2010

An Empirical Latent Heat Flux Parameterization for the Noah Land Surface Model

Christopher M. Godfrey
University of Oklahoma, cgodfrey@unca.edu

David J. Stensrud
NOAA/National Severe Storms Laboratory

An Empirical Latent Heat Flux Parameterization for the Noah Land Surface Model

CHRISTOPHER M. GODFREY*

School of Meteorology and Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

DAVID J. STENSRUD

NOAA/National Severe Storms Laboratory, Norman, Oklahoma

(Manuscript received 7 January 2009, in final form 9 March 2010)

ABSTRACT

Proper partitioning of the surface energy fluxes that drive the evolution of the planetary boundary layer in numerical weather prediction models requires an accurate representation of initial land surface conditions. Unfortunately, soil temperature and moisture observations are unavailable in most areas and routine daily estimates of vegetation coverage and biomass are not easily available. This gap in observational capabilities seriously hampers the evaluation and improvement of land surface parameterizations, since model errors likely relate to improper initial conditions as much as to inaccuracies in the parameterizations. Two unique datasets help to overcome these difficulties. First, 1-km fractional vegetation coverage and leaf area index values can be derived from biweekly maximum normalized difference vegetation index composites obtained from daily observations by the Advanced Very High Resolution Radiometer onboard NOAA satellites. Second, the Oklahoma Mesonet supplies multiple soil temperature and moisture measurements at various soil depths each hour. Combined, these two datasets provide significantly improved initial conditions for a land surface model and allow an evaluation of the accuracy of the land surface model with much greater confidence than previously. Forecasts that both include and neglect these unique land surface observations are used to evaluate the value of these two data sources to land surface initializations. The dense network of surface observations afforded by the Oklahoma Mesonet, including surface flux data derived from special sensors, provides verification of the model results, which indicate that predicted latent heat fluxes still differ from observations by as much as 150 W m^{-2}. This result provides a springboard for assessing parameterization errors within the model. A new empirical parameterization developed using principal-component regression reveals simple relationships between latent heat flux and other surface observations. Periods of very dry conditions observed across Oklahoma are used advantageously to derive a parameterization for evaporation from bare soil. Combining this parameterization with an empirical canopy transpiration scheme 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 observations.

1. Introduction

Proper partitioning of the surface energy fluxes that drive the evolution of the planetary boundary layer (PBL) requires an accurate representation of land surface conditions in numerical weather prediction (NWP) models. 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 daily remote sensing observations of σ_f and LAI are not easily available at high spatial resolution. This gap in our

* Current affiliation: Department of Atmospheric Sciences, University of North Carolina at Asheville, Asheville, North Carolina.

Corresponding author address: Dr. Christopher M. Godfrey, Dept. of Atmospheric Sciences, UNC Asheville, One University Heights, Asheville, NC 28804.
E-mail: cgodfrey@unca.edu

DOI: 10.1175/2010JAMC2180.1

observational capabilities seriously hampers the evaluation and improvement of land surface model (LSM) parameterizations, since improper initial conditions and inaccuracies in the model formulations very likely produce comparable model errors. Taking advantage of a unique set of soil and vegetation observations to specify an improved characterization of the initial land surface conditions, it becomes possible to understand the shortcomings in the National Centers for Environmental Prediction (NCEP)–Oregon State University–Air Force–Hydrologic Research Laboratory (Noah) LSM (Chen et al. 1996; Koren et al. 1999) by evaluating individual model components.

Land surface parameterizations direct the exchange of energy between the land surface and the atmosphere. Many different land surface parameterizations are available (e.g., Bhumralkar 1975; Blackadar 1976; Deardorff 1978; McCumber and Pielke 1981; Pan and Mahrt 1987; Noilhan and Planton 1989; Chen and Dudhia 2001), but all serve to 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 moisture conditions, as well as vegetation coverage, atmospheric conditions, and the physical properties of the soil. Soil moisture is a particularly 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. 2004). Root-zone soil moisture impacts PBL 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). Soil temperature modulates the distribution of heat near the soil surface (Dudhia 1996) and affects the surface radiation budget through its influence on the ground heat flux (Brotzge and Crawford 2003). Unfortunately, NCEP operational Eta Model analyses (Black 1994), which provide initial conditions to a variety of operational and research models, exhibit strong biases in soil temperature and soil moisture (Marshall et al. 2003; Godfrey and Stensrud 2008). This finding necessitates the inclusion of soil temperature and soil moisture observations in model initializations in order to address inaccuracies in LSM parameterizations.

Studies of the effect of vegetation density and coverage indicate the necessity for including vegetation parameterizations within NWP models (e.g., Pielke et al. 1991; McPherson et al. 2004). The characterization of vegetation in the Noah LSM requires two variables. 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 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. Together, vegetation density and coverage provide critical information on the partitioning of total evaporation between bare soil and canopy transpiration (Chen and Dudhia 2001). 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 and air temperature gradients in response to vegetation density gradients (e.g., 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 (Anthes 1984; Segele et al. 2005).

Observational studies report measurements of the effect of vegetation on the PBL. Fiebrich and Crawford (2001) trace anomalously cool air temperatures at a single Oklahoma Mesonet (hereinafter 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 (McPherson et al. 2004; Haugland and Crawford 2005). 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).

These studies highlight the need for accurate vegetation information, as well as soil temperature and moisture conditions, to properly initialize land surface models. The Noah LSM employed for this study implements a monthly climatology for σ_f and a constant LAI. Such coarse-resolution data based solely on climatology are insufficient to capture the important variations in surface characteristics (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 variations in rainfall, forest fires, irrigation, deforestation, desertification, crop harvesting, land usage, hail or tornado damage, and temporal variations in the growth and senescence of green vegetation. However, modeling studies implementing near real-time land surface characteristics from satellite observations show great promise for improving forecasts (e.g., Oleson and Bonan 2000; Zeng et al. 2000; Crawford et al. 2001; Kurkowski et al. 2003).

This study represents an effort to improve the specification of initial conditions and ultimately to facilitate improved model forecasts of air temperature and moisture, which directly affect PBL processes and convective development. A modified version of The Pennsylvania State University–National Center for Atmospheric Research fifth-generation Mesoscale Model (MM5), version 3.6 (Dudhia 1993; Grell et al. 1995; Dudhia 2003), coupled with the Noah LSM, assimilates soil temperature, soil moisture, σ_f, and LAI observations for several case studies. Model forecasts that both include and neglect these improved initial surface conditions highlight errors present in the physical parameterizations within the Noah LSM. Such errors specifically suggest the need for an improved latent heat flux parameterization. The combination of the complexity of the physical processes leading to evapotranspiration and the assumptions inherent in the current formulation make the latent heat flux a prime candidate for refinement.

There exist several methods for evaluating, calibrating, and improving land surface models. Artificial neural networks have proven useful in assessing the performance of land surface models and in correcting errors in latent heat flux and have shown that problems in model processes may lead to errors in surface energy fluxes (Abramowitz 2005; Abramowitz et al. 2006, 2007). In a different approach, multicriteria calibration methods can improve estimates of parameters within land surface models using observations of certain variables and lead to improved flux calculations (e.g., Bastidas et al. 1999; Gupta et al. 1999; Xia et al. 2002). This study instead develops a new parameterization scheme for latent heat flux using a principal-component regression technique. In a novel approach to determining latent heat flux, the new parameterization derives from surface observations rather than from theoretical formulations.

2. Observations

The Oklahoma Mesonet (Brock et al. 1995) is a network of over 117 automated surface observing stations (Fig. 1) that provides observations every 5 min. All Mesonet sites report the standard suite of surface observations plus soil temperature under native vegetation

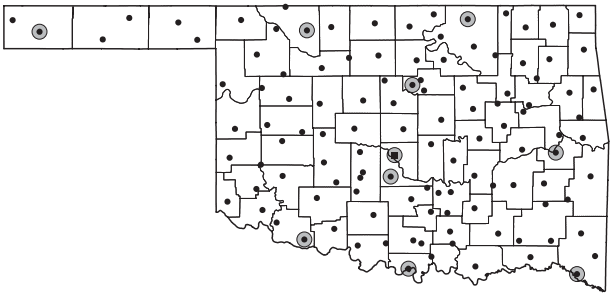

FIG. 1. Site locations for each of the 117 Mesonet sites operating between May 2004 and June 2006. Gray circles indicate OASIS super sites and the circle with the square identifies the Norman site.

at a depth of 10 cm, with approximately half of the sites also providing measurements at 5- and 30-cm depths. Over 100 sites record soil moisture every 30 min at depths of 5, 25, 60, and 75 cm. Mesonet sites are located in several different climate divisions across Oklahoma and in regions with very different vegetation types. All Oklahoma Mesonet data fall subject to rigorous quality assurance procedures in order to produce reliable research-quality data (Shafer et al. 2000), though all such data remain subject to inevitable uncertainties (e.g., Wilson et al. 2002).

