Retrieval of an Available Water-Based Soil Moisture Proxy from Thermal Infrared Remote Sensing. Part I: Methodology and Validation

CHRISTOPHER R. HAIN AND JOHN R. MECIKALSKI

Department of Atmospheric Science, University of Alabama in Huntsville, Huntsville, Alabama

MARTHA C. ANDERSON

Hydrology and Remote Sensing Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, Maryland

(Manuscript received 19 March 2008, in final form 11 October 2008)

ABSTRACT

A retrieval of available water fraction ($f_{AW}$) is proposed using surface flux estimates from satellite-based thermal infrared (TIR) imagery and the Atmosphere–Land Exchange Inversion (ALEXI) model. Available water serves as a proxy for soil moisture conditions, where $f_{AW}$ can be converted to volumetric soil moisture through two soil texture dependents parameters—field capacity and permanent wilting point. The ability of ALEXI to provide valuable information about the partitioning of the surface energy budget, which can be largely dictated by soil moisture conditions, accommodates the retrieval of an average $f_{AW}$ over the surface to the rooting depth of the active vegetation. For this method, the fraction of actual to potential evapotranspiration ($f_{PET}$) is computed from an ALEXI estimate of latent heat flux and potential evapotranspiration (PET). The ALEXI-estimated $f_{PET}$ can be related to $f_{AW}$ in the soil profile. Four unique $f_{PET}$ to $f_{AW}$ relationships are proposed and validated against Oklahoma Mesonet soil moisture observations within a series of composite periods during the warm seasons of 2002–04. Using the validation results, the most representative of the four relationships is chosen and shown to produce reasonable (mean absolute errors values less than 20%) $f_{AW}$ estimates when compared to Oklahoma Mesonet observations. Quantitative comparisons between ALEXI and modeled $f_{AW}$ estimates from the Eta Data Assimilation System (EDAS) are also performed to assess the possible advantages of using ALEXI soil moisture estimates within numerical weather prediction (NWP) simulations. This TIR retrieval technique is advantageous over microwave techniques because of the ability to indirectly sense $f_{AW}$—and hence soil moisture conditions—extending into the root-zone layer. Retrievals are also possible over dense vegetation cover and are available on spatial resolutions on the order of the native TIR imagery. A notable disadvantage is the inability to retrieve $f_{AW}$ conditions through cloud cover.

1. Introduction

The advent of sophisticated land surface models (LSMs) within a numerical weather prediction (NWP) framework has lead to an increase in accuracy in forecasting atmospheric phenomena, especially in the mid-latitude continental warm season (Chen and Dudhia 2001; Pleim and Xiu 1995). Although NWP has improved since the implementation of LSMs, the overall performance of a forecast is still highly dependent on the initialization of soil and vegetation variables, such as soil moisture, fraction of green vegetation cover ($f_v$), and leaf area index (LAI; Wetzel and Chang 1987; van den Hurk et al. 1997; Crawford et al. 2001; Kurkowski et al. 2003; Richter et al. 2004). These variables are not readily or easily observed on the scales needed to initialize NWP models, and therefore such models often rely on parameterized routines to provide a representation of the current land surface state. Soil moisture is an especially difficult LSM state variable because observations covering the United States cannot be obtained with adequate spatial resolution. Even within relatively dense observation networks (~1 observation site 35 km$^{-2}$) such as the Oklahoma Mesonet, it is challenging to implement observed soil moisture within a NWP model.
Soil moisture conditions can exhibit large variations even on small spatial scales (<1 km) because of heterogeneity in rainfall patterns and soil texture.

Because of the lack of high-resolution soil moisture observations, operational NWP models rely on sophisticated land surface data assimilation systems, which, in turn, depend on parameterized model physics and observed precipitation to estimate the current soil moisture state. In the United States, the most widely used systems include the Land Data Assimilation System (LDAS; Mitchell et al. 2004) and the Land Information System (LIS; Kumar et al. 2006). Errors in observed precipitation analyses and unrepresentative model physics or parameterizations can often cause the continually updated soil moisture fields within the model to “drift” from reality (Dirmeyer 2000; Li et al. 2005).

Issues resulting from the lack of an operational ground-based soil moisture observing system can potentially be alleviated by the development of soil moisture retrieval algorithms applied to satellite-based observations in both the thermal infrared (TIR; moisture retrieval algorithms applied to satellite-based data) and microwave [1.0 (L band); 6.9 (C band); 10.7 GHz (X band)] wavelengths. Several methods have been developed to use TIR observations of the earth’s surface to diagnose both partitioning of fluxes within the surface energy budget and soil moisture conditions (Idso et al. 1975; Price 1980; Carlson et al. 1981; Price 1983; Carlson 1986; Taconet et al. 1986; Carlson et al. 1994; McNider et al. 1994). Several of these methods have built upon the development of the “triangle” method initially developed by Price (1980) and later refined by Carlson et al. (1994), which relates a satellite-derived vegetation index, such as normalized difference vegetation index (NDVI), and surface skin temperature ($T_{\text{skin}}$) to a soil moisture availability term. Gilles and Carlson (1995) stated that a triangular distribution of pixels should emerge under the assumption that a full dynamic range of NDVI values and surface soil moisture conditions are present within the modeling domain, an assumption which may be difficult to fulfill on a consistent temporal and spatial basis. Additional methods have been developed by McNider et al. (1994) and Jones et al. (1998) in which satellite-derived TIR $T_{\text{skin}}$ heating rates are directly assimilated within the surface-layer parameterization of an NWP model. This assimilation method assumes that the largest errors in the modeled rate of $T_{\text{skin}}$ heating, computed with respect to the surface energy budget, result from the latent heat flux term, which is largely controlled by soil moisture conditions. Methods that use TIR observations of surface $T_{\text{skin}}$ to estimate components of surface energy budget have also been developed and are especially advantageous, because they provide information with respect to the partitioning of surface fluxes of sensible and latent heat (Norman et al. 1995; Anderson et al. 1997; Bastiaanssen et al. 1997; Mecikalski et al. 1999). In this approach, the diagnosed surface flux partitioning is used to retrieve available water ($f_{AW}$), which can be used as a proxy for soil moisture conditions.

