Islands.* Oxford at the Clarendon Press, 192 pp.

Beljaars, A. C. M., P. Viterbo, M. J. Miller, and A. Betts, 1996: The anomalous rainfall over the United States during July 1993: Sensitivity to land surface parameterization and soil moisture anomalies. *Mon. Wea. Rev.*, **124,** 362–383.

Betts, A. K., J. H. Ball, A. C. M. Beljaars, M. J. Miller, and P. A. Viterbo, 1996: The land-surface-atmosphere interaction: A review based on observational and global modeling perspectives. *J. Geophys. Res.,* **101,** 7209–7226.

——, F. Chen, K. E. Mitchell, and Z. I. Janjić, 1997: Assessment of the land surface and boundary layer models in two operational versions of the NCEP Eta model using FIFE data. *Mon. Wea. Rev.,* **125,** 2896–2916.

Bhumralkar, C. M., 1975: Numerical experiments on the computation of ground surface temperature in an atmospheric general circulation model. *J. Appl. Meteor.,* **14,** 1246–1258.

Black, T. L., 1994: The new NMC mesoscale Eta model: Description and forecast examples. *Wea. Forecasting,* **9,** 265–278.

Blackadar, A. K., 1976: Modeling the nocturnal boundary layer. Preprints, *Third Symp. on Atmospheric Turbulence, Diffusion and Air Quality,* Raleigh, NC, Amer. Meteor. Soc., 46–49.

Bratley, P., B. L. Fox, and L. E. Schrage, 1987: *A Guide to Simulation.* 2d ed. Springer-Verlag, 397 pp.

Brennan, M. J., G. M. Lackmann, and S. E. Koch, 2003: An analysis of the impact of a split-front rainband on Appalachian cold-air damming. *Wea. Forecasting,* **18,** 712–731.

Bright, D. R., and S. L. Mullen, 2002: The sensitivity of the numerical simulation of the southwest monsoon boundary layer to the choice of PBL turbulence parameterization in MM5. *Wea. Forecasting,* **17,** 99–114.

Brock, F. V., K. C. Crawford, R. L. Elliot, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, 1995: The Oklahoma Mesonet: A technical overview. *J. Atmos. Oceanic Technol.*, **12,** 5–19.

Brotzge, J. A., 2000: Closure of the surface energy budget. Ph.D. dissertation, University of Oklahoma, Norman, OK, 208 pp.

——, 2004: A two-year comparison of the surface water and energy budgets between two OASIS sites and NCEP–NCAR reanalysis data. *J. Hydrometeor.*, **5,** 311–326.

——, and K. C. Crawford, 2000: Estimating sensible heat flux from the Oklahoma Mesonet. *J. Appl. Meteor.,* **39,** 102–116.

——, and C. E. Duchon, 2000: A field comparison among a domeless net radiometer, two four-component net radiometers, and a domed net radiometer. *J. Atmos. Oceanic Technol.,* **17,** 1569–1582.

——, and K. C. Crawford, 2003: Examination of the surface energy budget: A comparison of eddy correlation and Bowen ratio measurement systems. *J. Hydrometeor.,* **4,** 160–178.

——, S. J. Richardson, K. C. Crawford, T. W. Horst, F. V. Brock, K. S. Humes, Z. Sorbjan, and R. L. Elliot, 1999: The Oklahoma atmospheric surface-layer instrumentation system (OASIS) project. Preprints, *13th Symp. on Boundary Layers and Turbulence,* Dallas, TX, Amer. Meteor. Soc., 612–615.

Chang, J.-T., and P. J. Wetzel, 1991: Effects of spatial variations of soil moisture and vegetation on the evolution of a prestorm environment: A numerical case study. *Mon. Wea. Rev.,* **119,** 1368–1390.

Chen, D., and W. Brutsaert, 1995: Diagnostics of land surface spatial variability and water vapor flux. *J. Geophys. Res.,* **100,** 25 595–25 606.

Chen, F., and R. Avissar, 1994a: The impact of land-surface wetness heterogeneity on mesoscale heat fluxes. *J. Appl. Meteor.,* **33,** 1323–1340.

——, and ——, 1994b: Impact of land-surface moisture variability on local shallow convective cumulus and precipitation in large-scale models. *J. Appl. Meteor.,* **33,** 1382–1401.

——, and J. Dudhia, 2001: Coupling an advanced land surface-hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.,* **129,** 569–585.

——, K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q. Y. Duan, M. Ek, and A. Betts, 1996: Modeling of land-surface evaporation by four schemes and comparison with FIFE observations. *J. Geophys. Res.*, **101,** 7251–7268.