Measurements of surface energy fluxes rely on instrumentation deployed at selected Mesonet sites by the Oklahoma Atmospheric Surface-layer Instrumentation System (OASIS) project (Brotzge et al. 1999; Brotzge 2000). A special suite of OASIS instruments augments the standard Mesonet instrumentation at 10 super sites (Fig. 1), measuring sensible heat flux, ground heat flux, and the four components of net radiation. 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 heat flux plates installed at a depth of 5 cm. Estimates of the storage ground heat flux derive from measurements of the soil moisture at 5 cm, an average volume fraction of minerals and organic matter, and the temperature within the 0–5-cm soil layer (Brotzge and Crawford 2003).

Brotzge and Crawford (2003) blame surface energy budget closure problems on systematic underestimates of latent heat flux using the eddy covariance method. The magnitude of these errors varies by solar time and season. Similarly, Wilson et al. (2002) find a closure imbalance on the order of 20% at FLUXNET sites, primarily due to underestimates of both sensible and latent heat flux using the eddy covariance technique. Rather than directly estimating latent heat fluxes from measurements, the residual of the surface energy balance instead provides a proxy for latent heat flux estimates. This residual

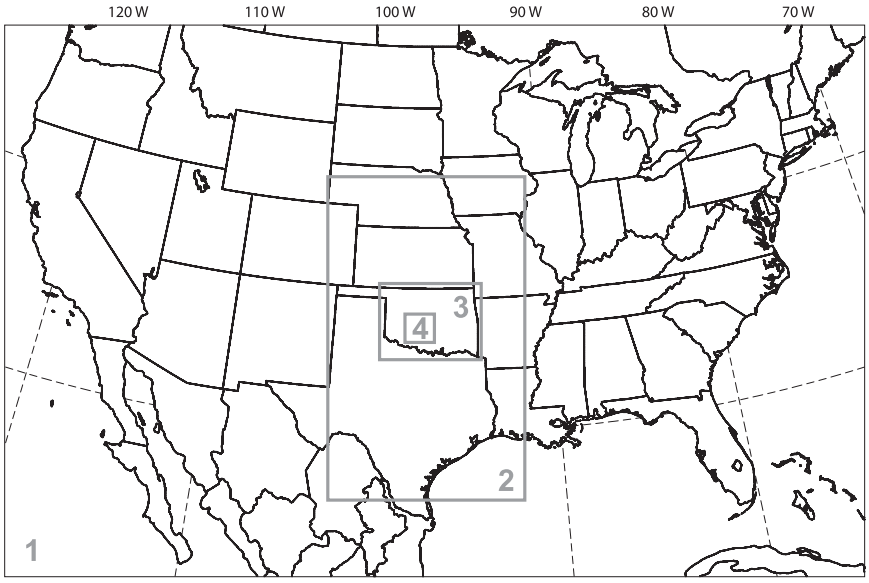

FIG. 2. Location of the four nested MM5 domains with 27-, 9-, 3-, and 1-km grid resolution.

approach may provide a more reasonable estimate of the latent heat flux than direct estimates from an eddy covariance system (Brotzge 2004).

Vegetation fraction and LAI are calculated following Chang and Wetzel (1991) and Yin and Williams (1997), respectively, using biweekly maximum value composites of normalized difference vegetation index (NDVI). The NDVI values are compiled from daily National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) data with a ground resolution of 1.09 km.

3. Model forecasts

a. Model description

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-h forecasts on four nested model domains with 27-, 9-, 3-, and 1-km grid resolution (Fig. 2) and 23 vertical half-sigma levels. NCEP operational Eta Model analyses and forecasts provide initial and boundary conditions. Specific user-defined options for MM5 include the Kain and Fritsch (1993) cumulus parameterization on domains one and two only, no shallow convection, the Hong and Pan (1996) PBL parameterization, 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 and is called every 5 min. This scheme systematically overestimates surface downward shortwave radiation, with overestimates exceeding 50 W m^{-2} under cloudless skies. In lieu of adding explicit formulations for ozone and aerosol absorption and Rayleigh and upward aerosol scattering (Zamora et al. 2003), the shortwave radiation is tuned for each case study to match the solar radiation observations from all nine OASIS super sites. Other options and parameters remain set to their default values.

The Noah LSM provides the multilayer soil physics and vegetation package. This LSM contains four soil layers depicting soil temperature and soil moisture and accounts for vegetation categories, monthly σ_f, and soil texture, and includes parameterizations for evaporation, soil drainage, runoff, the root zone, and canopy moisture (Skamarock et al. 2005). The Noah LSM functions as the primary driver for land surface processes in MM5 and contains nearly identical code to the land surface schemes found in both the operational Eta Model and the Weather Research and Forecasting (WRF) model. This allows for direct compatibility between the time-dependent soil variables and surface fluxes in MM5 forecasts and the Eta Model analyses that initialize the forecasts.

b. Surface fluxes in the Noah LSM

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). Since the predominant vegetation cover is grass at Oklahoma Mesonet sites, it is assumed that the canopy water

content is zero in the experiments that follow, 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 $E_{\rm dir}$ and E_t.

1) 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 Betts et al. (1997) in which

$$E_{\rm dir} = (1 - \sigma_f)\beta^2 E_p, \tag{1}$$

where

$$\beta^2 = \left(\frac{\Theta_1 - \Theta_w}{\Theta_{\rm ref} - \Theta_w}\right)^2 \tag{2}$$

represents a normalized soil moisture availability term, Θ_w is the wilting point, $\Theta_{\rm ref}$ is the field capacity, and Θ_1 is the volumetric water content of the top soil layer (Chen and Dudhia 2001). The β term is squared as suggested by Ek et al. (2003). 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 C_h (Mahrt and Ek 1984; Ek and Mahrt 1991).

2) Canopy transpiration

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

$$E_t = \sigma_f E_p P_c \left[1 - \left(\frac{W_c}{S}\right)^{0.5}\right], \tag{3}$$

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

$$P_c = \frac{r + \Delta}{r(1 + C_h R_c) + \Delta}, \tag{4}$$

where r and Δ are functions of the thermodynamic properties of the air at the lowest model level (Ek and Mahrt 1991), and

$$R_c = \frac{R_{\rm cmin}}{({\rm LAI}) F_1 F_2 F_3 F_4} \tag{5}$$

is the canopy resistance following the formulation of Jacquemin and Noilhan (1990) where $R_{\rm cmin}$ is the minimum stomatal resistance for each vegetation type. The canopy resistance factors F_1, F_2, F_3, and F_4 represent the effects of solar radiation, vapor pressure deficit, air temperature, and soil moisture (Chen and Dudhia 2001).

Canopy resistance is the most important factor contributing to canopy transpiration (e.g., Holtslag and Ek 1996; Ronda et al. 2001). Despite this physical importance, the canopy resistance formulation in Eq. (5) is arguably the most questionable term in the Noah LSM, since it simply multiplies together four physically important atmospheric and land surface effects. Jarvis (1976) proposed a very similar formulation for stomatal conductance (the inverse of resistance) based on the known independent influence of four variables. 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 is adopted and implemented in several land surface models with some modification (e.g., Noilhan and Planton 1989; Jacquemin and Noilhan 1990; Chen and Dudhia 2001), leads to the four canopy resistance factors in Eq. (5).

c. Selection of case studies

To aid in a selection of several case studies, MM5 is used to compute daily 48-h forecasts during the entire warm season of 2004. Results from 9-, 24-, and 33-h forecasts of 2-m air temperature and mixing ratio for each run are compared with corresponding Mesonet observations. The four selected cases provide a representative sample of the accuracy and typical errors seen in these warm season forecasts. These four case studies start at 1200 UTC 3 May, 20 July, 1 August, and 3 September 2004. In all cases, no strong synoptic features passed over Oklahoma. This assortment of forecasts under synoptically quiescent conditions provides several ideal cases for studying the impact of improved initial conditions on surface flux forecasts and maximizes the potential for isolating the effect of changes to the LSM on near-surface atmospheric variables.

d. Initial conditions

To explore the importance of the land surface on the model forecasts, two different sets of initial conditions for the soil and land surface are used. The control MM5 (CTRL) uses a 0.15° × 0.15° climatological σ_f produced

from a five-year climatology of NDVI observations (Gutman and Ignatov 1998). The model also assumes a constant LAI (set to 4.0) regardless of the season or location. Eta Model analyses provide initial soil temperature and soil moisture conditions.

The second MM5 (MM5VEGSOIL) initial condition includes the 1-km resolution σ_f and LAI observations derived from a 15- or 16-day NDVI composite. The composite windows end just prior to the start date for each case, thereby representing analyses that could be obtained in real time. Satellite-derived σ_f and LAI data cover a swath similar to the area of domain two (Fig. 2). 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.