This method will use the TIR-based Atmosphere–Land Exchange Inversion (ALEXI) surface energy balance model (Anderson et al. 1997; Mecikalski et al. 1999; Anderson et al. 2007a). ALEXI provides estimates of each component of the surface energy budget, which, in turn, are used to develop a retrieval of $f_{AW}$. The retrieved $f_{AW}$ are then validated against in situ soil moisture observations taken at sites in the Oklahoma Mesonet. Retrieval performance is also compared with Eta Data Assimilation System (EDAS) $f_{AW}$ estimates, which serve as a primary soil moisture initialization dataset for mesoscale NWP simulations.

Section 2 gives a brief description of the development and implementation of ALEXI. Section 3 provides a detailed methodology of the $f_{AW}$ retrieval. Section 4 outlines the datasets used in the retrieval and validation of ALEXI $f_{AW}$. Section 5 presents a detailed statistical verification of each proposed retrieval relationship and a quantitative comparison with an offline LSM-modeled $f_{AW}$.

2. ALEXI model description

The ALEXI model was formulated as an extension to the two-source (TSM) model of Norman et al. (1995). The TSM was developed to address many of the disadvantages with respect to the monitoring of surface energy fluxes from satellite-based platforms. Within the two-source approximation, the radiometric temperature ($T_{\text{rad}}$) of a vegetated surface is represented as the ensemble average of the individual temperatures of both the soil ($T_s$) and vegetation ($T_c$) components, partitioned by the fractional vegetation cover [$f(\phi)$] apparent from the thermal sensor zenith view angle ($\phi$):

$$T_{\text{RAD}}(\phi) \approx [f(\phi)T_c + (1 - f(\phi))T_s].$$

The TSM separately balances the energy budgets for the soil (subscript $s$) and canopy (subscript $c$) components of the composite system, solving for system and component fluxes of net radiation (RN), latent heat (LE), sensible heat (H), and ground heat conduction (G):

$$\text{RN} = H + LE + G,$$
$$\text{RN}_s = H_s + LE_s + G,$$  \[\text{(2)}\]
$$\text{RN}_c = H_c + LE_c.$$
Using an upper boundary condition of air temperature along with turbulent transport resistances, the values of $H$ for both the soil and canopy components of the system can be computed, using

$$\begin{align*}
H &= H_s + H_c = \rho c_p \frac{T_{ac} - T_a}{R_s}, \\
H_s &= \rho c_p \frac{T_s - T_{ac}}{R_s}, \quad \text{and} \\
H_c &= \rho c_p \frac{T_c - T_{ac}}{R_s},
\end{align*}$$

where $T_a$, $T_s$, and $T_c$ represent the air temperature above the canopy, soil temperature, and canopy temperature, respectively; $T_{ac}$ represents the air temperature inside the canopy; $\rho$ is air density; $c_p$ is the specific heat of air; and $R_s$, $R_a$, and $R_c$ represent the specific turbulent transport resistances for the atmosphere, soil surface, and bulk leaf boundary layer, respectively. The value of the $G$ is parameterized as 30% of the net radiation at the soil surface, $RN_s$ (Choudhury et al. 1987), which is related to $RN$ through a simple two-stream model of radiative transfer through the canopy (Kustas and Norman 2000). The canopy transpiration rate, $LE_c$, is initially estimated using a modified Priestley–Taylor approximation applied to the net radiation divergence within the canopy, $RN_c$, while the soil evaporation component, $LE_s$, is computed as the residual to the surface energy budget. Stressing canopies will not transpire at the potential Priestley–Taylor rate, causing $LE_s$ to be overestimated and inconsistent with the diagnosed canopy temperature, $T_c$, and $LE_s$ to become negative in compensation. Condensation onto the soil surface is unlikely midday, so such solutions are flagged as an indication of vegetation stress, and $LE_s$ is reduced until $LE_s$ is 0. Thus, when available water becomes limiting, both $LE_c$ and $LE_s$ will be less than the potential rate. The TSM is an advancement over simpler “single source” models that do not explicitly treat the differences in atmospheric coupling between the soil ($R_s$) and canopy ($R_c$), which result in poor performance over sparse and partial canopy cover (Anderson et al. 1997; Norman et al. 1995).

ALEXI couples the TSM land surface representation with a simple atmospheric boundary layer (ABL) model, alleviating the need for surface meteorological observations of air temperature and allowing application over large spatial regions. In this coupled mode, ALEXI simulates $T_a$ at the blending height internally within the ABL model, ensuring that $T_a$ is consistent with the modeled surface fluxes. The TSM is applied at 2 times during the morning hours—usually 1.5 h after local sunrise ($t_1$) and 5.5 h after local sunrise ($t_2$)—using radiometric temperature estimates derived from a geostationary satellite, such as the Geostationary Operational Environmental Satellite (GOES). The ABL component relates the rise in $T_a$ in the mixed layer over this interval to the time-integrated influx of $H$ from the surface, thus providing a means for energy closure with modeled surface fluxes from the TSM component (McNaughton and Spriggs 1986; Anderson et al. 1997).

The coupling of the ABL within ALEXI is advantageous because it moves the upper boundary conditions in temperature from the near-surface to the “blending height,” where conditions are more uniform at a spatial scale of a geostationary satellite thermal pixel (typically $>5$ km). The model is sensitive to the time-differential change in $T_{rad}$, reducing the effect of bias errors (BEs) in $T_{rad}$ as a result of atmospheric corrections, sensor calibration, and surface emissivity specification (Kustas et al. 2001).