Chen, Y., F. L. Ludwig, and R. L. Street, 2004: Stably stratified flows near a notched transverse ridge across the Salt Lake valley. *J. Appl. Meteor.,* **43,** 1308–1328.

Clapp, R. B., and G. M. Hornberger, 1978: Empirical equations for some soil hydraulic properties. *Water Resour. Res.,* **14,** 601–604.

Clark, C. A., and R. W. Arritt, 1995: Numerical simulations of the effect of soil moisture and vegetation cover on the development of deep convection. *J. Appl. Meteor.*, **34,** 2029–2045.

Cleveland, W. S., and S. J. Devlin, 1988: Locally weighted regression: An approach to regression analysis by local fitting. *J. Amer. Statist. Assoc.*, **83**, 596–610.

Colle, B. A., C. F. Mass, and D. Ovens, 2001: Evaluation of the timing and strength of MM5 and Eta surface trough passages over the eastern Pacific. *Wea. Forecasting,* **16,** 553–572.

Cosby, B. J., G. M. Hornberger, R. B. Clapp, and T. R. Ginn, 1984: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.,* **20,** 682–690.

Crawford, K. C., and G. R. Essenberg, 2006: Coop modernization: NOAA's Environmental Real-time Observation Network in New England, the Southeast, and addressing NIDIS in the West. Preprints, *10th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface,* Atlanta, GA, Amer. Meteor. Soc., CD-ROM, J5.9.

Crawford, T. M., D. J. Stensrud, T. N. Carlson, and W. J. Capehart, 2000: Using a soil hydrology model to obtain regionally averaged soil moisture values. *J. Hydrometeor.,* **1,** 353–363.

——, ——, F. Mora, J. W. Merchant, and P. J. Wetzel, 2001: Value of incorporating satellite-derived land cover data in MM5/PLACE for simulating surface temperatures. *J. Hydrometeor.,* **2,** 453–468.

Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.,* **124,** 1767–1785.

Curran, P. J., 1983: Multispectral remote sensing for the estimation of green leaf area index. *Phil. Trans. R. Soc. Lond.,* **309,** 257–270.

Deardorff, J. W., 1978: Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation. *J. Geophys. Res.,* **83,** 1889–1903.

DiMego, G. J., and E. Rogers, cited 2005: Spring 2005 upgrade package for North American Mesoscale (NAM) decision brief. [Available online at http://wwwt.emc.ncep.noaa.gov/mmb/Spring2005.NAMUpgrade.pdf.]

Dirmeyer, P. A., F. J. Zeng, A. Ducharne, J. C. Morrill, and R. D. Koster, 2000: The sensitivity of surface fluxes to soil water content in three land surface schemes. *J. Hydrometeor.,* **1,** 121–134.

Doran, J. C., W. J. Shaw, and J. M. Hubbe, 1995: Boundary layer characteristics over areas of inhomogeneous surface fluxes. *J. Appl. Meteor.,* **34,** 559–571.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.,* **46,** 3077–3107.

——, 1993: A nonhydrostatic version of the Penn State–NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.,* **121,** 1493–1513.

——, 1996: A multi-layer soil temperature model for MM5. Preprints, *Sixth PSU/NCAR Mesonet Model Users' Workshop,* Boulder, CO, NCAR, 49–50.

——, 2003: MM5 model status and plans. Preprints, *13th PSU/NCAR Mesoscale Model Users' Workshop,* Boulder, CO, NCAR, 1–2.

Durre, I., and J. M. Wallace, 2001: The warm season dip in diurnal temperature range over the eastern United States. *J. Climate,* **14,** 354–360.

Dutton, J. A., 2002: Opportunities and priorities in a new era for weather and climate services. *Bull. Amer. Meteor. Soc.,* **83,** 1303–1311.

Dyer, A. J., 1974: A review of flux-profile relationships. *Bound.-Layer Meteor.,* **7,** 363–372.

Ek, M. B., and L. Mahrt, 1991: OSU 1-D PBL model user's guide. Version 1.0.4, 118 pp. [Available from Dept. of Atmospheric Sciences, Oregon State University, Corvallis, OR 97331-2209.]

——, and A. A. M. Holtslag, 2004: Influence of soil moisture on boundary layer cloud development. *J. Hydrometeor.,* **5,** 86–99.

——, K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.,* **108,** 8851, doi: 10.1029/2002JD003296.

Entekhabi, D., G. R. Asrar, A. K. Betts, K. J. Beven, R. L. Bras, C. J. Duffy, T. Dunne, R. D. Koster, D. P. Lettenmaier, D. B. McLaughlin, W. J. Shuttleworth, M. T. van Genuchten, M.-Y. Wei, and E. F. Wood, 1999: An agenda for land surface hydrology research and a call for the second international hydrological decade. *Bull. Amer. Meteor. Soc.,* **80,** 2043–2058.

