Improving a Penman-Monteith evapotranspiration model by incorporating soil moisture control on soil evaporation in semiarid areas

Liang Sun\textsuperscript{ab}, Shunlin Liang\textsuperscript{abc}, Wenping Yuan\textsuperscript{ab} & Zhongxin Chen\textsuperscript{d}

\textsuperscript{a} State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Institute of Remote Sensing Applications, Chinese Academy of Science, Beijing, China

\textsuperscript{b} College of Global Change and Earth System Science, Beijing Normal University, Beijing, China

\textsuperscript{c} Department of Geography, University of Maryland, College Park, MD, USA

\textsuperscript{d} Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, China

Accepted author version posted online: 11 Mar 2013. Published online: 10 Apr 2013.

To cite this article: Liang Sun, Shunlin Liang, Wenping Yuan & Zhongxin Chen (2013) Improving a Penman-Monteith evapotranspiration model by incorporating soil moisture control on soil evaporation in semiarid areas, International Journal of Digital Earth, 6:sup1, 134-156, DOI: 10.1080/17538947.2013.783635

To link to this article: http://dx.doi.org/10.1080/17538947.2013.783635

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or
Improving a Penman–Monteith evapotranspiration model by incorporating soil moisture control on soil evaporation in semiarid areas

Liang Sun\textsuperscript{a,b*}, Shunlin Liang\textsuperscript{a,b,c}, Wenping Yuan\textsuperscript{a,b} and Zhongxin Chen\textsuperscript{d}

\textsuperscript{a}State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Institute of Remote Sensing Applications, Chinese Academy of Science, Beijing, China; \textsuperscript{b}College of Global Change and Earth System Science, Beijing Normal University, Beijing, China; \textsuperscript{c}Department of Geography, University of Maryland, College Park, MD, USA; \textsuperscript{d}Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, China

(Received 23 September 2012; final version received 6 March 2013)

Penman–Monteith (PM) theory has been successfully applied to calculate land surface evapotranspiration (ET) for regional and global scales. However, soil surface resistance, related to soil moisture, is always difficult to determine over a large region, especially in arid or semiarid areas. In this study, we developed an ET estimation algorithm by incorporating soil moisture control, a soil moisture index (SMI) derived from the surface temperature and vegetation index space. We denoted this ET algorithm as the PM-SMI. The PM-SMI algorithm was compared with several other algorithms that calculated soil evaporation using relative humidity, and validated with Bowen ratio measurements at seven sites in the Southern Great Plain (SGP) that were covered by grassland and cropland with low vegetation cover, as well as at three eddy covariance sites from AmeriFlux covered by forest with high vegetation cover. The results show that in comparison with the other methods examined, the PM-SMI algorithm significantly improved the daily ET estimates at SGP sites with a root mean square error (RMSE) of 0.91 mm/d, bias of 0.33 mm/d, and \( R^2 \) of 0.77. For three forest sites, the PM-SMI ET estimates are closer to the ET measurements during the non-growing season when compared with the other three algorithms. At all the 10 validation sites, the PM-SMI algorithm performed the best. PM-SMI 8-day ET estimates were also compared with MODIS 8-day ET products (MOD16A2), and the latter showed negligible bias at SGP sites. In contrast, most of the PM-SMI 8-day ET estimates are around the 1:1 line.

Keywords: ET; Penman–Monteith; soil moisture; MODIS

1. Introduction

Evapotranspiration (ET), which includes soil evaporation and vegetation transpiration, is a major component of the global water cycle, and plays a central role in climate and meteorology, plant community dynamics, and carbon and nutrient biogeochemistry (Vörösmarty, Federer, and Schloss 1998). ET returns 60% of the precipitation to the atmosphere (Oki and Kanae 2006) and imposes an important constraint on the water supply. In addition, more than half of the absorbed solar energy is utilized for ET (Trenberth, Fasullo, and Kiehl 2009). Accordingly, accurate
estimation of ET is extremely valuable for better irrigation management, especially in semiarid areas.

Remote sensing has been recognized as the most feasible means to provide a spatial distribution of land surface ET, as it can be used to estimate the land surface parameters required for estimation of ET, such as Normalized Difference Vegetation Index (NDVI), Leaf Area Index (LAI), albedo, and land surface temperature ($T_s$). Remote sensing-based methods can be divided into four types: (1) empirical statistical methods (Bella, Rebella, and Paruelo 2000; Carlson, Capehart, and Gillies 1995; Wang and Liang 2008); (2) residual methods based on the residual of surface energy balance through thermal infrared data (Anderson et al. 1997; Bastiaanssen et al. 1998; Norman, Kustas, and Humes 1995; Su 2002); (3) triangle/trapezoid methods linked with vegetation index and $T_s$ (Jiang and Islam 1999; Jiang and Islam 2001; Moran et al. 1994; Stisen et al. 2008; Tang, Li, and Tang 2010; Wang, Li, and Cribb 2006; Zhang et al. 2005); and (4) Penman–Monteith or Priestley–Taylor methods (Cleugh et al. 2007; Fisher, Tu, and Baldocchi 2008; Leuning et al. 2008; Mu et al. 2007a; Yuan et al. 2010; Zhang et al. 2010).

The empirical methods relate ET to vegetation parameters and key environmental control factors. These methods are naturally empirical, but their applications in other areas are limited because these methods do not provide an explicit formula to follow (Lu and Zhuang 2010). Some methods have proposed a universal empirical method that is suitable for different land cover (Wang and Liang 2008; Wang et al. 2010). However, experiments show that the composition of tree species strongly influence ET. Land cover dependent coefficients may be a better choice than a universal formulation (Wang and Dickinson 2012).

For the residual method, surface temperature ($T_s$), air temperature ($T_a$), and aerodynamic resistance ($r_a$) are first critically used to calculate sensible heat flux. Land surface temperature retrieved from thermal infrared data is strongly affected by clouds, resulting in discontinuous valid data with an error of about 1 K (Wan et al. 2002). Air temperature may have an unacceptable error depending on the elevation and number of distributed stations, if the ground-station measured values are interpolated spatially (Hamann and Wang 2005; Stahl et al. 2006). The difference between aerodynamic temperature and surface temperature results in the introduction of KB-1, which is a function of the structural characteristics of the vegetation, level of water stress, and climatic conditions (Lhomme et al. 1997), and is difficult to estimate precisely, especially on a regional or global scale. In addition, remote sensing can only directly capture instantaneous information, which may not meet the requirements of all applications (Li et al. 2009). The residual method may not be suitable for the estimation of period-averaged ET because of the difficulties in using 8- or 16-day composites of once-daily measurements of radiative surface temperature (Cleugh et al. 2007).