### 3. Available water retrieval methodology

The TSM partitions the bulk latent heat flux into soil evaporation ($LE_s$) and canopy transpiration ($LE_c$) components. These fluxes, in turn, are largely controlled by $f_{AW}$ in both the surface layer (0–5 cm), and the root zone layer (5–200 cm), respectively—with decreasing available water in these two pools serving to decrease $LE_s$ and $LE_c$ from their respective potential rates. Anderson et al. (2007b) describe a technique for simulating the effects of soil moisture on latent heat estimates from ALEXI using a soil moisture stress function relating the value of the fraction of potential evapotranspiration (ET) $f_{PET}$:

$$f_{PET} = \frac{LE}{PET}$$

to the fraction of available water ($f_{AW}$) in the soil profile:

$$f_{AW} = \frac{\theta - \theta_{wp}}{\theta_{fc} - \theta_{wp}},$$

where $\theta_{fc}$ and $\theta_{wp}$ are the volumetric soil water contents at field capacity and wilting point, respectively. This retrieval of $f_{AW}$ represents an integrated value from the top of the soil surface to the level at which root extraction is negligible. In this case, the stress function is applied to the $f_{PET}$ computed for the system (soil + canopy):

$$f_{PET(System)} = \frac{LE_s + LE_c}{PET_s + PET_c}$$
to probe average \( f_{AW} \) within the soil profile. The use of a soil moisture stress function can perform well when \( f_{AW} \) conditions are between \( \theta_c \) and \( \theta_{wp} \), though when \( f_{AW} \) conditions occur beyond this range, this particular methodology may not fully capture such conditions. For instance, shortly after a heavy precipitation event, \( \theta \) can be observed above \( \theta_c \) and can reach the saturation point (porosity); the use of a soil moisture stress function is not able to capture this particular signal and would lead to an underestimation of the actual \( \theta \) conditions. A similar error may occur when \( \theta \) is below \( \theta_{wp} \), which can lead to an overestimation of actual \( \theta \) conditions.

### a. Computation of potential evapotranspiration

The PET is the rate of evaporation that would occur under optimal conditions over a surface with no soil moisture stress. There are several documented methods for estimating PET, requiring varying amounts of surface and meteorological information (Penman 1948; Monteith 1975). To be consistent with the canopy transpiration computation in the TSM, a Priestley–Taylor approximation for PET (Priestley and Taylor 1972) is used, modified to partition PET between the canopy (PET$_c$) and the soil (PET$_s$) components of the system.

Potential canopy transpiration (PET$_c$) is related to RN intercepted by the canopy (Norman et al. 1995; Tanner and Sinclair 1986):

\[
\text{PET}_c = \alpha_c \frac{S}{S + \gamma} \text{RN}_c,
\]

where \( \alpha_c = 1.26 \), \( S \) is the slope of the curve of the saturation vapor pressure versus temperature, and \( \gamma \) is the psychrometric constant (0.066 kPa °C$^{-1}$).

Similarly, potential soil evaporation (PET$_s$) is related to RN$_s$, following Tanner and Jury (1976):

\[
\text{PET}_s = \alpha_s \frac{S}{S + \gamma} \text{RN}_s,
\]

where the value of \( \alpha_s \) is computed as a function of the value of a canopy transmission factor and approaches unity (equilibrium evaporation) under very dense canopies (see Anderson et al. 2007b).

### b. Relationships between \( f_{PET} \) and \( f_{AW} \)

A wide variety of relationships between \( f_{PET} \) and \( f_{AW} \) have been proposed in past studies. The most commonly used relationships include linear functions (Mahfouf and Noilhan 1991; Wetzel and Chang 1986), piecewise linear or threshold functions (Mahrt and Pan 1984), and nonlinear functions (Campbell and Norman 1998; Jarvis 1976). Stress functions requiring detailed information regarding soil and vegetation type are less easily implemented on a regional or continental scale.

Here, we consider four possible functionals relating \( f_{PET} \) and \( f_{AW} \), referred to as Noah, LINEAR, NONLINEAR, and BLEND, as described below.

#### 1) Noah

The Noah functional (Fig. 1d) is built on the relationships used in the Noah LSM to compute direct soil evaporation and canopy transpiration (Chen and Dudhia 2001). The formulation of direct soil evaporation uses a simple exponential relationship between \( f_{PET} \) and \( f_{AW} \). Canopy transpiration is computed from a linear relationship between \( f_{PET} \) and the plant coefficient \( B_c \), which contains a complex nonlinear canopy resistance formulation. Here, \( B_c \) represents the influence of stomatal control on canopy transpiration and can be expressed as

\[
B_c = \frac{1 + \Delta}{1 + R_c C_h + \frac{\Delta}{R_r}},
\]

where \( R_r \) is a function of surface air temperature, surface pressure, and the surface exchange coefficient \( (C_h) \); and \( R_c \) is the canopy resistance. The canopy resistance term is computed following the formulation of Jacquemin and Noilhan (1990) and can be expressed as

\[
R_c = \frac{R_{min}}{\text{LAI}(F_1 F_2 F_3 F_4)},
\]

where \( F_1, F_2, F_3, \) and \( F_4 \) have values that range from 0 to 1 and represent the effects of solar radiation, vapor pressure deficit, air temperature, and root-zone soil moisture on the value of canopy resistance (Chen and Dudhia 2001); and \( R_{min} \) is the minimum stomatal resistance, which is parameterized as a function of vegetation type and LAI. A complete explanation of each term within the formulation of \( R_c \) and \( B_c \) can be found in Jacquemin and Noilhan (1990) and Chen and Dudhia (2001). The soil moisture forcing term \( (F_4) \) is an integrated value of available water in the soil profile \( (f_{AW}) \) and can be solved for by rearranging Eqs. (12) and (13), yielding

\[
f_{AW} = \frac{R_{min}}{(\text{LAI})F_1 F_2 F_3} \left( 1 + \frac{\Delta}{f_{PET,LENI}} - \frac{\Delta}{R_c C_H} - \frac{1}{C_H} \right).
\]

Figure 1 shows a graphical representation relationship between the retrieved \( f_{AW} \) and \( f_{PET} \) from the Noah LSM method over the range of 0%–100% \( f_{AW} \).