The MM5VEGSOIL initial condition also includes Mesonet soil data interpolated to the model grid using a two-pass Barnes analysis (Barnes 1973). Soil temperature and moisture observations replace Eta Model analyses of soil fields in the Noah LSM on domains three and four in the top two layers for soil temperature and the top three layers for soil moisture. Observations of soil temperature at a depth of 5 cm replace the soil temperature in the 0–10-cm model layer. 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 depths of 5, 25, and 60 cm replaces the initial soil moisture field in the 0–10-, 10–40-, and 40–100-cm model layers, respectively. 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. All soil fields for domains one and two also remain unchanged from the Eta analyses. The MM5VEGSOIL initial condition represents a substantially more accurate specification of the land surface and soil conditions for the model.

e. Preliminary results

For the cases under consideration, the observed soil moisture is considerably wetter and the observed soil temperatures are generally cooler at the 1200 UTC start time than the Eta analyses (Godfrey and Stensrud 2008). The climatological σ_f and constant LAI values stand in stark contrast to the σ_f and LAI observations. The MM5 forecasts indicate that the greatly improved initial conditions in the MM5VEGSOIL forecast do not yield dramatic improvements over the CTRL forecasts (Fig. 3), although improvements of 20–40 W m^{-2} are seen in three of the four cases. However, the MM5VEGSOIL forecasts still consistently underestimate midday latent heat fluxes by 20%–40% compared with observations. This occurs even though the model is given a significantly improved characterization of the initial land surface conditions. 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. Further testing (not shown) indicates that including only soil observations in the initial conditions improves the latent heat flux forecasts, while including only vegetation observations considerably worsens the flux forecasts. The difference between these midday latent heat flux forecasts can exceed 225 W m^{-2}. Thus, the use of vegetation observations appears to offset the improvements gained from using soil observations, indicating that errors are present within the flux formulations. It appears that these errors primarily result from the incorrect partitioning between the fluxes of sensible and latent heat. Such errors highlight the need for an improved latent heat flux parameterization.

4. Empirical latent heat flux parameterization

Given the physical importance of canopy resistance in the canopy transpiration term of the Noah LSM, one approach to improving short-term latent heat flux forecasts is to focus on tuning the formulation for canopy resistance. An inverted form of the Noah LSM is constructed that uses Mesonet observations as input and calculates the values of plant coefficient P_c (and thereby canopy resistance R_c) needed to yield the observed latent heat fluxes from the OASIS super sites. Data over a five-month period during 2004 are used as input and results evaluated. Unfortunately, many of the resistance values are unphysical, including exceedingly large canopy resistances and unbounded plant coefficients. This occurs because either the $E_{\rm dir}$ term [Eq. (1)] is greater than the observed latent heat flux or the sum of $E_{\rm dir}$ and $\sigma_f E_p$ [Eq. (3)] is less than the observed latent heat flux. These problems persist even after adjusting for a $\pm$20 W m^{-2} error in the OASIS latent heat flux observations (Brotzge 2000). Thus, the $E_{\rm dir}$ and E_t terms clearly yield inappropriate values when forced with observations. Any 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. Tests indicate that least squares simple and multiple linear

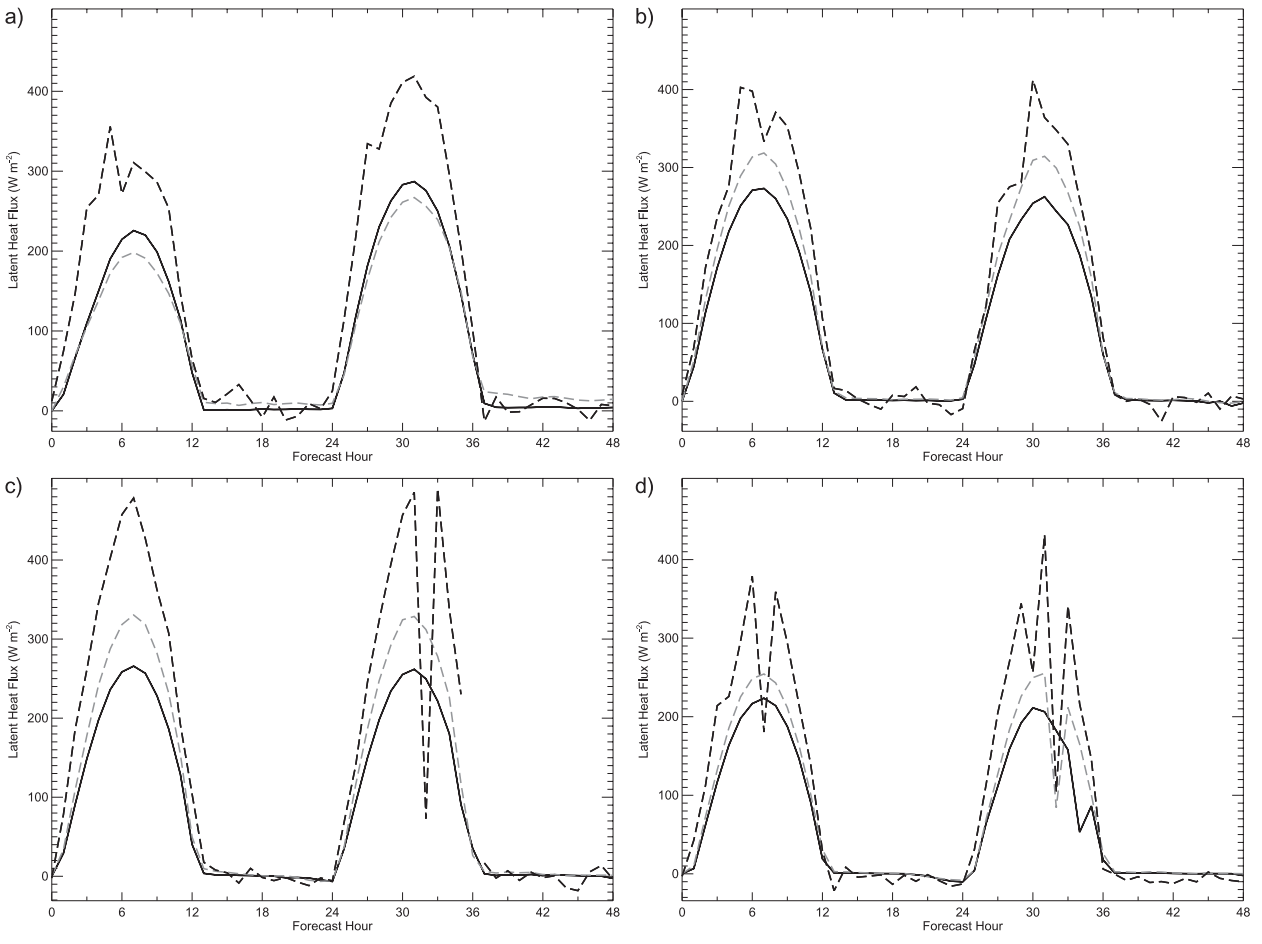

FIG. 3. Latent heat flux (W m^{-2}) at Norman, OK, for CTRL (black) and MM5VEGSOIL (gray dashed) forecasts for domain 4 initialized at 1200 UTC (a) 3 May, (b) 20 Jul, (c) 1 Aug, and (d) 3 Sep 2004 compared with the residual of the surface energy balance computed from Oklahoma Mesonet observations (black dashed).

regression models with automatic and manual predictor selection have limited potential to provide good results. Instead, a principal-component regression procedure is used for predictor selection.

a. Principal-component regression

Principal-component regression techniques (e.g., Richman 1986; Wilks 2006) 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), wheat yield from temperature and rainfall observations (Wigley and Qipu 1983), and surface ozone concentrations (Pryor et al. 1995) have all used this technique. In addition, principal-component regression has been used to determine source regions for fine particulates and sulfate (Wolff et al. 1984). However, the use of this technique to predict fluxes of latent heat from a wealth of surface observations represents a novel application.

Since the Noah LSM contains separate expressions for latent heat flux over bare soil and vegetated surfaces, separate principal-component regression analyses are conducted to develop the best possible expressions for both E_{dir} and E_t that match the observed latent heat fluxes. 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). 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 is used to correct for this problem, such that

$$\overline{R}^2 = 1 - (1 - R^2)\left(\frac{n-1}{n-p-1}\right), \quad (6)$$

where p is the number of predictors in the multiple regression model and n is the sample size (Yamane 1967). The $\overline{R}^2$ value justifies the results of each principal-component regression in each independent cross-validation dataset.

b. 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 combinations of variables that already exist within the Noah LSM. To remove the influence of very small nighttime latent heat fluxes in creating a new scheme for latent heat flux forecasts, the principal-component regression is provided only with sets of observations that include incoming shortwave radiation values in excess of 10 W m^{-2}. There are also several restrictions on the available observations from the Oklahoma Mesonet. Data available for analysis span the period May 2004–June 2006 with satellite-derived vegetation data spanning only 15 April–15 September 2004. The observations needed to determine the surface energy balance are available every 30 min from nine OASIS sites. Since precipitation is known to interfere with sensible heat flux measurements, periods of rainfall are removed beginning with the first nonzero daily precipitation total through local midnight on the day of the observation. This also allows elimination of the wet canopy evaporation term during the development of a new parameterization, such that the total latent heat flux is simply the sum of the direct evaporation and canopy transpiration terms.

c. Direct evaporation from bare soil

Since vegetated surfaces surround every observation site, direct measurements of evaporation from bare soil are unavailable. However, the long time series of available soil moisture observations 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 types is roughly where the matric potential $\psi = -1500$ kPa (Marshall et al. 1996). 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. By separating only those sets of observations where the soil has reached the permanent wilting point at the 5-cm level, as calculated from the matric potential formulation in Basara and Crawford (2000), the residual of the surface energy balance becomes a good approximation to the direct evaporation from bare soil. This collection of bare soil evaporation observations comprises more than 6300 sets of observations and is used to determine a new $E_{\rm dir}$ parameterization.