Although triangle/trapezoid methods are simple, and require only $T_s$ and the vegetation index, they have the following potential limitations: (1) determination of the dry and wet edges requires a certain degree of subjectivity; (2) a large number of pixels over a flat area with a wide range of soil wetness and fractional vegetation cover ($F_v$) are required to ensure that dry and wet limits exist in the triangle/trapezoid space (Sun et al. 2012). Zhang et al. (2005, 2008) determined the true dry edge based on an energy balance equation to overcome the limitations mentioned above. However, the estimated ET is still an instantaneous value.
The main challenge in the practical application of the Penman–Monteith or Priestley–Taylor equation is to determine the surface resistance, which is related to LAI, soil moisture status, leaf physiological attributes, and soil attributes. Based on some definitions and assumptions, Kelliher et al. (1995) derived an expression for surface resistance that considered the effects of soil and canopy. Leuning et al. (2008) introduced a factor in the surface resistance model that indicates the ratio of the actual ET to the potential ET. Cleugh et al. (2007) adopted the Penman–Monteith equation to estimate the 16-day ET based on a canopy resistance sub-model driven by MODIS-derived vegetation data and daily surface meteorological inputs including incoming solar radiation, surface air temperature, and vapor pressure deficit (VPD). Mu et al. (2007a) further modified the model developed by Cleugh et al. to estimate the global ET, and its partitioning into soil evaporation and vegetation transpiration. In their soil evaporation sub-model, soil resistance is controlled by air relative humidity, assuming that it reflects soil moisture status (Fisher, Tu, and Baldocchi 2008). Based on the model described by Mu et al. (2007a), Zhang et al. (2010) developed a model to estimate ET using remotely sensed NDVI data. Yuan et al. (2010) modified Mu et al.’s (2007a) model by adding the constraint of air temperature to stomatal conductance and calculating the vegetation cover fraction using LAI instead of EVI. Vinukollu et al. (2011) pointed out that Mu et al.’s (2007a) model underestimated ET for grassland, cropland, and woody savanna. Mu et al. (2011) improved Mu et al.’s (2007a) model by adding a nighttime ET component and wet soil surface component to enhance the ET estimates.

Soil resistance is largely controlled by soil moisture. However, directly monitoring large-scale soil moisture is always a challenge for remote sensing. Passive microwave sensors, such as TRMM and AMSR-E, have 1–1.5 passes per day but a coarse spatial resolution (25 km) that can only be used to estimate global or large-area soil moisture. Conversely, active microwave sensors (SAR: Synthetic Aperture Radar) have higher spatial resolution (10–30 m), but low repeat intervals (once in 16–25 days). Optical–thermal infrared sensors, such as MODIS and AVHRR/NOAA, have moderate resolution (250 m to 1 km) and one or two overpasses per day, which is suitable for regional- or large-scale applications. In addition, the triangle/trapezoid method formed by land surface temperature and vegetation index data is considered as a simple and suitable method to monitor soil moisture, and has been applied by many researchers (Gao et al. 2011; Mallick, Bhattacharya, and Patel 2009; Sandholt, Rasmussen, and Andersen 2002; Wang et al. 2004). However, this method has the same disadvantages associated with the determination of the dry edge, Zhang et al. (2008) developed the theoretical dry edge based on the energy balance equation. Then, Sun et al. (2012) showed that the theoretical dry edge could monitor soil moisture better than the traditional dry edge, which is fitted by scatterplots.

Mu et al. (2007b) have pointed out that the use of VPD alone is not enough to indicate the effects of the environmental water stress on plant growth for the arid and semiarid areas. For the similar reason, the VPD alone should not be enough to represent the water stress on ET especially in the arid and semiarid regions. In this study, we incorporated a soil moisture index (SMI), extracted from $T_r–F_r$ space, into the Penman–Monteith equation to estimate ET over semiarid areas, which is abbreviated as the PM-SMI method. We then compared this method with those developed by Mu et al. (2007a), Mu et al. (2011), and Yuan et al. (2010), which are
abbreviated as PM-Mu2007, PM-Yuan, and PM-Mu2011 below, in which soil evaporation is controlled by relative humidity.

2. ET algorithms

Soil evaporation in PM-Mu2007, PM-Mu2011, and PM-Yuan is limited by relative humidity, assuming that VPD (saturation vapor pressure deficit) can be used as an indicator of environmental water stress (Fisher, Tu, and Baldocchi 2008; Granger and Gray 1989). The PM-SMI algorithm is revised from PM-Mu2007, and we used SMI instead of the relative humidity, as soil moisture control.

2.1. PM-Mu2007 algorithm

Net radiation \( (R_n) \) is linearly partitioned between the canopy and the soil surface using \( F_v \):

\[
\begin{align*}
R_{nc} &= F_v \cdot R_n \\
R_{ns} &= (1 - F_v) \cdot R_n
\end{align*}
\]

where \( R_{nc} \) and \( R_{ns} \) are the total net incoming radiation \( (R_n) \) partitioned to the canopy and soil, respectively. \( F_v \) is estimated as follows:

\[
F_v = \frac{\text{NDVI} - \text{NDVI}_\text{min}}{\text{NDVI}_\text{max} - \text{NDVI}_\text{min}}
\]

where NDVI\text{min} and NDVI\text{max} are the signals from bare soil and dense green vegetation, which are set as seasonally and geographically invariant constants of 0.05 and 0.95, respectively (Zhang et al. 2009). The total latent heat flux is the sum of vegetation transpiration and soil evaporation:

\[
\lambda E = \lambda E_c + \lambda E_s
\]

where \( \lambda E \) is the latent heat flux, \( \lambda E_c \) is the vegetation transpiration \( (\text{W/m}^2) \), and \( \lambda E_s \) is the soil evaporation \( (\text{W/m}^2) \). Finally, ET (mm) is calculated as:

\[
\text{ET} = \frac{\lambda E}{\lambda}
\]

where \( \lambda \) \((2.43 \times 10^6 \text{ J/kg})\) is the latent heat of vaporization.

2.1.1. Vegetation transpiration

Vegetation transpiration is calculated as:

\[
\lambda E_c = \frac{\Delta R_{nc} + F_v \rho C_p \text{VPD}/r_a}{\Delta + \gamma(1 + r_c/r_a)}
\]

where \( \Delta \) (Pa/K) is the slope of the curve relating saturated water vapor pressure to temperature, \( \rho \) (kg/m\(^3\)) is the air density, \( C_p \) (J/kg/K) is the specific heat capacity of air, \( \gamma \) (Pa/K) is the psychrometric constant, VPD (Pa) is the saturation vapor pressure deficit \( (\text{VPD} = e_{sat} - e_0) \), where \( e_{sat} \) is the saturation vapor pressure, \( e_0 \) is the actual
vapor pressure), $r_a$ (s/m) is the aerodynamic resistance, and $r_c$ (s/m) is the canopy resistance, which is the reciprocal of canopy conductance $g_c$ ($g_c = 1/r_c$). LAI is used as a scalar to convert the stomatal conductance ($g_s$) calculated at the leaf level to a canopy conductance, and $g_s$ is estimated as described by Mu et al. (2007a), assuming that it is mainly affected by VPD and air temperature:

$$g_s = C_L \times m(T_{\text{min}}) \times m(\text{VPD})$$
$$g_c = g_s \times \text{LAI}$$