#### 2) LINEAR

The second proposed method uses a simplified linear relationship (Fig. 1a) between \( f_{PET} \) and \( f_{AW} \) suggested...
by Wetzel and Chang (1986), who set the value of $f_{\text{PET}}$ to be 85% of the value of $f_{\text{AW}}$, where the 85% represented a constant typical value of $B_c$. Here, the value for the plant coefficient ($B_c^*$) is allowed to vary across the domain according to Eq. (12), but it sets each of the four forcing terms in Eq. (13) to optimal conditions (i.e., 1.0), yielding

$$f_{\text{PET}} = B_c^* f_{\text{AW}}. \quad (14)$$

3) NONLINEAR

The third method, NONLINEAR (Fig. 1c), proposes a nonlinear relationship but to a lesser degree than that of the Noah relationship. The functional uses $B_c^*$ and the observed $f_{\text{PET}}$ to compute a coefficient that varies to produce the nonlinearity and can be expressed as

$$f_{\text{PET}} = \frac{f_{\text{PET}}}{(2B_c^* f_{\text{PET}})} f_{\text{AW}}. \quad (15)$$

4) BLEND

Analyses of the three aforementioned relationships showed significant biases with respect to soil moisture observations in certain moisture regimes. NONLINEAR and Noah showed a substantial dry bias when observed $f_{\text{AW}}$ or ALEXI $f_{\text{PET}}$ was high, while the LINEAR method has a substantial wet bias when observed $f_{\text{AW}}$ or ALEXI $f_{\text{PET}}$ was low. Using these findings, a fourth method was proposed, which weights LINEAR and NONLINEAR based on a function of the observed $f_{\text{PET}}$ value. The BLEND method (Fig. 1b) can be represented by

$$f_{\text{AW}} = \frac{f_{\text{PET}}}{B_c^*} (f_{\text{PET}}) + \frac{f_{\text{PET}}}{(2B_c^* f_{\text{PET}})} (1 - f_{\text{PET}}). \quad (16)$$

When the observed value of $f_{\text{PET}}$ is high, the BLEND method will heavily weight the result toward the LINEAR relationship, which exhibits a wet bias when $f_{\text{PET}}$ is low. However, when $f_{\text{PET}}$ is low, it will heavily weight the NONLINEAR relationship, thus attempting to correct its observed bias.

4. Data sources

a. ALEXI

Mecikalski et al. (1999) applied ALEXI spatially over the central United States using high-resolution remotely sensed surface temperature and $f_c$ data aggregated to a 10-km grid. The hourly GOES surface blackbody brightness temperatures ($T_b$) are subjected to a simple cloud-clearing algorithm, stipulating that the 10-km grid cell must be cloud-free over the interval $t_1$ to $t_2$ to avoid the corruption of the surface temperature heating curve. At-sensor brightness temperature data $T_b$ are corrected for atmospheric and surface emissivity effects following Anderson et al. (2007a) and converted to radiometric temperature values $T_{\text{rad}}$. The directional fraction of green vegetation cover [$f(\phi)$], needed for the partitioning of $T_{\text{rad}}$ with respect to the soil and canopy components within ALEXI [Eq. (1)], is computed from the 1-km Moderate Resolution Imaging Spectroradiometer (MODIS)/Terra 8-day composite LAI product (Myneni et al. 2002; Anderson et al. 2007a). Specific canopy characteristics, such as roughness length and leaf size, required to compute the canopy coupling coefficient $R_c$ are estimated from 1-km maps of $f_c$ and landcover class (Hansen et al. 2000), aggregated to the 10-km ALEXI grid following Anderson et al. (2007a). Downwelling solar and longwave radiation, required to compute net radiation, are estimated at each grid cell using GOES-based insolation products at spatial resolution of 20-km, which are interpolated to the 10-km ALEXI grid (Diak et al. 1996, 2000). The upwelling components of RN are predicted by the TSM, based on the diagnosed values of $T_s$ and $T_c$ and estimates of surface emissivity and albedo (Mecikalski et al. 1999). Surface observations of wind speed $T_c$, vapor pressure ($\varepsilon$), and surface pressure from Eta model initialization fields are analyzed to a 40-km grid using the Cooperative Institute for Meteorological Satellite Studies (CIMSS) Regional Assimilation System (CRAS; Diak et al. 1992).
This step also provides gridded atmospheric profiles of potential temperature needed by the ABL component, valid at \( t_1 \).

**b. ALEXI \( f_{\text{PET}} \) composites**

ALEXI was executed over the contiguous United States at a spatial resolution of 10 km for each day in the period of April–September 2002–04. However, the cloud-free constraint needed during the period of \( t_1 \) to \( t_2 \) can cause significant data coverage gaps caused by cloudy-sky conditions. To minimize resulting cloud-induced data gaps in the retrieved \( f_{\text{AW}} \), and to provide some noise reduction, \( f_{\text{PET}} \) values used in this study have been composited over 2–5-day intervals. Table 1 lists each composite period used in the validation exercises in section 5. The length of the compositing interval was selected to provide clear-sky coverage over the entire modeling domain. Over the compositing interval, clear-sky values of \( f_{\text{PET}} \) at a given pixel were averaged with equal weight. This weighting assumes that the day-to-day variability of ALEXI surface flux estimates at a given pixel is not substantially greater than the degree that the average profile value of soil moisture will change during the compositing interval.

c. EDAS soil moisture estimates

Retrievals of ALEXI \( f_{\text{AW}} \) are compared to the soil moisture estimates provided by EDAS, valid at 1200 UTC on the final day of each composite period. The EDAS provides estimates of volumetric soil moisture (m\(^3\) m\(^{-3}\)) in four layers: 0–10, 10–40, 40–100, and 100–200 cm. The volumetric soil moisture estimates are converted to \( f_{\text{AW}} \) by Eq. (8), using values of \( \theta_w \) and \( \theta_{wp} \) extracted from the EDAS soil texture input fields, representing parameterized values based on the Zobler 9-class soil texture classification (Zobler 1986) and then interpolated to the ALEXI grid. The EDAS soil moisture estimates are effectively rescaled to generate EDAS estimates of \( f_{\text{AW}} \), which are used in ETA in determining evaporative land surface fluxes. The conversion to \( f_{\text{AW}} \) allows for a consistent comparison between ALEXI and EDAS \( f_{\text{AW}} \).