Entin, J. K., A. Robock, K. Y. Vinnikov, S. E. Hollinger, S. Liu, and A. Namkhai, 2000: Temporal and spatial scales of observed soil moisture variations in the extratropics. *J. Geophys. Res.,* **105,** 11 865–11 877.

Fennessy, M. J., and J. Shukla, 1999: Impact of initial soil wetness on seasonal atmospheric prediction. *J. Climate,* **12,** 3167–3180.

Fiebrich, C. A., and K. C. Crawford, 2001: The impact of unique meteorological phenomena detected by the Oklahoma Mesonet and ARS Micronet on automated quality control. *Bull. Amer. Meteor. Soc.,* **82,** 2173–2187.

——, J. E. Martinez, J. A. Brotzge, and J. B. Basara, 2003: The Oklahoma Mesonet's skin temperature network. *J. Atmos. Oceanic Technol.,* **20,** 1496–1504.

Fulton, R. A., J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Wea. Forecasting,* **13,** 377–395.

Galewsky, J., and A. Sobel, 2005: Moist dynamics and orographic precipitation in northern and central California during the New Year's flood of 1997. *Mon. Wea. Rev.,* **133,** 1594–1612.

Gannon, P. T., 1978: Influences of earth surface and cloud properties in the south Florida sea breeze. NOAA Tech. Rep. ERL402-NHELM2, 91 pp. [NTIS PB-297398.]

Gao, X., S. Sorooshian, and H. V. Gupta, 1996: A sensitivity analysis of the Biosphere-Atmosphere Transfer Scheme (BATS). *J. Geophys. Res.,* **101,** 7279–7289.

Garrett, A. J., 1982: A parameter study of interactions between convective clouds, the convective boundary layer, and a forested surface. *Mon. Wea. Rev.,* **110,** 1041–1059.

Goodrum, G., K. B. Kidwell, and W. Winston, 2001: NOAA KLM User's Guide. National Oceanic and Atmospheric Administration/National Environmental Satellite, Data, and Information Service, 1414 pp. [Available from NOAA/NESDIS National Climatic Data Center, Satellite Data Services Division, Federal Office Building #3, Washington, DC 20233.].

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1995: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR/TN-398+STR, 122 pp. [Available from MMM Division, NCAR, P.O. Box 3000, Boulder, CO 80307.]

Gutman, G., and A. Ignatov, 1998: The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *Int. J. Remote Sens.,* **19,** 1533–1543.

——, D. Tarpley, A. Ignatov, and S. Olson, 1995: The enhanced NOAA global land dataset from the Advanced Very High Resolution Radiometer. *Bull. Amer. Meteor. Soc.,* **76,** 1141–1156.

Haan, C. T., 1977: *Statistical methods in hydrology.* The Iowa State University Press, 378 pp.

Harnack, R. P., 1979: A further assessment of winter temperature predictions using objective methods. *Mon. Wea. Rev.,* **107,** 250–267.

Hart, K. A., W. J. Steenburgh, D. J. Onton, and A. J. Siffert, 2004: An evaluation of mesoscale-model-based model output statistics (MOS) during the 2002 Olympic and Paralympic Winter Games. *Wea. Forecasting,* **19,** 200–218.

Haugland, M. J., and K. C. Crawford, 2002: The diurnal cycle of dewpoint across Oklahoma's winter wheat belt. Preprints, *13th Symp. on Global Climate Change and Climate Variations,* Orlando, FL, Amer. Meteor. Soc., 254–256.

Hoadley, J. L., K. Westrick, S. A. Ferguson, S. L. Goodrick, L. Bradshaw, and P. Werth, 2004: The effect of model resolution in predicting meteorological parameters used in fire danger rating. *J. Appl. Meteor.,* **43,** 1333–1347.

Holtslag, A. A. M., and M. Ek, 1996: Simulation of surface fluxes and boundary layer development over the pine forest in HAPEX-MOBILHY. *J. Appl. Meteor.,* **35,** 202–213.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan, 2004: Development of a 2001 national land-cover database for the United States. *Photogramm. Eng. Remote Sens.,* **70,** 829–840.

Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.,* **124,** 2322-2339.

Jacquemin, B., and J. Noilhan, 1990: Sensitivity study and validation of a land surface parameterization using the HAPEX-MOBILHY data set. *Bound.-Layer Meteor.,* **52,** 93–134.

Jarvis, P. G., 1976: The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Philos. Trans. Roy. Soc. London,* **273B,** 593–610.

Jin, M., and D.-L. Zhang, 2002: Observed variations of leaf area index and its relationship with surface temperatures during warm seasons. *Meteor. Atmos. Phys.,* **80,** 117–129.