where $m(T_{\text{min}})$ is a temperature stress factor, $m(\text{VPD})$ is a water/moisture stress factor, $C_L$ is the mean potential stomatal conductance per unit leaf area, and $m(\text{VPD})$ and $m(T_{\text{min}})$ are calculated as Mu et al. (2007a):

$$m(\text{VPD}) = \begin{cases} 1.0 & \text{VPD} \leq \text{VPD}_{\text{open}} \\ \frac{\text{VPD}_{\text{close}} - \text{VPD}}{\text{VPD}_{\text{close}} - \text{VPD}_{\text{open}}} & \text{VPD}_{\text{open}} < \text{VPD} < \text{VPD}_{\text{close}} \\ 0.1 & \text{VPD}_{\text{close}} \leq \text{VPD} \end{cases}$$

$$m(T_{\text{min}}) = \begin{cases} 1.0 & T_{\text{min}} \geq T_{\text{min,open}} \\ \frac{T_{\text{min}} - T_{\text{min,close}}}{T_{\text{min,open}} - T_{\text{min,close}}} & T_{\text{min,close}} < T_{\text{min}} < T_{\text{min,open}} \\ 0.1 & T_{\text{min}} \leq T_{\text{min,close}} \end{cases}$$

where $\text{VPD}_{\text{open}}$ and $T_{\text{min,open}}$ are the biome-specific critical value of VPD and minimum air temperature ($T_{\text{min}}$) at which the canopy stomata are completely open, respectively; $\text{VPD}_{\text{close}}$ and $T_{\text{min,close}}$ are the biome-specific critical value of VPD and $T_{\text{min}}$ at which canopy stomata are completely closed, respectively. The values of $C_L$, $\text{VPD}_{\text{open}}$, $\text{VPD}_{\text{close}}$, $T_{\text{min,open}}$, and $T_{\text{min,close}}$ are listed in a Biome Properties Look-Up Table (Mu et al. 2007a). Aerodynamic resistance is calculated as follows:

$$r_a = \frac{r_h \times r_r}{r_h + r_r}$$

$$r_r = \frac{\rho C_p}{4\sigma T_h^3}$$

where $r_r$ (s/m) is the resistance to radiative heat transfer and $r_h$ (s/m) is the resistance to convective heat transfer and assumed to be equal to the boundary layer resistance (Thornton 1998), which is set as 107 s/m (Mu et al. 2007a).

### 2.1.2. Soil evaporation

Relative humidity is used as a constraint for potential soil evaporation to calculate actual soil evaporation:

$$\dot{E}_s = \left(\frac{RH}{100}\right)^{\text{VPD}/100} \times \frac{\Delta(R_{ns} - G) + (1 - F_v)\rho C_p \text{VPD}/r_a}{\Delta + \gamma(1 + r_{tot}/r_a)}$$

The total aerodynamic resistance to vapor transport ($r_{tot}$) is the sum of surface resistance ($r_s$) and the aerodynamic resistance, which is the same value as $r_h$ (107 m/s) (Mu et al. 2007a; Thornton 1998).
2.2. PM-SMI algorithm

In the PM-SMI algorithm, vegetation transpiration is calculated the same way as PM-Mu2007. Soil evaporation is estimated by incorporating soil moisture control as follows:

\[
\lambda E_s = \frac{\Delta(R_{ns} - G) + (1 - F_v) \rho C_p VPD/r_a}{\Delta + \gamma(1 + r_s/r_a)}
\]  

(12)

where \(G\) is the soil heat flux and \(r_s\) (s/m) is the soil resistance, which is estimated from the surface soil water content (assumed to be representative of the 0–5 cm layer) with an exponential equation (Kustas, Zhan, and Schmugge 1998; Sellers, Heiser, and Hall 1992),

\[
r_s = \exp[a - b(W/W_s)]
\]  

(13)

where \(W\) is the actual surface soil water content, \(W_s\) is the water content at saturation, which is a function of the soil texture, \(a\) and \(b\) are the experimentally derived parameters, adopted as 8.4 and 5.9, respectively, here (Kustas, Zhan, and Schmugge 1998). In addition, \(W/W_s\) could also be substituted by a relative soil moisture index (SMI) in Equation (13), expressed as Equation (14):

\[
r_s = \exp[a - b \cdot \text{SMI}]
\]  

(14)

\(T_s - F_v\) triangle is adopted widely to estimate soil moisture. However, the dry edge observed in the remote sensing data is characterized by lower temperatures compared with the theoretical dry edge, and thus, the observed dry edge determined by scatter plots is usually lower than the theoretical dry edge (Stisen et al. 2008). In this study, SMI is calculated from an improved \(T_s - F_v\) triangle space (Sun et al. 2012), in which the dry edge is determined by the energy balance equation. Sun et al. (2012) demonstrated that the new dry edge set a standard for different days and that the drought index calculated from the improved \(T_s - F_v\) space could monitor soil moisture conditions better than the traditional \(T_s - F_v\) space could. The details could be found in Appendix 1. Soil heat flux (\(G\)) is calculated as:

\[
G = R_n(1 - F_v) \ast k
\]  

(15)

where \(k\) is the ratio between \(G\) and \(R_n\) for bare soil. Halliwell and Rouse (1987) reported a daily \(k\) value of 0.16–0.18 at Churchill, Manitoba, while Rouse (1984) found a daily \(k\) value of 0.16–0.24 in Hudson Bay. Therefore, we set the daily \(k\) value as 0.18 (Jacobsen 1999).

Figure 1 shows the logic behind the PM-SMI ET algorithm for estimating daily ET.

2.3. PM-Mu2011 algorithm

Mu et al. (2011) added two items to improve ET estimation, evaporation from the wet canopy surface (\(\lambda E_{\text{wet}}\)) and nighttime ET. The total \(\lambda E\) is expressed as:

\[
\lambda E = \lambda E_{\text{wet}} + \lambda E_c + \lambda E_s
\]  

(16)

The estimation of \(r_c\), \(r_a\), and \(r_{\text{tot}}\) are also revised. The details are provided in Mu et al. (2011).
2.4. PM-Yuan algorithm

PM-Yuan adopted the same equations as PM-Mu2007, except for \( m(T_{\text{min}}) \), which follows the equation detailed by Fisher, Tu, and Baldocchi (2008):

\[
m(T_{\text{min}}) = \exp \left( - \left( \frac{T_a - T_{\text{opt}}}{T_{\text{opt}}} \right)^2 \right)
\]

where \( T_{\text{opt}} (^{\circ}\text{C}) \) is the optimum air temperature for photosynthesis, which is set as 25°C (Yuan et al. 2010). In this algorithm, biome property parameters such as \( C_L \), \( \text{VPD}_{\text{open}} \), \( \text{VPD}_{\text{close}} \), and \( T_{\text{opt}} \) are invariant across the various land cover types.

3. Data

To compare the performance of ET algorithms applied to different vegetation cover conditions, we used two data-sets. One data-set included seven sites in the Southern Great Plain (SGP) collected from 2004. The SGP is located in the middle of the USA, and primarily covered by grassland and cropland with low vegetation cover. The other data-set included three sites of the AmeriFlux network sampled in 2005. These sites are located in eastern USA, and primarily covered by forest with high vegetation cover.