From an initial wide selection of possible variables, multiple passes through a principal component analysis lead to a reduced pool of possible predictors for $E_{\rm dir}$. In addition to the overarching goal of achieving the largest possible $\overline{R}^2$ in the cross-validation data, several other factors contribute to the decision to retain or eliminate variables from the regression. These factors include the ease of implementation of the resulting flux equation in the Noah LSM, 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. The existence of 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 resulting equation for direct evaporation from bare soil is

$$E_{\rm dir} = \left\{ 22.33 + 0.0226[R_g(1-\alpha)]^{(3/2)}\left(\frac{\Theta_1 - \Theta_w}{\Theta_{\rm ref} - \Theta_w}\right) - 3.426V + 3650w \right\}(1 - \sigma_f), \quad (7)$$

where R_g is the incoming solar radiation (W m^{-2}), 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 $\Theta_{\rm ref}$ is the field capacity, V is the 10-m wind speed (m s^{-1}), and w is the 2-m mixing ratio (kg kg^{-1}). As implemented in the Noah LSM, Θ_1 is the volumetric water content of the top (0–10 cm) soil layer, Θ_w and $\Theta_{\rm ref}$ 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 $\overline{R}^2$ 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 R^2 between the same predictand and the $E_{\rm dir}$ from the original Noah LSM formulation is 0.52. Compared with the existing $E_{\rm dir}$ parameterization in the Noah LSM, the forecasts from the new empirical scheme more closely match the total latent heat flux observations when the soil is dry enough to assume senescent vegetation, particularly for increased $E_{\rm dir}$ (Fig. 4).

As indicated by locally weighted regressions prior to the principal-component regression, each of the variables

in Eq. (7) exhibits a quasilinear relationship with the observed latent heat flux during dry conditions. The second and most important term on the right-hand side of Eq. (7) 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. (2).

The negative coefficient in the wind speed term in Eq. (7) may seem somewhat counterintuitive. However, studies have shown that the eddy correlation method underestimates the sensible heat flux (e.g., Barr et al. 1994; Lee and Black 1994) and this underestimation increases with decreasing friction velocity (cf. Barr et al. 2006). As the wind speed increases, the sensible heat flux increases and the residual of the surface energy balance decreases to more reasonable values. The small wind speed term in Eq. (7) likely accounts for this behavior. Stomatal stress in the vegetation that is at or near its permanent wilting point may provide a second possible explanation for the negative coefficient. Studies show that transpiration generally increases with increasing wind speed up to a point where it may then decrease slightly with wind speed (Chang 1968; Dixon and Grace 1984), particularly if the vegetation is already dry (e.g., Lydolph 1964). Regardless, the overall contribution of this term to the evaporation from bare soil is quite small. The small mixing ratio term in Eq. (7) likely appears because of the dependence of leaf-to-air vapor pressure deficit on mixing ratio. At large vapor pressure deficits, and by extension low mixing ratios, stomatal closure reduces transpiration (El-Sharkawy et al. 1985).

With the exception of the σ_f term, each term in Eq. (7) represents a single variable present in the principal component analysis. Since each component uniquely contains a very strong signal from one of these 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 variables justifies using the principal-component regression approach both to ascertain the significance of the mutual correlations and as a robust variable-selection method.

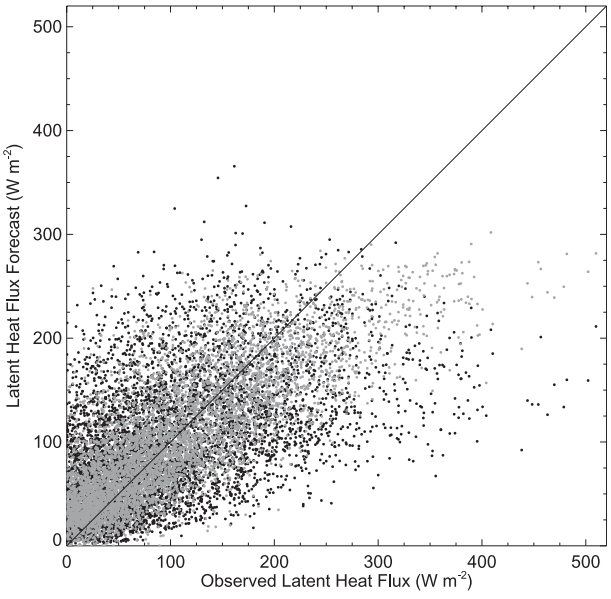

FIG. 4. Direct soil evaporation from the original Noah LSM formulation (black) and the empirical scheme (gray) compared with the observed total latent heat flux under dry soil conditions.

d. Canopy transpiration

With a proper parameterization for E_{dir} in place, a similar principal-component regression procedure based on 9239 sets of observations leads to a new empirical canopy transpiration scheme. The canopy transpiration term defined by

$$E_t = \frac{E_{\text{obs}} - E_{\text{dir}}}{\sigma_f} \tag{8}$$

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. (7) that already includes the σ_f 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 E_{dir} parameterization, a principal component analysis combined with physical, statistical, and practical considerations leads to the final regression equation for canopy transpiration,

$$E_t = \left\langle \begin{array}{l} -1392 + 0.9154\left\{ R_g(1-\alpha)\left[\left(\dfrac{\Theta_3 - \Theta_w}{\Theta_{\text{ref}} - \Theta_w}\right)^{1/2}\right]\right\} \\ +4.374T_{\text{air}} + 60.59\left[\dfrac{w}{w_s(T_{\text{air}})}\right] \end{array} \right\rangle \sigma_f + 6.116(\text{LAI}), \tag{9}$$

where Θ_3 is the volumetric water content ($m^3\ m^{-3}$) at 60-cm depth, T_{air} is the 9-m air temperature (K), $w_s(T_{air})$ is the saturation mixing ratio at the 9-m air temperature ($kg\ kg^{-1}$), and the remaining terms are the same as those defined for Eq. (7). The Θ_w and Θ_{ref} terms correspond with the measured soil textures at a depth of 60 cm at each Oklahoma Mesonet site. As implemented in the model, Θ_3 is the volumetric water content of the third (40–100 cm) 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.

The first term in Eq. (9) describes how the root-zone soil moisture availability scales the evaporative power of the sun. This is 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 LAI terms in the regression equation are less significant and their coefficients may serve as tunable parameters for different locations. Note, however, that Eq. (9) includes the effects of solar radiation, LAI, σ_f, 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 $\overline{R}^2$ 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. (8) and neglect the scaling by σ_f. Using only the independent cross-validation data and summing the E_t forecasts from Eq. (9) with the E_{dir} forecasts from Eq. (7) 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 R^2 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 E_{dir} and E_t parameterizations vastly improve the latent heat flux forecasts by the Noah LSM when driven by observations (Fig. 5). The original parameterization tends to overestimate latent heat fluxes under a variety of conditions, while the new parameterization corrects for this problem without introducing a negative bias.

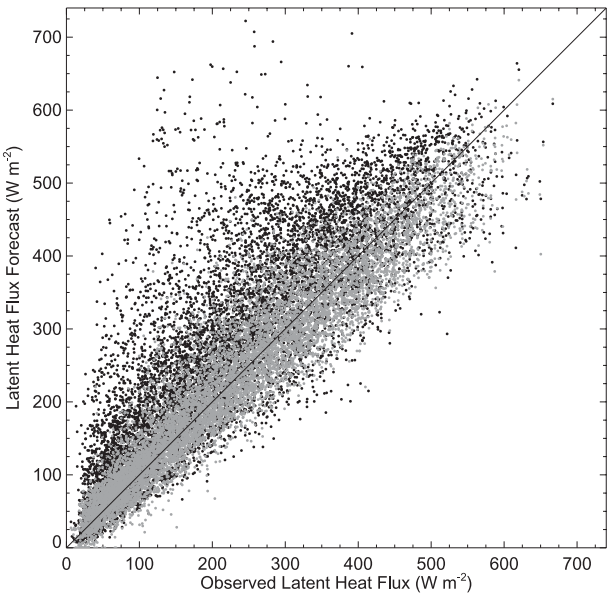

FIG. 5. Forecasts of total latent heat flux for 9239 forecast–observation pairs by the original Noah LSM formulation (black) and the new empirical direct soil evaporation and canopy transpiration schemes (gray) compared with the observed total latent heat flux for the period 15 Apr–15 Sep 2004.