Justice, C. O., E. Vermote, J. R. G. Townshend, R. Defries, D. P. Roy, D. K. Hall, V. V. Salomonson, J. L. Privette, G. Riggs, A. Strahler, W. Lucht, R. B. Myneni, Y. Knyazikhin, S. W. Running, R. R. Nemani, Z. Wan, A. R. Huete, W. van Leeuwen, R. E. Wolfe, L. Giglio, J.-P. Muller, P. Lewis, and M. J. Barnsley, 1998: The moderate resolution imaging spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Trans. Geosci. Remote Sens.,* **36,** 1228–1249.

Kain, J. S., and J. M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.,* No. 46, Amer. Meteor. Soc., 165–170.

Kaiser, H. F., 1958: The varimax criterion for analytic rotation in factor analysis. *Psychometrika,* **23,** 187–200.

Kidwell, K. B., 1998: NOAA polar orbiter data user's guide. National Oceanic and Atmospheric Administration/National Environmental Satellite, Data, and Information Service, 469 pp. [Available from NOAA/NESDIS National Climatic Data Center, Satellite Data Services Division, Federal Office Building #3, Washington, DC 20233.].

Koren, V., J. Schaake, K. Mitchell, Q.-Y. Duan, F. Chen, and J. M. Baker, 1999: A parameterization of snowpack and frozen ground intended for NCEP weather and climate models. *J. Geophys. Res.,* **104,** 19 569–19 585.

Koster, R. D., P. A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, C. T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, C.-H. Lu, S. Malyshev, B. McAvaney, K. Mitchell, D. Mocko, T. Oki, K. Oleson, A. Pitman, Y. C. Sud, C. M. Taylor, D. Verseghy, R. Vasic, Y. Xue, and T. Yamada, 2004a: Regions of strong coupling between soil moisture and precipitation. *Science,* **305,** 1138–1140.

Koster, R. D., M. J. Suarez, P. Liu, U. Jambor, A. Berg, M. Kistler, R. Reichle, M. Rodell, and J. Famiglietti, 2004b: Realistic initialization of land surface states: Impacts on subseasonal forecast skill. *J. Hydrometeor.,* **5,** 1049–1063.

Kurkowski, N. P., D. J. Stensrud, and M. E. Baldwin, 2003: Assessment of implementing satellite-derived land cover data in the Eta model. *Wea. Forecasting,* **18,** 404–416.

Kustas, W. P., J. L. Hatfield, and J. H. Prueger, 2005: The Soil Moisture–Atmosphere Coupling Experiment (SMACEX): Background, hydrometeorological conditions, and preliminary findings. *J. Hydrometeor.,* **6,** 791–804.

Lazo, J. K., P. Larsen, and D. Waldman, 2007: Sensitivity of the United States economy to weather variability. Preprints, *Second Symp. on Policy and Socio-Economic Research,* San Antonio, TX, Amer. Meteor. Soc., CD-ROM, 1.4.

Leese, J., T. Jackson, A. Pitman, and P. Dirmeyer, 2001: GEWEX/BAHC International Workshop on Soil Moisture Monitoring, Analysis, and Prediction for Hydrometeorological and Hydroclimatological Applications. *Bull. Amer. Meteor. Soc.,* **82,** 1423–1430.

Li, B., and R. Avissar, 1994: The impact of spatial variability of land-surface characteristics on land-surface heat fluxes. *J. Climate,* **7,** 527–537.

Lin, Y., K. E. Mitchell, E. Rogers, and G. J. DiMego, 2005: Using hourly and daily precipitation analyses to improve model water budget. Preprints, *Ninth Symp. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface,* San Diego, CA, Amer. Meteor. Soc., CD-ROM, 3.3.

Liu, Y., and R. Avissar, 1999a: A study of persistence in the land-atmosphere system using a general circulation model and observations. *J. Climate,* **12,** 2139–2153.

——, and ——, 1999b: A study of persistence in the land-atmosphere system with a fourth-order analytical model. *J. Climate,* **12,** 2154–2168.

——, D. Z. Ye, and J. J. Ji, 1993: Influence of soil moisture and vegetation on climate. Part II: Numerical experiments on persistence of short-term climatic anomalies. *Sci. China,* **36B,** 102–109.

Maddox, R. A., 1980: An objective technique for separating macroscale and mesoscale features in meteorological data. *Mon. Wea. Rev.,* **108,** 1108–1121.

Mahfouf J.-F., and J. Noilhan, 1991: Comparative study of various formulations of evaporation from bare soil using in situ data. *J. Appl. Meteor.,* **30,** 1354–1365.