3.1. SGP

The SGP site was the first field measurement site established by the Atmospheric Radiation Measurement Program. More than 30 instrument clusters have been placed...
around the SGP site, at the Central Facility and at Boundary, Extended, and Intermediate Facilities. The locations for the instruments were selected so that the measurements reflect conditions over the typical distribution of land uses within the site. The SGP covers most of Oklahoma and the southern part of Kansas, which is a heterogeneous land cover area characterized by mixed farming, interrupted forest, and tall and short grass. Major soil types are silt loam, loamy sand, and loam. The region produces much of the nation’s grain and fiber, including more than 60% of the wheat and 36% of the cotton. Water supplied as irrigation in the region amounts to billions of cubic meters ($2 \times 10^{10} \text{ m}^3$ in 1990) annually and is governed by the determination of ET in the region. An area including seven study sites (Table 1), was selected in this study, and the Energy Balance Bowen Ratio (EBBR) system was placed to obtain ET measurements.

EBBR produces 30-minute estimates of the vertical fluxes of sensible and latent heat at the local surface. Flux estimates are calculated from the observations of net radiation, soil surface heat flux, and the vertical gradients of temperature and relative humidity. Meteorological data collected by the EBBR are used to calculate the bulk aerodynamic fluxes, which are used in the Bulk Aerodynamic Technique EBBR value-added product to replace sunrise and sunset spikes in the flux data. A unique aspect of the system is the automatic exchange mechanism which helps to reduce errors from instrument offset drift.

$\lambda E$ is estimated as a function of the Bowen ratio $\beta$:

$$\lambda E = \frac{R_n - G}{1 + \beta} \quad (18)$$

where $\beta$ is estimated from the vertical gradients of temperature ($\Delta T_a$) and specific humidity ($\Delta e_o$).

3.2. AmeriFlux sites

An area in the eastern USA including three flux towers, Duke Forest Hardwoods, North Carolina Loblolly Pine, and Walker Branch, were selected. Detailed information regarding these sites is listed in Table 1.

The ET at the AmeriFlux sites was measured by the eddy covariance (EC) method (http://public.ornl.gov/ameriflux/data-get.cfm). The eddy covariance method is accepted as the best method for direct measurement of heat fluxes and is widely used.

Table 1. Location, soil type and land cover of study sites.

<table>
<thead>
<tr>
<th>Site</th>
<th>Lat</th>
<th>Lon</th>
<th>Land cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>E7_Elk_Falls</td>
<td>37.383</td>
<td>96.18 W</td>
<td>Grass</td>
</tr>
<tr>
<td>E8_Coldwater</td>
<td>37.333</td>
<td>99.309 W</td>
<td>Grass</td>
</tr>
<tr>
<td>E9_Ashton</td>
<td>37.133</td>
<td>97.266 W</td>
<td>Grass</td>
</tr>
<tr>
<td>E13_Lamont</td>
<td>36.605</td>
<td>97.485 W</td>
<td>Crop</td>
</tr>
<tr>
<td>E15_Ringwood</td>
<td>36.431</td>
<td>98.284 W</td>
<td>Grass</td>
</tr>
<tr>
<td>E20_Meeker</td>
<td>35.564</td>
<td>96.988 W</td>
<td>Grass</td>
</tr>
<tr>
<td>E27_Earlboro</td>
<td>35.269</td>
<td>96.740 W</td>
<td>Grass</td>
</tr>
<tr>
<td>Duke Forest Hardwoods</td>
<td>35.974</td>
<td>79.100 W</td>
<td>Mixed forest</td>
</tr>
<tr>
<td>North Carolina Loblolly Pine</td>
<td>35.803</td>
<td>76.668 W</td>
<td>Mixed forest</td>
</tr>
<tr>
<td>Walker Branch</td>
<td>35.959</td>
<td>84.287 W</td>
<td>Deciduous broad-leaf forest</td>
</tr>
</tbody>
</table>
used in global measurement experiments (Baldocchi et al. 2001). Unfortunately, it suffers from an energy imbalance problem, defining the energy closure ratio (ECR) as:

$$ECR = \frac{\lambda E_{ec} + H_{ec}}{R_n - G}$$  \hspace{1cm} (19)

where $\lambda E_{ec}$ and $H_{ec}$ are the original latent heat flux ($\lambda E$) and sensible heat flux ($H$) measured by the eddy covariance method. The ECR of the three sites used in this study is listed in Table 3. Wilson et al. (2002) argued that ECR is about 0.8 for the global FluxNet measurements, which results in a substantial underestimation of $\lambda E$ and $H$. However, ECR should be equal to unity according to the energy balance law. Although several reasons are proposed to explain the energy imbalance issue, the mechanism for this imbalance remains unclear (Gao 2005; Wilson et al. 2002; Yang et al. 2004). We applied the correction method, assuming that Bowen ratio is correctly measured, proposed by Twine et al. (2000) to obtain the corrected $\lambda E$ measurements for validation:

$$\lambda E = \frac{\lambda E_{ec}}{ECR}$$  \hspace{1cm} (20)

### 3.3. Remote sensing data

Four MODIS products, including the 16-day vegetation indices at 1-km resolution (MOD13A2), daily $T_s$ products at 1-km resolution (MOD11A1), 8-day albedo products at 1-km resolution (MCD43B3), and 8-day ET products (MOD16A2) at 1-km resolution were used. These data were acquired through the EOS data gateway (https://wist.echo.nasa.gov), except for the MOD16A2 which were downloaded from frp://ftp.ntsg.umt.edu. Two algorithms were used to retrieve $T_s$ from the MODIS thermal and middle infrared spectral regions: the generalized split window algorithm (Wan and Dozier 1996) and the MODIS day/night land surface temperature algorithm (Wan and Li 1997). The accuracy of the MODIS $T_s$ product is better than 1 K (Wan et al. 2002; Wan et al. 2004). The 16-day NDVI extracted from MOD13A2 was time series smoothed by the locally adjusted cubic spline capping (LACC) method (Chen, Feng, and Chen 2006), which can produce flexible and mathematically smooth capping curves and fit rapid seasonal changes, for the purpose of identifying the atmosphere-contaminated data points and their replacements through temporal interpolation. The data were then linear interpolated to give the daily value. The albedo product consists of black sky and white sky albedo. To derive the overall surface albedo, we followed the method described by Su et al. (2007) by averaging the above two estimates. The MOD16 ET data-sets were estimated using the ET algorithm described by Mu et al. (2011). The MOD16 ET algorithm uses NASA’s MERRA GMAO (GEOS-5) daily meteorological reanalysis data, collection 4 MOD12Q1, collection 5 MCD43B2/B3 and MOD15A2 as input. The 8-day ET is the sum of ET during these 8-day time periods. All these products described above were converted to UTM projection using the Modis Reprojection Tool.