5. Results

Latent heat flux forecasts from model predictions implementing the new empirical latent heat flux scheme during the daytime and initialized with both satellite-derived vegetation indexes and soil temperature and moisture observations (MM5LATENT) highlight the potential value of this empirical parameterization by showing vast improvement for all four cases over both the CTRL and MM5VEGSOIL forecasts when compared with observations at Norman, Oklahoma (Fig. 6). The maximum day one values of latent heat flux are increased by an average of 160 W m^{-2} compared to the CTRL forecast in much better agreement with the observations. Thus, the MM5LATENT forecast no longer severely underestimates daytime latent heat fluxes.

Initial testing (not shown) indicates that nighttime fluxes using the new flux scheme may exceed observations by nearly 50 W m^{-2}, especially shortly after sunset. This is attributed to limiting the observations in the principal-component regression to those associated with incoming solar radiation values in excess of 10 W m^{-2}. To overcome this limitation, the latent heat flux parameterization reverts to the original canopy resistance approach when modeled downward shortwave radiation falls below 10 W m^{-2}. Results from this combined new daytime and old nighttime flux scheme are shown in Fig. 6.

With reasonable latent heat flux forecasts, the previously overestimated sensible heat flux forecasts more

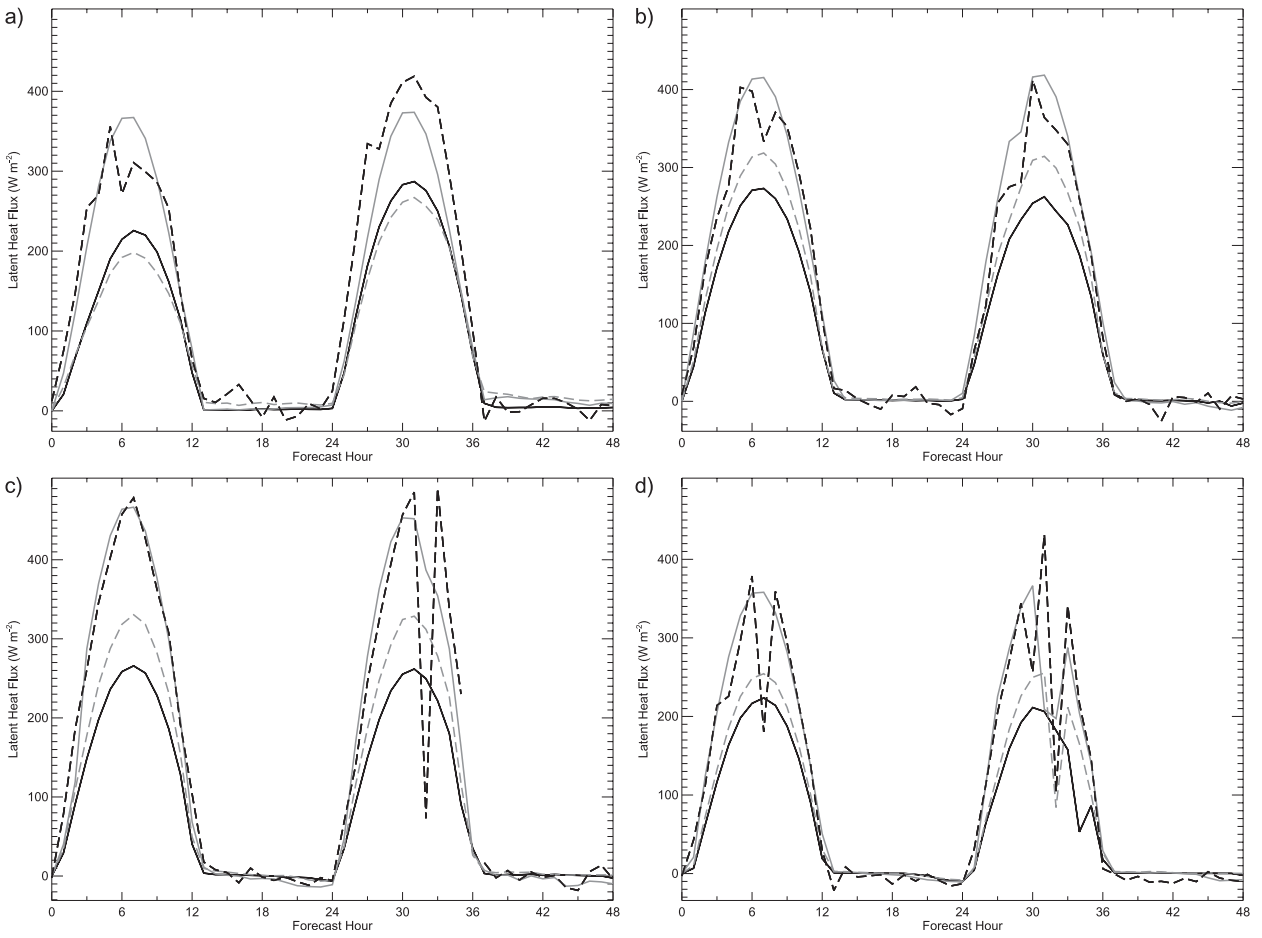

FIG. 6. Latent heat flux (W m^{-2}) at Norman, OK, for CTRL (black), MM5VEGSOIL (gray dashed), and MM5LATENT (gray) domain four forecasts initialized at 1200 UTC (a) 3 May, (b) 20 Jul, (c) 1 Aug, and (d) 3 Sep 2004 compared with the residual of the surface energy balance computed from Oklahoma Mesonet observations (black dashed).

closely resemble the observations (Fig. 7). At Norman, the maximum day one values of sensible heat flux are decreased by an average of 134 W m^{-2} compared to the CTRL forecasts. However, the MM5LATENT ground heat flux forecasts show little to no overall improvement over CTRL forecasts (not shown), suggesting that the main benefit to the new latent heat flux scheme is a more accurate partitioning of the sensible and latent heat fluxes.

The original 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. The original version of the Noah LSM implemented here does not force closure of the surface energy budget, though the most recent version more adequately addresses surface energy budget closure problems. To force closure of the surface energy budget in the modified Noah LSM for use in long-term climate modeling applications, one possible 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. Tests using both approaches in coupled MM5 forecasts 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 modified Noah LSM calculates each component of the surface energy balance individually and, like the original model formulation, does not force closure of the surface energy budget.

The new latent heat flux scheme also yields improved 2-m temperature forecasts (Fig. 8), which agree particularly well with the observations during the first 6 to 9 h after sunrise during the times of maximum warming. Cumulative errors in the surface energy balance likely cause the air temperature to start decreasing an hour or