——, E. Richard, and P. Mascart, 1987: The influence of soil and vegetation on the development of mesoscale circulations. *J. Climate Appl. Meteor.,* **26,** 1483–1495.

Mahrt, L., and M. Ek, 1984: The influence of atmospheric stability on potential evaporation. *J. Climate Appl. Meteor.,* **23,** 222–234.

——, and H. L. Pan, 1984: A two-layer model of soil hydrology. *Bound.-Layer Meteor.,* **29,** 1–20.

Marshall, C. H., K. C. Crawford, K. E. Mitchell, and D. J. Stensrud, 2003: The impact of the land surface physics in the operational NCEP Eta model on simulating the diurnal cycle: Evaluation and testing using Oklahoma Mesonet data. *Wea. Forecasting,* **18,** 748–768.

Marshall, T. J., J. W. Holmes, and C. W. Rose, 1996: *Soil Physics.* 3d ed. Cambridge University Press, 453 pp.

Mass, C. F., and Y.-H. Kuo, 1998: Regional real-time numerical weather prediction: Current status and future potential. *Bull. Amer. Meteor. Soc.,* **79,** 253–263.

Matsui, T., V. Lakshmi, and E. E. Small, 2005: The effects of satellite-derived vegetation cover variability on simulated land-atmosphere interactions in the NAMS. *J. Climate,* **18,** 21–40.

McCumber, M. C., and R. A. Pielke, 1981: Simulation of the effects of surface fluxes of heat and moisture in a mesoscale numerical model. *J. Geophys. Res.,* **86,** 9929–9938.

McPherson, R. A., D. J. Stensrud, and K. C. Crawford, 2004: The impact of Oklahoma's winter wheat belt on the mesoscale environment. *Mon. Wea. Rev.,* **132,** 405–421.

Miller, D. A., and R. A. White, 1998: A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. *Earth Interactions,* **2,** 1–26.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.,* **102,** 16 663–16 682.

Namias, J., 1952: The annual course of month-to-month persistence in climatic anomalies. *Bull. Amer. Meteor. Soc.,* **33,** 279–285.

——, 1959: Persistence of mid-tropospheric circulations between adjacent months and seasons. *The Atmosphere and the Sea in Motion: Scientific Contributions to the Rossby Memorial Volume,* B. Bolin, Ed., The Rockefeller Institute Press and Oxford University Press, 240–248.

National Imagery and Mapping Agency, 2000: Department of Defense World Geodetic System 1984: Its Definition and Relationships with Local Geodetic Systems. Tech. Rep. TR8350.2, 3d. ed., amendment 1, 3 January 2000, Bethesda, MD, 175 pp. [Available from National Imagery and Mapping Agency, ATTN: ISDFR, Mail Stop D-17, 4600 Sangamore Road, Bethesda, MD 20816-5003.]

Nelson, J. A., 1999: The Eta Data Assimilation System. WR Tech. Attachment 99-14, 6 pp. [Available from National Weather Service Western Region, P.O. Box 11188, Salt Lake City, UT 84147.]

Noilhan, J., and S. Planton, 1989: A simple parameterization of land-surface processes for meteorological models. *Mon. Wea. Rev.,* **117,** 536–549.

Oleson, K. W., and G. B. Bonan, 2000: The effects of remotely sensed plant functional type and leaf area index on simulations of boreal forest surface fluxes by the NCAR land surface model. *J. Hydrometeor.,* **1,** 431–446.

Ookouchi, Y., M. Segal, R. C. Kessler, and R. A. Pielke, 1984: Evaluation of soil moisture effects on the generation and modification of mesoscale circulations. *Mon. Wea. Rev.,* **112,** 2281–2292.

Pan, H. L., and L. Mahrt, 1987: Interaction between soil hydrology and boundary-layer development. *Bound.-Layer Meteor.,* **38,** 185–202.

Paulson, C. A., 1970: The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *J. Appl. Meteor.,* **9,** 857–861.

Penman, H. L., 1948: Natural evaporation from open water, bare soil and grass. *Proc. Roy. Soc. London,,* **193A,** 120–145.

Peters-Lidard, C. D., E. Blackburn, X. Liang, and E. F. Wood, 1998: The effect of soil thermal conductivity parameterization on surface energy fluxes and temperatures. *J. Atmos. Sci.,* **55,** 1209–1224.

Pielke, R. A., and X. Zeng, 1989: Influence on severe storm development of irrigated land. *Natl. Wea. Dig.,* **14,** 16–17.

——, G. A. Dalu, J. S. Snook, T. J. Lee, and T. G. F. Kittel, 1991: Nonlinear influence of mesoscale land use on weather and climate. *J. Climate,* **4,** 1053–1069.