Two study regions were selected for this study. One is in the middle of the USA, containing seven SGP sites. The selected window size for the remote sensing image was 562 × 526 km, extended in longitude from 95.3°W to 99.5°W and in latitude from 34.5°N to 38.5°N, which is the spatial subset from tile H10V05 of 2004 and...
covered most of the SGP. Another is in eastern USA, containing three AmeriFlux sites, for which the selected window size of the remote sensing image was 884 × 592 km. This area extended in longitude from 75.1°W to 85.4°W and in latitude from 34.5°N to 39.8°N, which is the spatial subset from tile H11V05 of 2005.

3.4. Methodology

In this study, although SMI distribution could be estimated based on $T_s-F_v$ space, we use in situ meteorological data such as $T_a$, RH, and $R_n$. Therefore, only in situ ET is estimated using measured meteorological data and $3 \times 3$ 1-km pixels of remote sensing data surrounding each site. Below is the description of each step.

1. Deriving the dry edge using instantaneous value of air temperature, vapor pressure ($e_0$), solar radiation ($S_0$), $T_s$, albedo, and NDVI;
2. Calculating the SMI distribution based on the improved $T_s-F_v$ space and extracting $3 \times 3$ pixels value of SMI, NDVI, and LAI images surrounding each study site;
3. Estimating daily ET using in situ daily value of air temperature, relative humidity (RH), $R_n$, and each pixel value of SMI, NDVI, and LAI;
4. The PM-SMI ET average over the $3 \times 3$ pixels was validated with tower ET observations;
5. Compared to the other three ET algorithms with the same input measurement data.

4. Results and discussion

Table 2 shows the mean NDVI, mean LAI, max NDVI, max LAI, and mean soil water content for each site. The vegetation cover is low for SGP sites, which were

<table>
<thead>
<tr>
<th>Site</th>
<th>Mean NDVI</th>
<th>Mean LAI</th>
<th>Max NDVI</th>
<th>Max LAI</th>
<th>Mean soil water content (m$^3$/m$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E7</td>
<td>0.54</td>
<td>1.4</td>
<td>0.75</td>
<td>3.1</td>
<td>0.37</td>
</tr>
<tr>
<td>E8</td>
<td>0.45</td>
<td>0.8</td>
<td>0.52</td>
<td>1.5</td>
<td>0.14</td>
</tr>
<tr>
<td>E9</td>
<td>0.49</td>
<td>0.6</td>
<td>0.62</td>
<td>1.1</td>
<td>0.32</td>
</tr>
<tr>
<td>E13</td>
<td>0.53</td>
<td>0.9</td>
<td>0.67</td>
<td>2.1</td>
<td>0.32</td>
</tr>
<tr>
<td>E15</td>
<td>0.48</td>
<td>0.64</td>
<td>0.60</td>
<td>1.0</td>
<td>0.11</td>
</tr>
<tr>
<td>E20</td>
<td>0.55</td>
<td>1.2</td>
<td>0.73</td>
<td>2.6</td>
<td>0.32</td>
</tr>
<tr>
<td>E27</td>
<td>0.56</td>
<td>1.1</td>
<td>0.74</td>
<td>2.3</td>
<td>0.18</td>
</tr>
<tr>
<td>Duke FH</td>
<td>0.67</td>
<td>2.42</td>
<td>0.85</td>
<td>5.9</td>
<td>0.33</td>
</tr>
<tr>
<td>North CLP</td>
<td>0.63</td>
<td>2.27</td>
<td>0.86</td>
<td>6.3</td>
<td>0.33</td>
</tr>
<tr>
<td>Walker B</td>
<td>0.68</td>
<td>2.44</td>
<td>0.88</td>
<td>6</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table 2. The mean NDVI, mean LAI, max NDVI, max LAI and mean soil water content for each site. The mean and max values of NDVI and LAI were obtained from the MODIS NDVI and LAI products. The mean soil water content was obtained from surface soil moisture measurements taken at 5 cm. The data from SGP sites was collected from 2004 and the data from AmeriFlux sites was collected from 2005.
covered by grassland and cropland. There is a mix of good density improved pasture in site E7; therefore, the value of NDVI and LAI in this site is relatively high (max NDVI = 0.75, max LAI = 3.1). The vegetation cover is high for the three forest sites. The soil type of E8 and E15 is sandy; therefore, the mean soil water content is lower than 0.15 m$^3$/m$^3$.

Since $T_s$ is affected by clouds, the results of the PM-SMI algorithm were only available for less than half of the year. As shown in Table 3, the PM-SMI algorithm significantly enhanced the ET estimates for all of the SGP sites, which were covered by grassland and cropland, while the other three ET algorithms had relatively large errors. For example, at site E9, the root mean square error (RMSE) of PM-Mu2007, PM-Mu2011, PM-Yuan, and PM-SMI is 1.69, 1.64, 1.67, and 0.39 mm, respectively, while the bias is −1.49, −1.34, −1.45, and −0.03 mm, respectively, and the $R^2$ is 0.94, 0.87, 0.95, and 0.95, respectively. Although the $R^2$ values of the PM-SMI ET estimates are close to those of the PM-Mu2007 and PM-Yuan ET estimates, the RMSE and bias were greatly reduced, indicating that PM-SMI ET estimates capture the magnitude of ET measurements better than the other three algorithms. For all the seven SGP sites, the RMSE and bias of PM-Mu2011 (1.37 mm and −0.93 mm) are lower than that of PM-Mu2007 (1.4 mm and −1.06 mm). However, the $R^2$ of PM-Mu2011 (0.64) is lower than the PM-Mu2007 (0.69). The RMSE, bias, and $R^2$ of PM-Yuan fall between those of PM-2007 and PM-2011. Although PM-Mu2011 improved the ET estimates when compared with PM-Mu2007 at grasslands and croplands as reported by Mu et al. (2011), results of this study indicate that PM-Mu2011 still underestimates ET.

Among the forest sites, in Duke Forest Hardwoods (Mixed forest), North Carolina Loblolly Pine (Mixed forest) and Walker Branch (Deciduous Broad-leaf forest), there were no large differences among ET estimates of PM-Mu2007, PM-Yuan, and PM-SMI. Although PM-Mu2011 showed the lowest RMSE and bias and the highest $R^2$ in Walker Branch, PM-Mu2011 showed the lowest correlation and highest RMSE and bias across all the three forest sites.

Across all 10 sites, the PM-SMI algorithm also had the best performance, with an RMSE of 0.88 mm, bias of 0.24 mm and $R^2$ of 0.78. PM-Yuan ET estimates had a slightly lower RMSE and bias (1.28 mm, −0.78 mm), and higher $R^2$ (0.66) than PM-Mu2007 (1.31 mm, −0.82 mm and 0.66) and PM-Mu2011 (1.33 mm, −0.8 mm and 0.64). PM-Mu2011 had a lower bias than PM-Mu2007, but the RMSE was higher and the $R^2$ was lower. Mu et al. (2011) also indicated that PM-Mu2011 had a reduced bias relative to PM-Mu2007, although it had a slightly increased RMSE and decreased $R^2$.