The EDAS analyses are chosen for comparison because they provide the initial soil moisture conditions for the operational National Centers for Environmental Prediction (NCEP)Eta Model (1996–2006), and a similar system is currently used for soil moisture initializations in the operational NCEP North American Mesoscale (NAM) model (2006–present).

d. Oklahoma Mesonet soil moisture observations

The Oklahoma Mesonet is an automated network of 115 remote meteorological stations distributed more or less uniformly across the state. Soil moisture measurements are taken at 60 of these sites (Fig. 2), providing estimates of soil moisture at the depths of 5, 25, 60, and 75 cm. The Mesonet provides the highest spatial and temporal density of soil moisture observations of any soil moisture network in the United States (Brock et al. 1995).

To collect soil moisture information, the Oklahoma Mesonet uses Campbell Scientific model 229-L heat dissipation sensors to retrieve soil matric potential (Basara and Crawford 2000). The fractional water index (FWI), which is a normalized signal of the sensor response, can be computed from the observed \( \Delta T \). Values of FWI range from 0.0 to 1.0, where a value of 0.0 represents soil at the permanent wilting point and a value of 1.0 represents soil at field capacity (Schneider et al. 2003). FWI values are computed by

\[
\text{FWI} = \frac{\Delta T_d - \Delta T_{\text{ref}}}{\Delta T_d - \Delta T_w},
\]

where \( \Delta T_{\text{ref}} \) is the measured temperature change by the sensor, \( \Delta T_d \) is the sensor response when dry (3.96°C), and \( \Delta T_w \) is the sensor response when wet (1.38°C; Schneider et al. 2003). The observations at the four sensor depths have been weighted and averaged to represent an average \( f_{\text{AW}} \) in the layer of 0–100 cm. Instantaneous observations taken at 1200 UTC on the

---

**Table 1. List of composite periods in the study.**

<table>
<thead>
<tr>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>15–19 Jun 2002</td>
</tr>
<tr>
<td>11–12 May 2003</td>
</tr>
<tr>
<td>28–29 May 2003</td>
</tr>
<tr>
<td>4–5 Jul 2003</td>
</tr>
<tr>
<td>27–31 Jul 2003</td>
</tr>
<tr>
<td>6–7 May 2004</td>
</tr>
<tr>
<td>1–2 Jun 2004</td>
</tr>
<tr>
<td>1–3 Aug 2004</td>
</tr>
</tbody>
</table>

---

**Fig. 2. Geographical locations of Oklahoma Mesonet FWI observations used in the validation.**
final day of each composite period are compared to composited ALEXI $f_{AW}$ retrievals. This temporal co-
dordination assumption was tested during each composite
day to quantify the average change in 0–100-cm FWI
across the study domain. The domain-averaged 0–100-cm
FWI only changes on the order of 1%–2% during a 5-day
compositing period, while domain-averaged ALEXI $f_{PET}$
varies on the order of 5%–10% during the same period.
The larger daily variability with respect to ALEXI $f_{PET}$
is caused by errors in atmospheric correction and cloud
detection, and interpolation of necessary input fields.
Additionally, the comparison of FWI observations with
ALEXI and EDAS $f_{AW}$ estimates is facilitated by the
assumption that FWI observations are equivalent to
$f_{AW}$. Basara and Crawford (2000) describe in detail the
retrieval of soil matric potential and subsequently vol-
umetric soil moisture from $\Delta T_{eff}$ measurements. Re-
trieval of soil matric potential from each site can be
used to compute site-specific variables of field capacity
(matric potential equal to $-33$ kPa) and permanent
wilting point (matric potential equal to $-1500$ kPa) in
terms of volumetric soil moisture. Using these site-
specific values, $f_{AW}$ can be computed at each observa-
tion site and compared to the FWI observation valid at
each site. A quantitative comparison between the
computed $f_{AW}$ and FWI show a correlation coefficient
of 0.97 (not shown), which validates the assumption of
equivalence between FWI and ALEXI $f_{AW}$.

5. Results

The validation of each of the four $f_{PET}$ to $f_{AW}$ rela-
tionships proposed for the retrieval of ALEXI $f_{AW}$,
along with $f_{AW}$ from EDAS, was performed for each of
the composites periods listed in Table 1. Statistical
metrics of BE, mean absolute error (MAE), and root-
mean-square difference (RMSD) were calculated to
assess the accuracy of each $f_{PET}$ and $f_{AW}$ relationship
and EDAS $f_{AW}$.

a. Validation of the $f_{PET}$ to $f_{AW}$ relationships

Retrieved $f_{AW}$ from each of the four proposed $f_{PET}$ to
$f_{AW}$ relationships were compared to observations of
FWI. Scatterplot comparisons are shown in Fig. 3, with
associated statistics in Table 2. In general, the BLEND
method performs best, exhibiting the smallest MAE
(18% $f_{AW}$ and an $R^2$ value of 0.48; Fig. 3a). The LINEAR
and NONLINEAR methods yielded larger errors
(19% $f_{AW}$ and 22% $f_{AW}$, respectively; Figs. 3b and 3c).
The Noah relationship exhibits the largest retrieval er-
ror, with an MAE of 29% $f_{AW}$ and an $R^2$ value of 0.22
(Fig. 3d). Neither BLEND nor LINEAR exhibited any
substantial bias over the entire validation dataset,
whereas both NONLINEAR and Noah showed large
negative biases. Frequency histograms of MAE values
for each of the four relationships shows MAE values
exceeding 30% $f_{AW}$ for both BLEND and LINEAR are
less frequent than those associated with both NON-
LINEAR and Noah (Fig. 4). It is also evident that
BLEND and LINEAR exhibit a larger number of ob-
servations, with a MAE value less than 15% $f_{AW}$.