——, G. E. Liston, J. L. Eastman, L. Lu, and M. Coughenour, 1999: *J. Geophys. Res.,* **104,** 19 463–19 479.

Pinty, J.-P., P. Mascart, E. Richard, and R. Rosset, 1989: An investigation of mesoscale flows induced by vegetation inhomogeneities using an evapotranspiration model calibrated against HAPEX-MOBILHY data. *J. Appl. Meteor.,* **28,** 976–992.

Pryor, S. C., I. G. McKendry, and D. G. Steyn, 1995: Synoptic-scale meteorological variability and surface ozone concentrations in Vancouver, British Columbia. *J. Appl. Meteor.,* **34,** 1824–1833.

Rabin, R. M., S. Stadler, P. J. Wetzel, D. J. Stensrud, and M. Gregory, 1990: Observed effects of landscape variability on convective clouds. *Bull. Amer. Meteor. Soc.,* **71,** 272–280.

Rawls, W. J., D. L. Brakensiek, and K. E. Saxton, 1982: Estimation of soil water properties. *Trans. Amer. Soc. Agric. Eng.,* **25,** 1316–1320.

Reece, C. F., 1996: Evaluation of a line heat dissipation sensor for measuring soil matric potential. *Soil Sci. Soc. Amer. J.,* **60,** 1022–1028.

Ren, D., M. Xue, and A. Henderson-Sellers, 2004: Incorporating hydraulic lift into a land surface model and its effects on surface soil moisture prediction. *J. Hydrometeor.,* **5,** 1181–1191.

Richman, M. B., 1986: Rotation of principal components. *J. Climatol.,* **6,** 293–335.

Rind, D., 1982: The influence of ground moisture conditions in North America on summer climate as modeled in the GISS GCM. *Mon. Wea. Rev.,* **110,** 1487–1494.

Robock, A., K. Y. Vinnikov, G. Srinivasan, J. K. Entin, S. E. Hollinger, N. A. Speranskaya, S. Liu, and A. Namkhai, 2000: The global soil moisture data bank. *Bull. Amer. Meteor. Soc.,* **81,** 1281–1299.

——, L. Luo, E. F. Wood, F. Wen, K. E. Mitchell, P. R. Houser, J. C. Schaake, D. Lohmann, B. Cosgrove, J. Sheffield, Q. Duan, R. W. Higgins, R. T. Pinker, J. D. Tarpley, J. B. Basara, and K. C. Crawford, 2003: Evaluation of the North American Land Data Assimilation System over the southern Great Plains during the warm season. *J. Geophys. Res.,* **108,** 8846, doi:10.1029/2002JD003245.

Rodell, M., P. R. Houser, A. A. Berg, and J. S. Famiglietti, 2005: Evaluation of 10 methods for initializing a land surface model. *J. Hydrometeor.,* **6,** 146–155.

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

Rowntree, P. R., and J. A. Bolton, 1983: Simulation of the atmospheric response to soil moisture anomalies over Europe. *Quart. J. Roy. Meteor. Soc.,* **109,** 501–526.

Santanello, J. A., and T. N. Carlson, 2001: Mesoscale simulation of rapid soil drying and its implications for predicting daytime temperature. *J. Hydrometeor.,* **2,** 71–88.

Schlatter, T. W., 1975: Some experiments with a multivariate statistical objective analysis scheme. *Mon Wea. Rev.,* **103,** 246–257.

Scurlock, J. M. O., G. P. Asner, and S. T. Gower, 2001: Worldwide historical estimates of leaf area index, 1932–2000. Technical Memorandum ORNL/TM-2001/268. Oak Ridge National Laboratory, Oak Ridge, TN, 40 pp.

Segal, M., and R. W. Arritt, 1992: Nonclassical mesoscale circulations caused by surface sensible heat-flux gradients. *Bull. Amer. Meteor. Soc.,* **73,** 1593–1604.

——, R. Avissar, M. C. McCumber, and R. A. Pielke, 1988: Evaluation of vegetation effects on the generation and modification of mesoscale circulations. *J. Atmos. Sci.,* **45,** 2268–2292.

——, W. E. Schreiber, G. Kallos, J. R. Garratt, A. Rodi, J. Weaver, and R. A. Pielke, 1989: The impact of crop areas in northeast Colorado on midsummer mesoscale thermal circulations. *Mon. Wea. Rev.,* **117,** 809–825.

Segele, Z. T., D. J. Stensrud, I. C. Ratcliffe, and G. M. Henebry, 2005: Influence of a hailstreak on boundary layer evolution. *Mon. Wea. Rev.,* **133,** 942–960.