Figure 2 shows the scatterplots between the observed and the estimated ET with four algorithms. PM-Mu2007, PM-Mu2011, and PM-Yuan significantly underestimate ET and PM-SMI performs the best among the four ET algorithms. Figure 3 shows comparisons of the ET measurements and the ET estimates with four algorithms. The PM-SMI ET estimates match well across the annual variation for all SGP sites. PM-Mu2007, PM-Mu2011, and PM-Yuan algorithms obviously underestimated ET across SGP sites, where soil evaporation plays a dominant role. However, for E7, E20, and E27, which have relatively good vegetation condition (Table 2), these three algorithms’ ET estimates are closer to the ET measurements, with a lower bias (Table 3). No significant difference was observed among the four algorithms’ ET estimates for the three forest sites during the growing season.
Table 3. Available days, root mean square error (RMSE: mm/day), absolute bias (bias: mm/day), and square of correlation coefficient ($R^2$) of four ET algorithms for each site. (1) PM-Mu2007; (2) PM-Mu2011; (3) PM-Yuan; and (4) PM-SMI.

<table>
<thead>
<tr>
<th>Site</th>
<th>ECR</th>
<th>Available days</th>
<th>RMSE1</th>
<th>RMSE2</th>
<th>RMSE3</th>
<th>RMSE4</th>
<th>Bias1</th>
<th>Bias2</th>
<th>Bias3</th>
<th>Bias4</th>
<th>$R^2_1$</th>
<th>$R^2_2$</th>
<th>$R^2_3$</th>
<th>$R^2_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>E7</td>
<td>–</td>
<td>148</td>
<td>1.30</td>
<td>1.08</td>
<td>1.22</td>
<td>0.74</td>
<td>−1.01</td>
<td>−0.71</td>
<td>−0.95</td>
<td>−0.05</td>
<td>0.79</td>
<td>0.81</td>
<td>0.81</td>
<td>0.86</td>
</tr>
<tr>
<td>E8</td>
<td>–</td>
<td>161</td>
<td>1.48</td>
<td>1.49</td>
<td>1.44</td>
<td>1.05</td>
<td>−1.14</td>
<td>−1.20</td>
<td>−1.11</td>
<td>0.43</td>
<td>0.44</td>
<td>0.55</td>
<td>0.46</td>
<td>0.58</td>
</tr>
<tr>
<td>E9</td>
<td>–</td>
<td>162</td>
<td>1.69</td>
<td>1.64</td>
<td>1.67</td>
<td>0.39</td>
<td>−1.49</td>
<td>−1.34</td>
<td>−1.45</td>
<td>−0.03</td>
<td>0.94</td>
<td>0.87</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>E13</td>
<td>–</td>
<td>153</td>
<td>1.89</td>
<td>1.95</td>
<td>1.88</td>
<td>0.97</td>
<td>−1.35</td>
<td>−1.23</td>
<td>−1.29</td>
<td>0.18</td>
<td>0.60</td>
<td>0.40</td>
<td>0.54</td>
<td>0.75</td>
</tr>
<tr>
<td>E15</td>
<td>–</td>
<td>161</td>
<td>1.38</td>
<td>1.49</td>
<td>1.39</td>
<td>0.98</td>
<td>−1.05</td>
<td>−1.07</td>
<td>−1.04</td>
<td>0.57</td>
<td>0.73</td>
<td>0.69</td>
<td>0.74</td>
<td>0.76</td>
</tr>
<tr>
<td>E20</td>
<td>–</td>
<td>149</td>
<td>1.02</td>
<td>0.87</td>
<td>0.96</td>
<td>0.83</td>
<td>−0.76</td>
<td>−0.46</td>
<td>−0.71</td>
<td>0.34</td>
<td>0.82</td>
<td>0.81</td>
<td>0.84</td>
<td>0.85</td>
</tr>
<tr>
<td>E27</td>
<td>–</td>
<td>127</td>
<td>1.22</td>
<td>1.01</td>
<td>1.18</td>
<td>0.90</td>
<td>−0.91</td>
<td>−0.67</td>
<td>−0.85</td>
<td>0.26</td>
<td>0.83</td>
<td>0.86</td>
<td>0.84</td>
<td>0.81</td>
</tr>
<tr>
<td>SGP</td>
<td>–</td>
<td></td>
<td>1.40</td>
<td>1.37</td>
<td>1.37</td>
<td>0.91</td>
<td>−1.06</td>
<td>−0.93</td>
<td>−1.02</td>
<td>0.33</td>
<td>0.69</td>
<td>0.64</td>
<td>0.69</td>
<td>0.77</td>
</tr>
<tr>
<td>Duke FH</td>
<td>0.81</td>
<td>180</td>
<td>0.56</td>
<td>0.93</td>
<td>0.53</td>
<td>0.58</td>
<td>−0.05</td>
<td>−0.43</td>
<td>−0.02</td>
<td>0.09</td>
<td>0.93</td>
<td>0.84</td>
<td>0.93</td>
<td>0.91</td>
</tr>
<tr>
<td>North CLP</td>
<td>0.70</td>
<td>158</td>
<td>0.88</td>
<td>1.46</td>
<td>0.86</td>
<td>0.68</td>
<td>−0.67</td>
<td>−0.89</td>
<td>−0.66</td>
<td>−0.38</td>
<td>0.91</td>
<td>0.72</td>
<td>0.91</td>
<td>0.89</td>
</tr>
<tr>
<td>Walker B</td>
<td>0.64</td>
<td>98</td>
<td>1.31</td>
<td>0.82</td>
<td>1.30</td>
<td>1.42</td>
<td>0.76</td>
<td>0.37</td>
<td>0.78</td>
<td>0.98</td>
<td>0.74</td>
<td>0.80</td>
<td>0.73</td>
<td>0.68</td>
</tr>
<tr>
<td>Forest</td>
<td>–</td>
<td></td>
<td>0.99</td>
<td>1.20</td>
<td>0.97</td>
<td>0.96</td>
<td>−0.12</td>
<td>−0.45</td>
<td>−0.10</td>
<td>0.15</td>
<td>0.75</td>
<td>0.69</td>
<td>0.75</td>
<td>0.74</td>
</tr>
<tr>
<td>All</td>
<td>–</td>
<td></td>
<td>1.31</td>
<td>1.33</td>
<td>1.28</td>
<td>0.88</td>
<td>−0.82</td>
<td>−0.80</td>
<td>−0.78</td>
<td>0.24</td>
<td>0.66</td>
<td>0.64</td>
<td>0.66</td>
<td>0.78</td>
</tr>
</tbody>
</table>
However, the PM-SMI ET estimates are closer to the ET measurements in the non-growing season when compared with the other three algorithms.