The dependence of retrieval error on $f_c$ and FWI is
also examined for each of the $f_{PET}$ to $f_{AW}$ relationships.
Figures 5a and 5b show frequency histograms of ob-
served $f_c$ and $f_{AW}$ that depict a wide range of observed
$f_c$ values, along with a wide range of $f_{AW}$ conditions
during the eight composite periods. The value of $f_c$ is a
core input in ALEXI, governing the partitioning of
surface temperature and fluxes between the soil and
canopy components of the system. Canopy cover also
influences the rate at which surface temperature will
respond to soil moisture deficiencies. The surface layer
typically dries and heats up rapidly after a rainfall
event, while root-zone moisture and the associated
canopy temperature typically evolve over a longer time
scale. In the thermal band, we have the potential of
obtaining a moisture signal over a broad range of $f_c$ as
root-zone moisture depletion is reflected in the ob-
served canopy temperature. Therefore, it is advanta-
geous to test any retrieval of $f_{AW}$ using ALEXI against
the range of observed values of $f_c$. The MODIS-
derived LAI values varied between 0.5 and 4.0 over the
domain during the study period, translating to $f_c$ be-
tween 20% and 90% (Fig. 5a). In general, western
Oklahoma is much less densely vegetated than eastern
Oklahoma (Fig. 6).

BLEND generally yields smaller negative biases
when $f_c$ is $>40\%$ and generally larger positive biases
$<40\%$ (Fig. 7). NONLINEAR and Noah show in-
creasing negative biases as $f_c$ increases, the magnitude
of these biases are much larger than either BLEND or
LINEAR. BLEND also exhibits lower RMSD values
across the observed range of $f_c$, where even LINEAR
shows larger errors when $f_c$ is $<40\%$ (Fig. 8). NON-
LINEAR and Noah show large RMSD values across the

| Table 2. Error statistics for each of the four proposed $f_{PET}$ to $f_{AW}$ relationships used in the retrieval of ALEXI $f_{AW}$. |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| BE (% $f_{AW}$) | MAE (% $f_{AW}$) | RMSD (% $f_{AW}$) | $R^2$ |
| BLEND          | $-4.3\%$        | $17.7\%$        | $21.3\%$        | $0.48$          |
| LINEAR         | $+5.0\%$        | $18.8\%$        | $23.1\%$        | $0.33$          |
| NONLINEAR      | $-14.1\%$       | $22.1\%$        | $26.8\%$        | $0.33$          |
| Noah           | $-25.7\%$       | $28.6\%$        | $33.3\%$        | $0.26$          |
observed range of $f_c$, with a general increase in RMSD as $f_c$ increases. The performance is also quantified as a function of observed FWI. BLEND shows the smallest RMSD values for a majority of the FWI bins, although LINEAR had slightly lower RMSD errors when FWI was around 50% (Fig. 9). LINEAR shows increasingly large RMSD values when compared to BLEND as FWI decreases to less than 40%. NONLINEAR and EDAS shows some lower RMSD values for observed dry conditions (FWI <30%), but more substantial errors are seen when FWI is greater than 50%.

Using these analyses, the BLEND method appears to be the most representative $f_{PET}$ to $f_{AW}$ relationship. LINEAR performed only slightly worse than BLEND, especially during dry $f_{AW}$ and low $f_c$ conditions. NONLINEAR exhibits a general negative (dry) bias for both wet and dry $f_{AW}$ and across all observed values of $f_c$. Noah shows the largest values of negative (dry) bias, MAE, and RMSD of the four methods. Therefore, the BLEND method will be used as the $f_{PET}$ to $f_{AW}$ relationship for the retrieval of ALEXI $f_{AW}$ for the remainder of this paper.
b. Quantitative comparison of ALEXI and EDAS $f_{\text{AW}}$

A comparison of ALEXI and EDAS $f_{\text{AW}}$ was performed for the same eight composite periods. This section will investigate the differences between ALEXI and EDAS $f_{\text{AW}}$ and illustrate the possible advantages for the implementation of ALEXI $f_{\text{AW}}$ within NWP simulations. Figures 3a and 3e show scatterplots for ALEXI and EDAS $f_{\text{AW}}$ as compared to FWI observations from the Oklahoma Mesonet. EDAS $f_{\text{AW}}$ shows considerably larger scatter, with an $R^2$ of 0.22 as compared to an $R^2$ of 0.48 for ALEXI. EDAS $f_{\text{AW}}$ also exhibits a large negative bias ($-24\% f_{\text{AW}}$), with the largest biases evident during observed wet conditions. The MAE and RMSD values for EDAS are larger than those for ALEXI, with values of 29\% $f_{\text{AW}}$ and 34\% $f_{\text{AW}}$, respectively.

Error frequency histograms for ALEXI and EDAS $f_{\text{AW}}$ are shown in Figs. 4a and 4e, demonstrating that errors greater than 30\% $f_{\text{AW}}$ are much more likely in EDAS than in ALEXI. Bias (Figs. 7a and 7e) and RMSD (Figs. 8a and 8e) values are also shown for ALEXI and
EDAS $f_{AW}$ as a function of $f_c$. The performance of EDAS improves as $f_c$ decreases, where the differences between ALEXI and EDAS are not as great. Large differences in bias and RMSD are evident when $f_c$ is >50%, with the largest EDAS errors occurring between 70% and 90%. ALEXI (Fig. 9a) and EDAS (Fig. 9e) $f_{AW}$ are also compared as a function of observed FWI. ALEXI $f_{AW}$ retrievals performed better than EDAS when the observed FWI was greater than 40%, while EDAS performed better under drier conditions. Large RMSD differences are evident under observed wet conditions, where EDAS RMSD values exceed 36% $f_{AW}$, while errors in ALEXI $f_{AW}$ are around 18%. A significant upgrade was made to the precipitation assimilation scheme used with the EDAS system in early 2004, with a motivation to improve the representation of the land surface parameters, including soil moisture (Fulton et al. 1998). The accuracy of EDAS $f_{AW}$ was investigated for both the preupgrade (2002–03) and postupgrade (2004) periods. During the eight composite periods, there are 215 observations during the preupgrade period and 209 observations for the postupgrade period. The large dry biases are still evident, and the averaged bias and RMSD decreased nearly 7% $f_{AW}$ for the postupgrade observations. These findings are consistent with those of Godfrey and Stensrud (2008), who compared Oklahoma Mesonet soil moisture observations with EDAS soil moisture estimates during the postupgrade period and found substantial dry biases in the EDAS soil moisture estimates. The study showed that the largest dry biases were found in the root-zone (10–100 cm) layer.