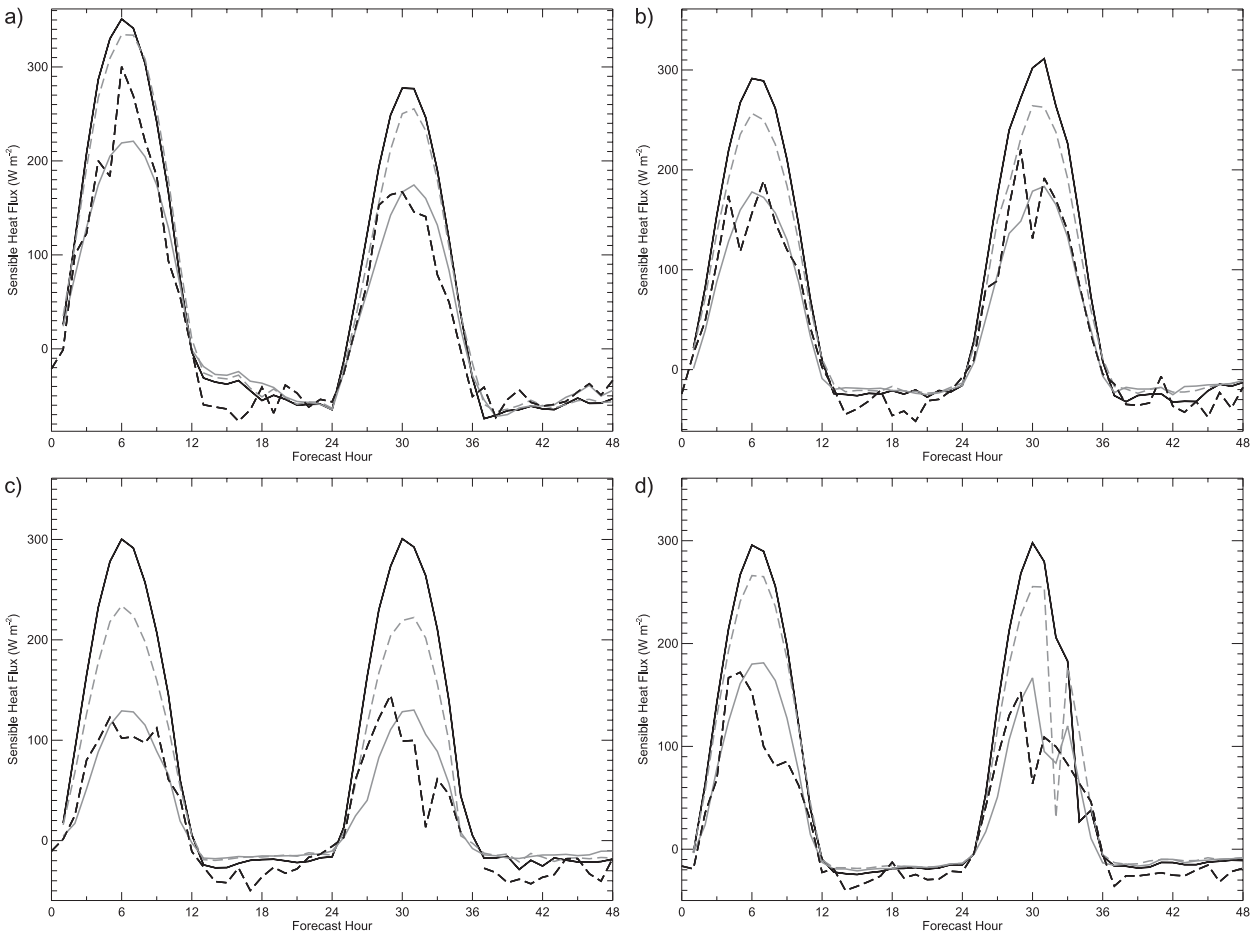

FIG. 7. As in Fig. 6, but for sensible heat flux.

two too early in the diurnal cycle as seen in all the forecasts. The sharp drop in 2-m air temperature near sunset is a consequence of the extrapolation errors during PBL regime transitions from free convection to stable conditions and not from surface energy flux errors.

The results for 2-m mixing ratio forecasts also indicate that the new latent heat flux parameterization yields improved predictions (Fig. 9). MM5 typically 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 ratios during PBL regime transitions, again caused by extrapolation errors, mixing ratio forecast errors decrease for the MM5LATENT forecasts compared with the other predictions.

Comparisons between the model and surface observations from all Mesonet stations across the main body of Oklahoma show similar results. However, observations from Oklahoma serve as the training data for the empirical latent heat flux parameterization in the Noah LSM. Sets of independent observations from two locations outside of Oklahoma help accomplish a simple verification of the new scheme by comparing modeled fluxes with data that do not compose any portion of the training data used to develop the empirical equations. Two meteorological-flux towers, maintained by the United States Department of Agriculture (USDA) Agricultural Research Service (ARS) National Soil Tilth Laboratory (NSTL), directly measure the four components of the surface energy balance near Ames, Iowa. Similar terrain to that in Oklahoma surrounds both sites: one tower stands over a soybean field and the other tower resides over a corn field. Direct measurements of sensible and latent heat flux occur roughly 2 m above the vegetation canopy at each location [see Kustas et al. (2005) for more information]. Data are available for the 20 July, 1 August, and 3 September 2004 case studies. The simultaneously measured fluxes over the corn and soybean fields may differ by more than 100 W m^{-2} on these three days, highlighting 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.

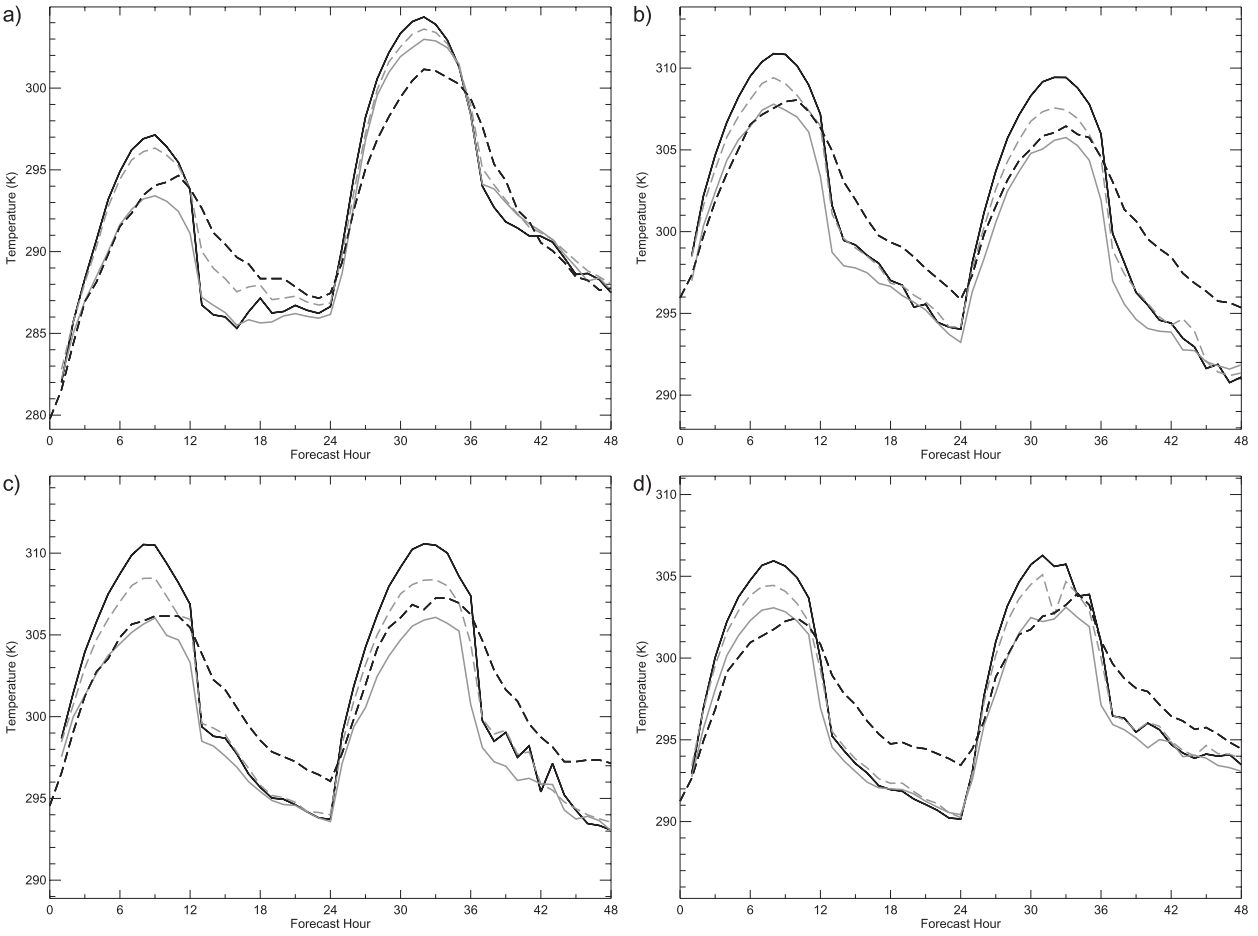

FIG. 8. As in Fig. 6, but for the 2-m air temperature (K).

Since soil temperature and moisture observations are only available from the Oklahoma Mesonet, the improved initial conditions in the MM5LATENT forecasts only include the satellite-derived vegetation indexes. As with the CTRL forecasts, the soil initial conditions in MM5LATENT over Iowa derive from the Eta analyses. 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 virtually identical. Despite lacking accurate initial soil temperature and moisture conditions, the 20 July and 3 August 2004 MM5LATENT forecasts perform remarkably well compared with the latent heat fluxes measured over both corn and soybeans. The MM5LATENT forecasts have smaller flux errors than seen in the CTRL forecast by as much as 115 W m^{-2} (Fig. 10). It is curious that the CTRL forecast overestimates rather than underestimates the observed latent heat flux (as seen over Oklahoma), yet the new empirical latent heat flux scheme realistically captures the total evapotranspiration. Admittedly, the MM5LATENT forecast underestimates by roughly 80 W m^{-2} the observed latent heat fluxes in the forecast initialized on 3 September 2004 with the CTRL providing a slightly better forecast. Without soil moisture observations, it is difficult to ascertain the reasons for the differences between the flux forecasts. However, on average the new latent heat flux formulation improves the flux forecasts over Ames more than it degrades them, providing further evidence that this new latent heat flux formulation is beneficial.