Figure 4 shows comparisons of the 8-day ET measurements with the PM-SMI 8-day ET estimates and MODIS 8-day ET product. For the PM-SMI ET estimates, it was difficult to obtain continuous daily ET owing to the lack of continuous $T_s$. Therefore, the valid PM-SMI 8-day ET was combined for every 8-day period, during which there are more than six valid daily ET values. The MODIS 8-day ET was extracted directly from the MOD16A2 product. As shown in Figure 4, the MODIS 8-day ET was significantly lower than the ET measurements for the SGP sites, with $R^2$, RMSE, and bias values of 0.51, 10.58 mm, and −6.72 mm, respectively. In contrast, most of the PM-SMI 8-day ET estimates were around the 1:1 line. However, the PM-SMI 8-day ET estimates were slightly larger than the ET measurements, with a bias of 4.02 mm. The PM-SMI daily ET was estimated based on clear day conditions with no clouds, resulting in larger ET values when compared to those obtained for cloudy days. Therefore, it is reasonable that the PM-SMI 8-day ET estimates combined from the cloudless daily ET were slightly larger than the measurements. The $R^2$ and RMSE of the PM-SMI 8-day ET estimates were 0.68 and 7.82 mm, respectively.

However, relatively large differences between PM-SMI ET estimates and ET measurements still exist for a few sites, including E8, E13, E15, and Walker Branch. These discrepancies may be explained by the following:
(1) Estimation error of soil heat flux. Surface available energy used in the Penman–Monteith equation is estimated from net radiation minus soil heat flux. However, soil heat flux, which is correlated with soil properties, soil moisture, and vegetation condition, has high spatial variability. The ratio between $G$ and $R_n$ over bare soil ranges from 0.20 to 0.50 depending on the soil moisture (Idso, Aase, and Jackson 1975), and from 0.05 to 0.30 depending on the vegetation cover (Clothier et al. 1986; Kustas and Daughtry 1990). Therefore, it is a challenge to estimate soil heat flux. In our study,

Figure 3. The ET measurements (gray line) and ET estimates with PM-Mu2007 (cyan line), PM-Mu2011 (blue line), PM-Yuan (yellow line), and PM-SMI algorithms (red line) for 10 sites.
relative bias of $G$ ranged from $-30\%$ to $30\%$, which may result in some uncertainties associated with the ET estimates.

(2) Algorithm limitations. Biophysical parameters ($C_L$, $VPD_{\text{close}}$, and $VPD_{\text{open}}$) used in this study are the same as those used in the PM-Mu 2011 algorithm. However, Mu et al. (2007a), Mu et al. (2011), Yuan et al. (2010), and Zhang et al. (2010), all have recalibrated the biophysical parameters of Penman–Monteith ET algorithms based on their studies. The parameters such as $a$ and $b$, which are used to compute soil resistance ($r_s$) from SMI, are also the same.
as those used by Kustas, Zhan, and Schmugge (1998). However, these values are a function of soil texture, which is heterogeneous in our study region, and consists of silt loam, sandy loam, clay loam, and loamy sand. Unfortunately, ground measurement sites for different biome types or soil textures are not sufficient for analysis. Therefore, it is not possible to refit these parameters.

3) SMI. The SMI incorporated in the PM-SMI algorithm is estimated based on improved $T_x - F_v$ trapezoid space (Sun et al. 2012). The assumption of monitoring soil moisture based on $T_x - F_v$ space is that meteorological parameters and land surface attributes are homogeneous (Friedl and Davis 1994; Lambin and Ehrlich 1995; Zhang et al. 2008). However, Sun et al. (2012) showed that sites with sandy soil, such as E8 and E15, could result in an evident difference in heat property and relationships between $T_x$ and soil moisture when compared to other soil types, which was likely a major cause of large errors at E8 and E15.

4) Inaccuracy of ET measurements. Surface energy flux data from the SGP sites measured by Bowen ratio do not have energy balance non-closure problem. However, data from three forest sites, measured by the eddy covariance flux towers, have lack of energy balance closure. Although latent heat flux was corrected using the method of Twine et al. (2000), correcting error and reducing uncertainty in the ET measurements is still uncertain due to the closure error (Shuttleworth 2007). The large errors at the Walker Branch site may result from a low energy balance ratio, which is about 50% on most days.

5) Validation. We found that some research used one pixel containing the study site to validate (Yuan et al., 2010), some research used $3 \times 3$ pixels to validate (Mu et al. 2007a), and some research used footprint model to validate (Jia et al. 2012). To reduce the effect of heterogeneity of land surface, we used $3 \times 3$ pixels surrounding the study sites to validate the results. Actually, this is a simplified method while the area of these pixels could mostly represent the source area of flux measurements. The source area is always varying with the observation height, wind velocity and direction, atmospheric stability, and the roughness length of land surface. For the EC and EBBR measurements, which used in this study, the source area of them might either located in only one MODIS pixel or covered some parts of two to four pixels, with different relative contribution within the source area.

5. Conclusions
Penman–Monteith equation is a classical method used to calculate ET. However, surface resistance, which is related to soil moisture, is always difficult to obtain. In this study, we developed an ET algorithm by incorporating the SMI derived from an improved $T_x - F_v$ space. The PM-Mu2007 algorithm has been successfully applied on the global scale (Mu et al. 2007a; Mu et al. 2009), and was recently adjusted by Yuan et al. (2010) and Zhang et al. (2009). Mu et al. (2011) pointed out that PM-Mu2007 underestimates ET for cropland, grassland, and woody savanna, which was also confirmed by other studies (Vinukollu et al. 2011). Mu et al. (2011) improved this ET algorithm and enhanced the ET estimates most for woody savannas, grasslands, and croplands. Soil evaporation in these ET algorithms was controlled by VPD. However, the VPD alone is not enough to represent the water stress on ET especially in the arid
and semiarid regions (Mu et al. 2007b). Generally in the arid and semiarid areas, the vegetation cover fraction is low, the plant transpiration is low as well, and hence the soil evaporation is the largest part of the ET components.

In this study, we tried to improve Mu et al.’s 2007 ET algorithm by incorporating the SMI to account for the soil moisture constraint on soil evaporation to develop the PM-SMI ET algorithm. ET estimates derived from PM-SMI algorithm were compared with PM-Mu2007, PM-Mu2011, and PM-Yuan algorithm and validated by seven SGP sites that were covered by cropland and grassland with low vegetation cover, as well as three AmeriFlux sites covered by forest with high vegetation cover. The results showed that the PM-SMI algorithm significantly improved daily ET estimates at SGP sites with an RMSE of 0.91 mm, bias of 0.33 mm, and $R^2$ of 0.77. PM-Mu2011 reduced the RMSE and bias of ET estimates when compared with PM-Mu2007, but still underestimated ET. For the three forest sites, no large difference was observed among the ET estimates produced by these four algorithms during the growing season. However, the PM-SMI ET estimates are closer to the ET measurements in the non-growing season when compared with the other three algorithms. For all 10 validation sites, PM-SMI also had the best performance, followed by PM-Yuan.

PM-SMI 8-day ET estimates were also compared with MODIS 8-day ET product (MOD16A2). MODIS 8-day ET was significantly lower than the ET measurements for the SGP sites, with $R^2$, RMSE, and bias values of 0.51, 10.58 mm, and $-6.72$ mm, respectively. In contrast, most PM-SMI 8-day ET estimates were around the 1:1 line, with $R^2$, RMSE, and bias values of 0.68, 7.82 mm, and 4.02 mm, respectively.