c. Conversion of ALEXI and EDAS $f_{AW}$ error statistics to volumetric soil moisture

The validation of ALEXI and EDAS moisture variables in terms of $f_{AW}$ (as opposed to volumetric soil moisture) is advantageous, because it minimizes the effect of errors in defining the soil textural fields, such as field capacity and permanent wilting point. However, soil moisture profiles in most LSMs are initialized in terms of volumetric soil moisture content, with units of cubic meters per cubic meter. ALEXI retrievals of $f_{AW}$ can be converted to volumetric soil moisture estimates using soil mapping data, such as the Zobler (1986) 9-class texture type classification and the State Soil Geographic (STATSGO) 19-class texture type classification, which are both commonly used in LSM and NWP simulations. Table 3 shows BE, MAE, and RMSD values for ALEXI and EDAS in terms of volumetric soil moisture ($m^3/m^3$), computed using wilting point and field capacity values associated with each of the nine soil texture type classifications from Zobler (1986). There are only small differences in the dynamic range between $u_{fc}$ and $u_{wp}$ for each of the texture classes, but it does slightly change the errors statistics across texture type. In general, on average, ALEXI $f_{AW}$ retrievals exhibit BE, MAE, and RMSD of $-0.01$, $0.05$, and $0.06 m^3/m^3$ in volumetric soil moisture units, respectively, while EDAS exhibits BE, MAE, and RMSD of $-0.06$, $0.08$, and $0.09 m^3/m^3$, respectively.

d. Spatial comparison of ALEXI and Oklahoma Mesonet $f_{AW}$

To demonstrate the capabilities of ALEXI in retrieving spatial patterns in soil moisture conditions,
Figs. 10a–h show spatial comparisons of (left) Oklahoma Mesonet FWI and (right) ALEXI $f_{AW}$ during each of the eight composite periods. For the 19 June 2002 composite period (Fig. 10a), the observations indicated very wet conditions across the eastern and central Oklahoma and relatively dry conditions across extreme western Oklahoma. Retrievals of ALEXI $f_{AW}$ show very good agreement with these observations, with a MAE value for this composite period of only 12% $f_{AW}$. ALEXI also accurately retrieves the relatively dry conditions existing across most of Oklahoma during the 1 August 2003 composite period (Fig. 10e), with an MAE value of 15% $f_{AW}$. Generally, each of the eight composite periods showed good spatial agreement between observed FWI and ALEXI $f_{AW}$ values. The BLEND algorithm generated $f_{AW}$ maps that do not suffer from any large spatial biases during the eight composite periods used in this study and performed well during both observed wet and dry FWI conditions. However, there are still some large differences between some point-scale soil moisture observations and ALEXI $f_{AW}$. These differences can likely be attributed to retrieval error, scale mismatch between the point observation and a satellite pixel, or observational quality.

6. Discussion and conclusions

This study has shown that surface flux estimates from the TIR-based ALEXI model are able to provide valuable information about the average available water fraction ($f_{AW}$) in the soil profile. The retrieval of the
Profile-averaged $f_{AW}$ was facilitated by establishing an empirical relationship between $f_{PET}$ and $f_{AW}$. The $f_{PET}$ to $f_{AW}$ formulation found to be most representative of observations from the Oklahoma Mesonet was a weighted average between a simple linear and a nonlinear relationship, namely, the BLEND method. The LINEAR method performs well under high $f_{PET}$ and moderate to dense vegetation but accuracy decreased as $f_{PET}$ and $f_c$ decreased. However, the NONLINEAR method, which exhibited substantial dry bias for high $f_{PET}$ and moderate vegetation, showed increased accuracy as $f_{PET}$ and $f_c$ decreased. This facilitated the development of the BLEND method, as high $f_{PET}$ and $f_c$ conditions favored a linear relationship between $f_{PET}$ and $f_{AW}$, while as $f_{PET}$ decreases nonlinearly between $f_{AW}$ increased. The Noah method, which shows the greatest nonlinearity between $f_{PET}$ and $f_{AW}$, along with being the most complex relationship, exhibited the most substantial errors. A possible explanation for the approach toward linearity from the Noah method may be that...
satellite pixels are essentially spatial integrations of the specific land surface conditions within each pixel, in this case representing an area of 100 km². The blending of the linear and nonlinear relationship may provide the best results because of the varying vegetation types within each pixel, where some vegetation may not be actively transpiring, while other vegetation types may be transpiring at their potential rate. This may especially be the case during “dry down” conditions, when vegetation with shallow root systems begin to experience stress, whereas deeper-rooted vegetation can extract water from deeper in the soil where $f_{AW}$ can still be quite high. Crow and Wood (2002) have shown that relationships, such as canopy transpiration, that are nonlinear at the subpixel scale often approach linearity as the pixel scale increases. Song et al. (2000) also found that a linear relationship between root-zone soil moisture and a soil water extraction term produced more realistic spatial variations in root-zone soil moisture than did nonlinear relationships.

Retrievals of ALEXI $f_{AW}$ were validated against FWI observations (0–75 cm) from the 60 soil moisture
monitoring sites in the Oklahoma Mesonet during eight composite periods during 2002–04. The FWI observations were averaged to provide a $f_{\text{AW}}$ value characteristic of the top meter of the soil profile and compared to ALEXI $f_{\text{AW}}$. The MAE between modeled and measured $f_{\text{AW}}$ for all sites and composite periods was found to be 18%.