6. Discussion

While recent advances in NWP models have improved short-term forecasts, the Noah LSM still inaccurately predicts 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 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

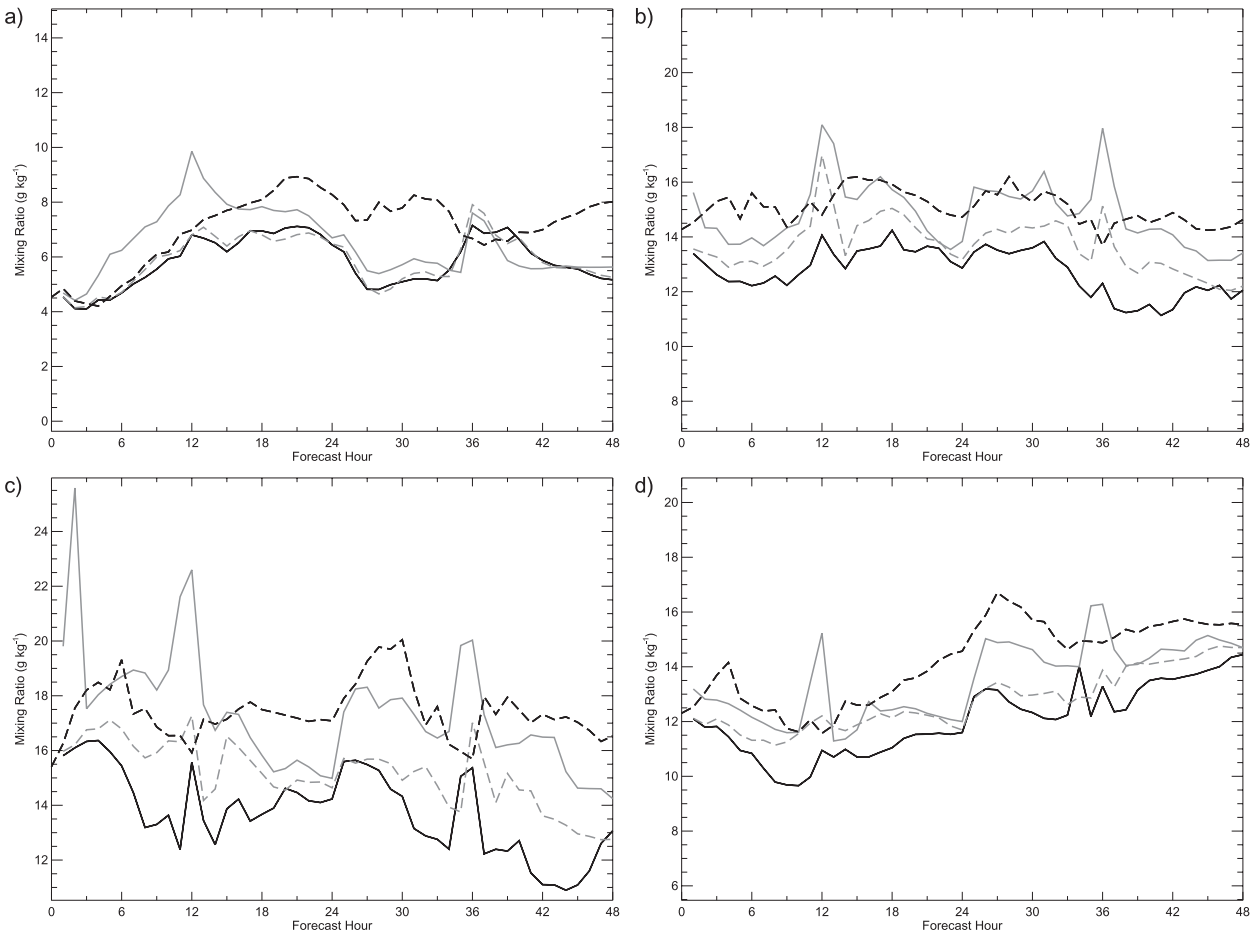

FIG. 9. As in Fig. 6, but for the 2-m mixing ratio (g kg^{-1}).

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 real-time satellite observations, provides a unique opportunity to begin the process of improving Noah LSM parameterizations by initializing the model with a considerably improved characterization of the land surface. Indeed, for the case studies discussed here, soil moisture and vegetation conditions strongly impact model forecasts.

Despite providing the Noah LSM with improved initial conditions, the model forecasts fail to capture surface energy fluxes realistically. 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 120 W m^{-2} in a 48-h forecast period. This leads to temperature errors in excess of 2°C and mixing ratio errors that exceed 3 g kg^{-1}. While these errors may seem large, they represent an improvement over the CTRL forecasts that do not use any Mesonet observations or satellite-derived vegetation information. These results emphasize the potential forecast value of minimizing errors in land surface initial conditions, while illustrating the profound difficulty in evaluating individual model components when all of the schemes are interdependent. Because the land surface physics determine the partitioning of the surface energy budget, improvements for forecasts with excellent soil and vegetation initial conditions require a careful calibration of many of these interdependent parameterization schemes within the Noah LSM.

In an attempt to improve the latent heat flux predictions, a new empirical parameterization is developed from a wealth of unique surface, soil, and vegetation observations that dramatically improves latent heat flux forecasts in the Noah LSM when compared with OASIS flux observations. Applying a completely new approach, this scheme replaces the usual theoretical formulations for latent heat flux. For one case study, the error in the maximum daily latent heat flux falls from close to 150 W m^{-2} for the MM5VEGSOIL forecast to approximately 12 W m^{-2} using the new empirical parameterization for latent heat

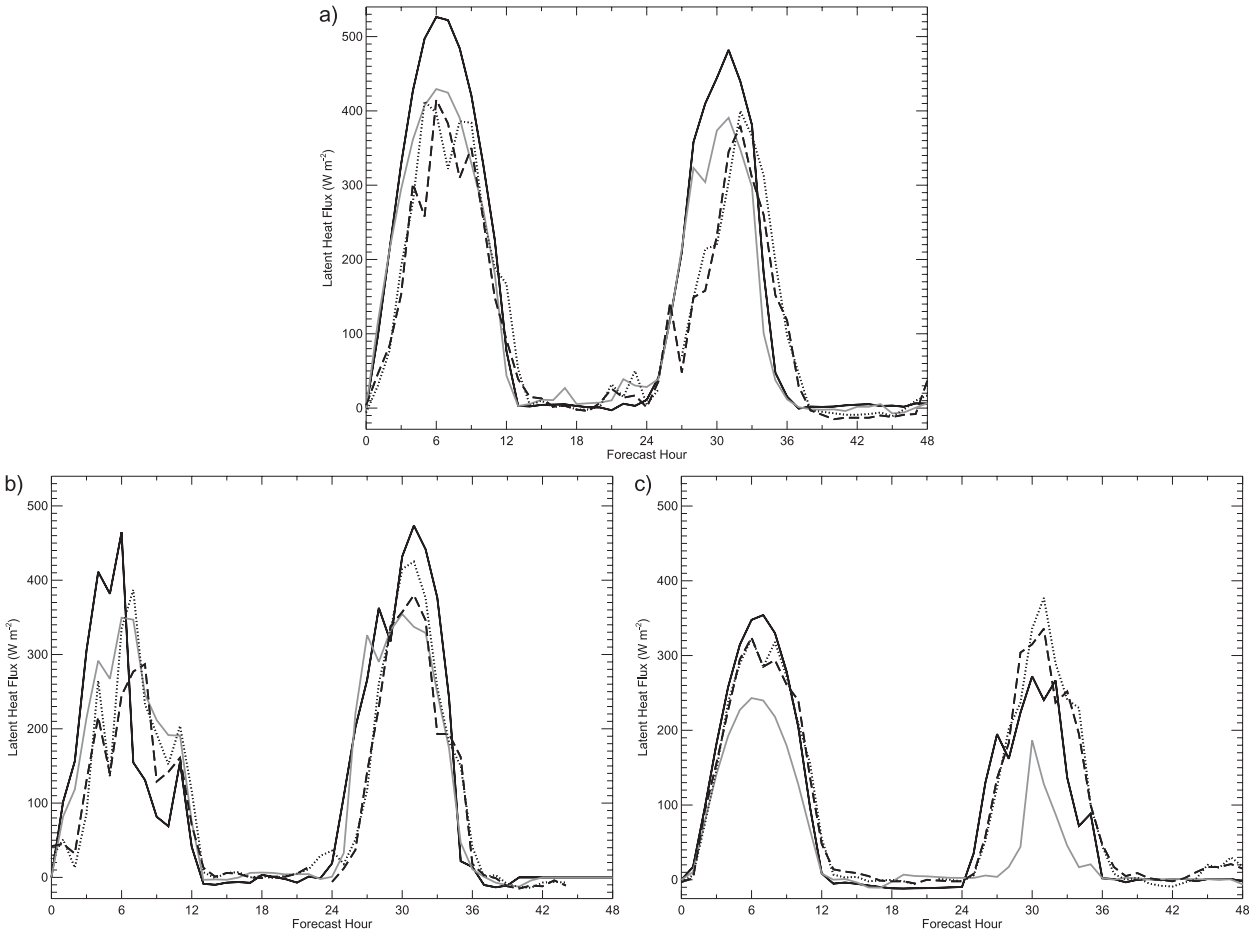

FIG. 10. Latent heat flux (W m^{-2}) near Ames, IA, for CTRL (black) and MM5LATENT (gray) forecasts initialized at 1200 UTC (a) 20 Jul, (b) 1 Aug, and (c) 3 Sep 2004 compared with observations of latent heat flux over a soybean field (dotted) and over a corn field (dashed).

flux starting with the same set of initial conditions. That this relatively simple approach outperforms the traditional theoretical formulations for several case studies highlights the weakness in the current physically based methods for calculating latent heat flux. Forecasts using the new latent heat flux formulation show improvements in daily maximum air temperature and mixing ratio forecasts of greater than 4°C and 2 g kg^{-1} over CTRL forecasts with the same initial conditions and model formulations present in the operational version of the Noah LSM. While physically based parameterizations are generally preferred for a variety of reasons, the formulation that yields the best forecasts should take priority. Improving upon this empirical formulation with a theoretical approach may prove quite difficult. However, other statistical methods or artificial neural networks may yield similar or better results than the present approach.