ET distribution is necessary in many applications, and the method proposed in this manuscript would be a meaningful attempt to get the ET maps. Unfortunately, the input meteorological data, such as net radiation ($R_n$) was only collected from study sites. We could not get a good product of $R_n$ distribution of the study area. We will try to make some analysis about the ET distribution in the future. The application of PM-SMI is limited by land surface temperature, which is used to obtain the SMI, since surface temperature is largely affected by clouds and it is difficult to capture consistent and valid data. Passive microwave sensors, such as AMSR-E, which have all-weather surface sensing capability, have been used to retrieve soil water content (Bindlish et al. 2003; Njoku et al. 2003). However, the application of this technique is limited by its coarse spatial resolution (25 km). In the future, we plan to downscale the soil moisture product derived from AMSR-E to fine scales matching MODIS 1 km resolution.

Acknowledgements
This work was funded by the High-Tech Research and Development Program of China (No. 2009AA122100) and the National Science and Technology Ministry (2012BAH29B02). We thank Dr. Tongren Xu for providing the measurement data. SGP data of the Atmospheric Radiation Measurement (ARM) Project of the US Department of Energy is downloaded from http://www.arm.gov/. Eddy covariance flux tower sites are part of the AmeriFlux. We gratefully acknowledge the efforts of researchers at these sites. Research at the Duke Forest Hardwoods, North Carolina Loblolly Pine and Walker Branch sites was led by Ram Oren, Asko Noormets and Tilden Meyers.
References


Appendix A: Calculation of relative soil moisture based on $T_s - F_v$ space.

Relative soil moisture index (SMI) is calculated as Equation (A1) and determined from $T_s - F_v$ space as shown in Figure A1. In this $T_s - F_v$ space, the dry edge is determined based on the energy balance equation (Zhang et al. 2008). Sun et al. (2012) showed that the new dry edge could monitor soil moisture better than the traditional dry edge, which is fitted by scatterplots.

$$\text{SMI} = \frac{T_{\text{max}} - T_s}{T_{\text{max}} - T_{\text{min}}} = \frac{AC}{AB}$$  \hspace{1cm} \text{(A1)}

The dry edge was determined from energy balance formula:

$$R_n - G = H + LE$$  \hspace{1cm} \text{(A2)}

For the extreme dry condition, there is no $LE$, so Equation (A2) could be expressed as:

$$R_n - G = H$$  \hspace{1cm} \text{(A3)}

The sensible heat flux $H$ is estimated as:

$$H = \frac{\rho C_p (T_s - T_a)}{r_a}$$  \hspace{1cm} \text{(A4)}

Net radiation $R_n$ can be expressed as:

$$R_n = S_0(1 - \alpha) + \sigma v_a T_a^4 - \sigma e_a T_s^4$$  \hspace{1cm} \text{(A5)}

Figure A1. $T_s - F_v$ trapezoid space. The upper solid line is the dry edge, estimated from the energy balance formula, representing extreme water stressed conditions. The lower solid line is the wet edge, representing saturated soil water conditions.
where $S_0$ is the solar radiation reaching the earth’s surface, $\alpha$ is albedo, $\sigma$ is Stefan–Boltzmann constant, $\epsilon_s$ is the atmospheric emissivity under clear skies, calculated as follows (Brutsaert 1975):

$$
\epsilon_a = 1.24 \left( \frac{e_0}{T_a} \right)^{1/7}
$$

(A6)

$\epsilon_s$ is the emissivity of land surface, estimated by (Zhang et al. 2008):

$$
\epsilon_s = F_v \epsilon_v + (1 - F_v) \epsilon_{sv}
$$

(A7)

where $\epsilon_v$ and $\epsilon_{sv}$ are emissivity for vegetation (0.97) and soil (0.95), respectively. Instantaneous $G$ at midday can be calculated as follows (Su 2002):

$$
G = R_n \Gamma_v + \left[ \frac{\Gamma_s}{\Gamma_v} \right] \left( \frac{1}{\Gamma_v} F_v \right) \frac{1}{\Gamma_s}
$$

(A8)

where $\Gamma_v$ and $\Gamma_s$ are the ratios between $G$ and $R_n$ for bare soil and fully covered vegetation surface, respectively. The value of $\Gamma_s$ is taken as 0.315 (Kustas and Daughtry 1990), while 0.05 for $\Gamma_v$ (Monteith 1973). For dry bare soil, the vegetation cover $F_v = 0$, taking Equations (A4), (A5), and (A8) into Equation (A3), surface temperature is derived as follows:

$$
T_{sd} = 0.7 S_0 \frac{1 - \alpha_{sd}}{\alpha_{sd}} + \sigma e_s T_{sd}^4 + \frac{\rho C}{\tau_{vd}} T_{vd}^3
$$

(A9)

where the subscript $sd$ denotes dry bare soil. For full-cover vegetation with no available water, the vegetation cover $F_v = 1$, taking Equations (A4), (A5), and (A8) into Equation (A3), surface temperature is given by:

$$
T_{vd} = 0.95 S_0 \frac{1 - \alpha_{vd}}{\alpha_{vd}} + \sigma e_v T_{vd}^4 + \frac{\rho C}{\tau_{vd}} T_{vd}^3
$$

(A10)

where the subscript $vd$ denotes true dry full cover vegetation. $T_{sd}$ or $T_{vd}$ can be iteratively computed if the $\alpha_{sd}$, $\alpha_{vd}$, $T_{sd}$, $T_{vd}$, $\tau_{sd}$, $\tau_{vd}$, $e_0$, and $S_0$ are acquired.

We adopt an empirical method proposed by Zaksek and Schroeder-Homscheidt (2009) to estimate instantaneous $T_a$:

$$
T_a = T_v + 1.82 - 10.66 \cdot \cos z_v \cdot (1 - \text{NDVI}) \quad + 0.566 \cdot a_v - 3.72 \cdot (1 - z) \cdot (\cos s_v + \cos s - (\pi - s)/\pi)
$$

where $z_v$ is the solar zenith angle, $a_v$ is the solar azimuth angle, $z$ is the albedo, $s_v$ is the solar incidence angle, $s$ is the slope, and $\Delta h$ is the difference between the station elevation and the mean elevation within a 20 km vicinity. Assuming that the theoretically driest atmospheric conditions is close to the observed driest conditions, the pixel with the greatest air temperature around the upper-left corner in the trapezoid represents the driest bare soil, and the pixel with the greatest air temperature around the upper-right corner in the trapezoid represents the driest full-cover vegetation. The temperatures of these two pixels are represented as $T_{sd}$ and $T_{vd}$.

For $\alpha_{sd}$ and $\alpha_{vd}$ we select the maximum albedo from the bare soil and full cover vegetation, corresponding to soil and vegetation dry conditions, respectively.

Assuming that there is not much variation for $S_0$ on a regional scale (Mueller et al. 2004) for a clear sky, the averaged value of measured $S_0$ is adopted as the input. The location of the theoretical dry edge is not sensitive to $\epsilon_s$ (Sun et al. 2012); thus, the averaged value of $\epsilon_s$ is also adopted as the input.