Sellers, P. J., Y. Mintz., Y. C. Sud, and A. Dalcher, 1986: A simple biosphere model (SiB) for use within general circulation models. *J. Atmos. Sci.,* **43,** 505–531.

Seuffert, G., H. Walker, P. Viterbo, M. Drusch, and J.-F. Mahfouf, 2004: The usage of screen-level parameters and microwave brightness temperature for soil moisture analysis. *J. Hydrometeor.,* **5,** 516–531.

Shafer, M. A., C. A. Fiebrich, D. S. Arndt, S. E. Fredrickson, and T. W. Hughes, 2000: Quality assurance procedures in the Oklahoma Mesonetwork. *J. Atmos. Sci.,* **17,** 474–494.

Shaw, B., and L. D. Baver, 1939: Heat conductivity as an index of soil moisture. *J. Amer. Soc. Agron.,* **31,** 886–891.

Silberstein, R. P., M. Sivapalan, and A. Wyllie, 1999: On the validation of a coupled water and energy balance model at small catchment scales. *J. Hydrology,* **220,** 149–168.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the advanced research WRF version 2. NCAR/TN-468+STR, 88 pp. [Available from MMM Division, NCAR, P.O. Box 3000, Boulder, CO 80307.]

Smith, C. B., M. N. Lakhtakia, W. J. Capehart, and T. N. Carlson, 1994: Initialization of soil-water content in regional-scale atmospheric prediction models. *Bull. Amer. Meteor. Soc.,* **74,** 585–593.

Smith, E. A., M. M.-K. Wai, H. J. Cooper, M. T. Rubes, and A. Hsu, 1994: Linking boundary-layer circulations and surface processes during FIFE 89. Part I: Observational analysis. *J. Atmos. Sci.,* **51,** 1497–1529.

Sodano, E. M., 1965: General non-iterative solution of the inverse and direct geodetic problems. *Bull. Géod.,* **75,** 69–89.

Stensrud, D. J., and S. J. Weiss, 2002: Mesoscale model ensemble forecasts of the 3 May 1999 tornado outbreak. *Wea. Forecasting,* **17,** 526–543.

——, N. Yussouf, M. E. Baldwin, J. T. McQueen, J. Du, B. Zhou, B. Ferrier, G. Manikin, F. M. Ralph, J. M. Wilczak, A. B. White, I. Djlalova, J.-W. Bao, R. J. Zamora, S. G. Benjamin, P. A. Miller, T. L. Smith, T. Smirnova, and M. F. Barth, 2006: The New England high-resolution temperature program. *Bull. Amer. Meteor. Soc.,* **87,** 491–498.

Sud, Y. C., W. C. Chao, and G. K. Walker, 1993: Dependence of rainfall on vegetation: Theoretical considerations, simulation experiments, observations, and inferences from simulated atmospheric soundings. *J. Arid. Environ.,* **25,** 5–18.

Sun, W.-Y., and Y. Ogura, 1979: Boundary-layer forcing as a possible trigger to a squall-line formation. *J. Atmos. Sci.,* **36,** 235–254.

Troen, I., and L. Mahrt, 1986: A simple model of the boundary layer: Sensitivity to surface evaporation. *Bound.-Layer Meteor.,* **37,** 129–148.

Tsonis, A. A., 2002: The problem of extracting precipitation information in the tropics from the UWM/COADS data. *J. Appl. Meteor.,* **41,** 1153–1162.

van Genuchten, M. T., 1980: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Amer. J.,* **44,** 892–898.

Viña, A., G. M. Henebry, and A. A. Gitelson, 2004: Satellite monitoring of vegetation dynamics: Sensitivity enhancement by the wide dynamic range vegetation index. *Geophys. Res. Lett.,* **31,** L04503, doi:10.1029/2003GL019034.

Vinnikov, K. Y., and I. B. Yeserkepova, 1991: Soil moisture: Empirical data and model results. *J. Climate,* **4,** 66–79.

——, A. Robock, N. A. Speranskaya, and C. A. Schlosser, 1996: Scales of temporal and spatial variability of midlatitude soil moisture. *J. Geophys. Res.,* **101,** 7163–7174.

Viterbo, P., and A. C. M. Beljaars, 1995: An improved land surface parameterization scheme in the ECMWF model and its validation. *J. Climate,* **8,** 2716–2748.

——, and A. K. Betts, 1999: Impact of the ECMWF reanalysis soil water on forecasts of the July 1993 Mississippi flood. *J. Geophys. Res.,* **104,** 19 361–19 366.

Walker, J. M., and P. R. Rowntree, 1977: The effect of soil moisture on circulation and rainfall in a tropical model. *Quart. J. Roy. Meteor. Soc.,* **103,** 29–46.