The empirical formulation determined through a principal-component regression contains many of the same variables present in the more complicated theoretical schemes for latent heat flux. This result implies that the statistical approach reveals predictor variables that match the underlying physical processes driving the flux of latent heat. However, this is just one of many possible results; the empirical relations may need slight modifications for different land cover types, particularly for the less influential terms in the regression, the coefficients of which may serve as tuning parameters for different locations. The current study, while limited in scope, provides some evidence that such an approach would work well, particularly in light of the positive results of independent tests of the empirical scheme in Iowa. Future studies could apply a similar approach to flux measurements collected in different vegetation regimes to create empirical flux formulations for several biomes. This could lead to rapid improvements in model flux predictions.

The dominant term in both the E_{dir} and E_t equations in the empirical latent heat flux parameterization requires

a measure of soil moisture. This suggests that including soil moisture alone in model initializations has the potential to improve maximum daily air temperature forecasts by 2°–4°C. This result underscores the importance of deploying a widespread soil moisture monitoring network.

While the observations from Oklahoma comprise a wide range of temperature, moisture, wind, and vegetation and soil conditions, indicating the applicability of the new latent heat flux parameterization to new locations across North America, the behavior of the new scheme remains unclear during precipitation events and when the ground lies under snow cover. These remaining questions warrant further testing. Nevertheless, the scheme improves short-term forecasts of surface energy fluxes over crops and grassland for the selected case studies. These improved forecast fluxes, which directly affect more tangible temperature and moisture variables, have many implications for agriculture, energy, transportation, and other weather-sensitive industries.

Unfortunately, the new latent heat flux formulation does not alleviate the remaining problems in the predicted sensible and ground heat fluxes. More accurate sensible and ground heat fluxes are needed for realistic surface energy balances and further-improved air temperature and moisture forecasts. It may be that using a principal-component regression approach can yield improvements in the predictions of sensible heat flux. For ground heat flux, the four soil layers currently in the Noah LSM may fall short of the number of soil levels required to accurately represent soil processes (e.g., Santanello and Carlson 2001). Further research using unique datasets to accurately specify the land surface state is needed to develop and evaluate new parameterizations for surface fluxes.

Acknowledgments. The authors thank the Oklahoma Climatological Survey for providing Oklahoma Mesonet data and Drs. Ken Crawford, Peter Lamb, Lance Leslie, Michael Richman, and May Yuan and three anonymous reviewers for their helpful comments. Dr. Jim Merchant and Roberto Bonifaz of the University of Nebraska—Lincoln provided the fractional vegetation coverage and leaf area index data. Funding was provided under NSF Grant ATM-0243720 and by NOAA/Office of Oceanic and Atmospheric Research under NOAA—University of Oklahoma Cooperative Agreement NA17RJ1227, U.S. Department of Commerce.

REFERENCES

Abramowitz, G., 2005: Towards a benchmark for land surface models. *Geophys. Res. Lett.,* **32,** L22702, doi:10.1029/2005GL024419.

——, H. Gupta, A. Pitman, Y. Wang, R. Leuning, H. Cleugh, and K.-L. Hsu, 2006: Neural error regression diagnosis (NERD): A tool for model bias identification and prognostic data assimilation. *J. Hydrometeor.,* **7,** 160–177.

——, A. Pitman, H. Gupta, E. Kowalczyk, and Y. Wang, 2007: Systematic bias in land surface models. *J. Hydrometeor.,* **8,** 989–1001.

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

Barnes, S. L., 1973: Mesoscale objective analysis using weighted time-series observations. NOAA Tech. Memo. ERL NSSL-62, National Severe Storms Laboratory, 60 pp.

Barr, A. G., K. M. King, T. J. Gillespie, G. Den Hartog, and H. H. Neumann, 1994: A comparison of Bowen ratio and eddy correlation sensible and latent heat flux measurements above deciduous forest. *Bound.-Layer Meteor.,* **71,** 21–41.

——, K. Morgenstern, T. A. Black, J. H. McCaughey, and Z. Nesic, 2006: Surface energy balance closure by the eddy-covariance method above three boreal forest stands and implications for the measurements of the CO_2 flux. *Agric. For. Meteor.,* **140,** 322–337.

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.

Bastidas, L. A., H. V. Gupta, S. Sorooshian, W. J. Shuttleworth, and Z. L. Yang, 1999: Sensitivity analysis of a land surface scheme using multicriteria methods. *J. Geophys. Res.,* **104D,** 19 481–19 490.

Betts, A. K., 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.

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, 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, 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.-H., 1968: *Climate and Agriculture: An Ecological Survey*. Aldine Publishing Company, 304 pp.

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.,* **100D,** 25 595–25 606.

Chen, F., 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.

——, and Coauthors, 1996: Modeling of land-surface evaporation by four schemes and comparison with FIFE observations. *J. Geophys. Res.,* **101,** 7251–7268.

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.

Crawford, T. M., D. J. Stensrud, 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.

Curran, P. J., 1983: Multispectral remote sensing for the estimation of green leaf area index. *Philos. Trans. Roy. Soc. London,* **309B,** 257–270.

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

Dixon, M. J., and J. Grace, 1984: Effect of wind on the transpiration of young trees. *Ann. Bot.,* **53,** 811–819.

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.

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

——, 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.

El-Sharkawy, M. A., J. H. Cock, A. del P. Hernandez, 1985: Stomatal response to air humidity and its relation to stomatal density in a wide range of warm climate species. *Photosynth. Res.,* **7,** 137–149.

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.

Godfrey, C. M., and D. J. Stensrud, 2008: Soil temperature and moisture errors in operational Eta Model analyses. *J. Hydrometeor.,* **9,** 367–387.

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.]

Gupta, H. V., L. A. Bastidas, S. Sorooshian, W. J. Shuttleworth, and Z. L. Yang, 1999: Parameter estimation of a land surface scheme using multicriteria methods. *J. Geophys. Res.,* **104D,** 19 491–19 503.

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.

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

Haugland, M. J., and K. C. Crawford, 2005: The diurnal cycle of land–atmosphere interactions across Oklahoma's winter wheat belt. *Mon. Wea. Rev.,* **133,** 120–130.

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.

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.

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.

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.,* **104D,** 19 569–19 585.

Koster, R. D., and Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Science,* **305,** 1138–1140.

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.

Lee, X., and A. T. Black, 1994: Relating eddy correlation sensible heat flux to horizontal sensor separation in the unstable atmospheric surface layer. *J. Geophys. Res.,* **99D,** 18 545–18 553.

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.

Lydolph, P. E., 1964: The Russian Sukhovey. *Ann. Assoc. Amer. Geogr.,* **54,** 291–309.

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.

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

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.* 3rd ed. Cambridge University Press, 453 pp.

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.,* **86C,** 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.

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.,* **102D,** 16 663–16 682.

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.

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

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

Pielke, R. A., 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.

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.

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

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.

——, and Coauthors, 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.

Ronda, R. J., H. A. R. de Bruin, and A. A. M. Holtslag, 2001: Representation of the canopy conductance in modeling the surface energy budget for low vegetation. *J. Appl. Meteor.,* **40,** 1431–1444.

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.

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.

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. Oceanic Technol.,* **17,** 474–494.

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, 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.

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

Wei, M.-Y., Ed., 1995: Soil moisture: Report of a workshop held in Tiburon, California, 25–27 January 1994. NASA Conference Publ. 3319, 80 pp.

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.

Wilson, K., and Coauthors, 2002: Energy balance closure at FLUXNET sites. *Agric. For. Meteor.,* **113,** 223–243.

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.

Xia, Y., A. J. Pitman, H. V. Gupta, M. Leplastrier, A. Henderson-Sellers, and L. A. Bastidas, 2002: Calibrating a land surface model of varying complexity using multicriteria methods and the Cabauw Dataset. *J. Hydrometeor.,* **3,** 181–194.

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

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. *Photogramm. Eng. Remote Sens.,* **63,** 69–77.

Zamora, R. J., and Coauthors, 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.

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.