Using the BLEND approach, good agreement was found between the TIR-based $f_{\text{AW}}$ retrievals and Mesonet observations averaged across 0–100-cm depth, even at sites with dense vegetation cover ($f_c > 80%;$ Fig. 10). This suggests an advantage over passive microwave-based retrievals of soil moisture, where retrieval accuracy improves over light vegetation cover and reflects moisture conditions in only the top few centimeters of the soil profile (Njoku et al. 2003; Jackson et al. 2007; Crow and Zhan 2007). In addition, thermal data are available at much higher spatial resolution than are microwave products—tens of meters as compared with tens of kilometers. Unlike microwave remote sensing, however, thermal imagery cannot be collected through cloud cover. It appears that there is also a potential
synergy between thermal and microwave soil moisture mapping techniques that has yet to be exploited. Methods for thermal microwave soil moisture product integration should be explored—through joint assimilation, disaggregation, or other data fusion approaches.

Famiglietti et al. (1999) showed that the heterogeneity of soil moisture can show large spatial variations within remote sensing footprints. Some of the scatter apparent in Fig. 3 is likely a result of a mismatch in the scale of the model pixel (10 km) and the scale representative of the soil moisture measurement (on the order of meters). Anderson et al. (2004, 2007a) encountered a similar scale mismatch when comparing ALEXI surface flux estimates valid at \( t_2 \) (\( R_N, H, LE, G \)) directly to flux tower observations (sampling a “footprint” on the land surface of dimension \( \sim 100 \) m). The RMSD between ALEXI model flux estimates (\( R_N, H, LE, G \)) and measurements was significantly reduced from 59 to 34 W m\(^{-2}\) by spatially disaggregating the ALEXI flux maps using higher-resolution TIR and vegetation cover information from the Landsat satellite (TIR resolution of 60 m for \textit{Landsat-7}). The DisALEXI algorithm can also be used to spatially disaggregate ALEXI \( f_{AW} \) maps over measurement sites. Disaggregation experiments are underway to determine what percentage of the errors reported above in the ALEXI \( f_{AW} \) validation is a result of scale mismatch and not retrieval error.

It should also be noted that although these validation results may only represent the Oklahoma Mesonet region, the wide variation of vegetation and land use types across the study region are characteristic of vegetation patterns common to large sections of the continental United States. The validation of the retrieval of ALEXI \( f_{AW} \) was assessed as a function of observed \( f_c \) during the study, and it was shown that there is no substantial increase or decrease in error associated with changes in observed \( f_c \). Additional validation of the ALEXI \( f_{AW} \) retrieval is being conducted in comparison with soil moisture observational networks in other states, including the Nebraska Automated Weather Data Network operated by the High Plains Regional Climate Center. Comparisons with other satellite-based retrieval methods for soil moisture or \( f_{AW} \) (e.g., microwave) and model-based soil moisture estimates are also underway. The feasibility of using ALEXI \( f_{AW} \) retrievals to initialize volumetric soil moisture within mesoscale NWP models will be addressed in a separate paper (Hain et al. 2009, submitted to \textit{J. Hydrometeor.}).

**Acknowledgments.** We gratefully acknowledge Jeff Basara for providing the Oklahoma Mesonet soil moisture dataset (2002–04), which was used in the validation of ALEXI soil moisture retrievals. We also acknowledge Wade Crow for providing insightful comments on the preliminary draft of this manuscript.

---

**TABLE 3. Error statistics for ALEXI and EDAS \( f_{AW} \) converted to volumetric soil moisture (m\(^3\) m\(^{-3}\)) for each of the nine texture classes in the Zobler (1986) soil type classification. Values of field capacity (\( \theta_{fc} \)) and permanent wilting point (\( \theta_{wp} \)) for each texture class are used to convert between \( f_{AW} \) and volumetric soil moisture.**

<table>
<thead>
<tr>
<th>Texture Class</th>
<th>Loamy sand</th>
<th>Silty clay loam</th>
<th>Light clay</th>
<th>Sandy loam</th>
<th>Sandy clay</th>
<th>Clay loam</th>
<th>Sandy clay loam</th>
<th>Loam</th>
<th>Loamy sand</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_{fc} )</td>
<td>0.283 ( \pm ) 0.029</td>
<td>0.387 ( \pm ) 0.119</td>
<td>0.412 ( \pm ) 0.139</td>
<td>0.312 ( \pm ) 0.047</td>
<td>0.338 ( \pm ) 0.100</td>
<td>0.382 ( \pm ) 0.103</td>
<td>0.315 ( \pm ) 0.069</td>
<td>0.329 ( \pm ) 0.066</td>
<td>0.283 ( \pm ) 0.029</td>
</tr>
<tr>
<td>( \theta_{wp} )</td>
<td>0.011 ( \pm ) 0.064</td>
<td>0.012 ( \pm ) 0.065</td>
<td>0.011 ( \pm ) 0.064</td>
<td>0.011 ( \pm ) 0.064</td>
<td>0.010 ( \pm ) 0.061</td>
<td>0.012 ( \pm ) 0.065</td>
<td>0.011 ( \pm ) 0.064</td>
<td>0.011 ( \pm ) 0.064</td>
<td>0.011 ( \pm ) 0.064</td>
</tr>
<tr>
<td>BE (m(^3) m(^{-3}))</td>
<td>-0.010</td>
<td>-0.061</td>
<td>0.044</td>
<td>0.011</td>
<td>0.044</td>
<td>0.011</td>
<td>0.047</td>
<td>0.017</td>
<td>0.047</td>
</tr>
<tr>
<td>MAE (m(^3) m(^{-3}))</td>
<td>-23.9</td>
<td>-21.7</td>
<td>28.9</td>
<td>28.9</td>
<td>28.9</td>
<td>28.9</td>
<td>28.9</td>
<td>28.9</td>
<td>28.9</td>
</tr>
<tr>
<td>RMSD (m(^3) m(^{-3}))</td>
<td>0.054</td>
<td>0.057</td>
<td>0.058</td>
<td>0.057</td>
<td>0.054</td>
<td>0.058</td>
<td>0.057</td>
<td>0.057</td>
<td>0.057</td>
</tr>
</tbody>
</table>
Fig. 10. Spatial comparisons of (left) Oklahoma Mesonet FWI and (right) ALEXI $f_{AW}$ for (a)–(h) each of the eight composite periods.
FIG. 10. (Continued)
REFERENCES