Walsh, J. E., W. H. Jasperson, and B. Ross, 1985: Influence of snow cover and soil moisture on monthly air temperature. *Mon. Wea. Rev.,* **113,** 756–768.

Walter-Shea, E. A., B. L. Blad, C. J. Hays, M. A. Mesarch, D. W. Deering, and E. M. Middleton, 1992: Biophysical properties affecting vegetative canopy reflectance and absorbed photosynthetically active radiation at the FIFE site. *J. Geophys. Res.,* **97,** 18 925–18 934.

Wei, M.-Y., Ed., 1995: Soil Moisture: Report of a Workshop Held in Tiburon, California, 25–27 January 1994. NASA Conference Publication 3319, 80 pp.

Westrick, K. J., P. Storck, and C. F. Mass, 2002: Description and evaluation of a hydrometeorological forecast system for mountainous watersheds. *Wea. Forecasting,* **17,** 250–262.

Wigley, T. M. L., and T. Qipu, 1983: Crop-climate modeling using spatial patterns of yield and climate. Part 1: Background and an example from Australia. *J. Climate Appl. Meteor.,* **22,** 1831–1841.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* International Geophysics Series, Vol. 91, Academic Press, 627 pp.

Wolff, G. T., M. L. Morrissey, and N. A. Kelly, 1984: An investigation of the sources of summertime haze in the Blue Ridge Mountains using multivariate statistical methods. *J. Climate Appl. Meteor.,* **23,** 1333–1341.

World Meteorological Organization, 1996: *Guide to Meteorological Instruments and Methods of Observation.* 6th ed. WMO-No. 8, World Meteorological Organization, unpaged.

Xu, Q., and B. Zhou, 2003: Retrieving soil water contents from soil temperature measurements by using linear regression. *Adv. Atmos. Sci.,* **20,** 849–858.

Xue, Y., M. J. Fennessy, and P. J. Sellers, 1996: Impact of vegetation properties on U.S. summer weather prediction. *J. Geophys. Res.,* **101,** 7419–7430.

Yamane, T., 1967: *Statistics, An Introductory Analysis.* 2d ed. Harper and Row, 919 pp.

Yan, H., and R. A. Anthes, 1988: The effect of variations in surface moisture on mesoscale circulations. *Mon. Wea. Rev.,* **116,** 192–208.

Yeh, T.-C., R. T. Wetherald, and S. Manabe, 1984: The effect of soil moisture on the short-term climate and hydrology change—A numerical experiment. *Mon. Wea. Rev.,* **112,** 474–490.

Yin, Z., and T. H. L. Williams, 1997: Obtaining spatial and temporal vegetation data from Landsat MSS and AVHRR/NOAA satellite images for a hydrologic model. *Photogr. Eng. Remote Sens.,* **63,** 69–77.

Zamora, R. J., S. Solomon, E. G. Dutton, J. W. Bao, M. Trainer, R. W. Portmann, A. B. White, D. W. Nelson, and R. T. McNider, 2003: Comparing MM5 radiative fluxes with observations gathered during the 1995 and 1999 Nashville southern oxidants studies. *J. Geophys. Res.,* **108,** 4050, doi: 10.1029/2002JD002122.

——, E. G. Dutton, M. Trainer, S. A. McKeen, J. M. Wilczak, and Y.-T. Hou, 2005: The accuracy of solar irradiance calculations used in mesoscale numerical weather prediction. *Mon. Wea. Rev.,* **133,** 783–792.

Zehnder, J. A., 2002: Simple modifications to improve fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model performance for the Phoenix, Arizona, metropolitan area. *J. Appl. Meteor.,* **41,** 971–979.

Zeng, X., R. E. Dickinson, A. Walker, M. Shaikh, R. S. DeFries, and J. Qi, 2000: Derivation and evaluation of global 1-km fractional vegetation cover data for land modeling. *J. Appl. Meteor.,* **39,** 826–839.

Zhang, D.-L., and W.-Z. Zheng, 2004: Diurnal cycles of surface winds and temperatures as simulated by five boundary layer parameterizations. *J. Appl. Meteor.,* **43,** 157–169.

Zhong, S., H.-J. In, X. Bian, J. Charney, W. Heilman, and B. Potter, 2005: Evaluation of real-time high-resolution MM5 predictions over the Great Lakes region. *Wea. Forecasting,* **20,** 63–81.

Ziegler, C. L., W. J. Martin, R. A. Pielke, and R. L. Walko, 1995: A modeling study of the dryline. *J. Atmos. Sci.,* **52,** 263–285.

Zobler, L., 1986: A world soil file for global climate modeling. NASA Tech. Memo. 87802, 32 pp.