ENVIROMENTAL RESEARCH LETTERS

Retrieving gap-free daily root zone soil moisture using surface flux equilibrium theory

Pushpendra Raghav and Mukesh Kumar

Department of Civil, Construction, and Environmental Engineering, University of Alabama, Tuscaloosa, AL, United States of America

* Author to whom any correspondence should be addressed.

E-mail: mkumar4@eng.ua.edu

Keywords: soil moisture proxy, fraction of potential evapotranspiration, kernel regression, surface flux equilibrium theory, ALEXI

Supplementary material for this article is available online

Abstract

Root zone soil moisture (RZSM) is a dominant control on crop productivity, land-atmosphere feedbacks, and the hydrologic response of watersheds. Despite its importance, obtaining gap-free daily moisture data remains challenging. For example, remote sensing-based soil moisture products often have gaps arising from limits posed by the presence of clouds and satellite revisit period. Here, we retrieve a proxy of daily RZSM using the actual evapotranspiration (ETa) estimates from Surface Flux Equilibrium Theory (SFET). Our method is calibration-less, parsimonious, and only needs widely available meteorological data and standard land-surface parameters. Evaluation of the retrievals at Oklahoma Mesonet sites shows that our method, overall, matches or outperforms widely available RZSM estimates from three markedly different approaches, viz. remote sensing data based Atmosphere-Land EXchange Inversion (ALEXI) model, the Variable Infiltration Capacity (VIC) model, and the Soil Moisture Active Passive (SMAP) mission RZSM data product. When compared with in-situ observations, unbiased root mean square difference of retrieved RZSM were 0.03 (m$^3$ m$^{-3}$), 0.06 (m$^3$ m$^{-3}$), and 0.05 (m$^3$ m$^{-3}$) for our method, the ALEXI model, and the VIC model, respectively. Better performance of our method is attributed to the use of both SFET for the estimation of ETa and non-parametric kernel-based method used to relate the RZSM with ETa. RZSM from our method may serve as a more accurate and temporally-complete alternative for a variety of applications including mapping of agricultural droughts, assimilation of RZSM for hydrometeorological forecasting, and design of optimal irrigation schedules.

1. Introduction

surface parameterization. In addition, the remote sensing derived surface soil moisture products that may be used in these methods, e.g. by (Chauhan et al 2003, Njoku et al 2003, Entekhabi et al 2010, Kerr et al 2010, Torres et al 2012, Wagner et al 2013), often suffer from temporal and spatial data gaps due to presence of cloud cover, narrow swath, and sparse revisit schedules (Walker and Houser 2004, Sabaghy et al 2018, Mao et al 2019). For example, the temporal resolution of the latest 3 km SM product from the National Aeronautics and Space Administration Soil Moisture Active Passive (SMAP) mission, the SMAP/Sentinel-1 L2_SM_SP SM product (Das et al 2018) is reported to vary between 3 and 12 days depending on the revisit schedules of backscatter measurements of Sentinel-1A and Sentinel-1B sensors (Entekhabi et al 2010, Das et al 2016, Mao et al 2019). The percentage of missing days for SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture Version 6 (SPL3SMP) product and SMAP Enhanced L3 Radiometer Global Daily 9 km EASE-Grid Soil Moisture Version 3 (SPL3SMP_E) product at Oklahoma Mesonet sites, the study area under consideration, are over 50% for ascending and descending overpass and over 15% for composite data during 2015–2019 (see figure S1 in supporting information (available online at stacks.iop.org/ERL/16/104007/mmedia)). Another promising approach for estimating RZSM is based on the use of Atmosphere-Land Exchange Inversion (ALEXI) surface energy balance model (Anderson et al 1997, Mecikalski et al 1999, Anderson et al 2011), which to a large extent reduces the need for model calibration. ALEXI is a state-of-the-science tool that has been frequently used to track soil moisture stress in crops and forests (Anderson et al 2007a, 2007b, 2016b, Mishra et al 2013, Knipper et al 2019) and forms the basis for next generation of moisture stress measurements (Anderson et al 2016a, Guan et al 2017, Cawse-Nicholson et al 2020, Fisher et al 2020). However, Anderson et al (2007a) reported that good quality thermal infrared (TIR) imagery, which is often used for moisture retrieval over the continental US, was only available around 30% of the time in their study area.

Here, we propose a method based on the surface flux equilibrium theory (SFET) to retrieve fraction of available water (fAW), a proxy for root-zone soil moisture. The proxy is then used to also obtain volumetric soil moisture (VSM) in the root zone using data of soil properties. The method yields gap-free daily estimate of fAW, while only needing widely available meteorological data and standard land-surface parameters. The remainder of this paper is organized as follows: section 2 presents details of our methodology and a concise overview of the study area and datasets. Results are presented in section 3. Section 4 presents conclusions and related discussion.

2. Methods

2.1. SFET

McColl et al (2019) presented SFET where it was hypothesized that in the regions with no or minimal advective moisture convergence (e.g. inland continental regions), the near-surface atmosphere is in the state of ‘surface flux equilibrium’, i.e. the surface heating and surface moistening terms in the near-surface relative humidity budget are in the state of equilibrium at daily to monthly time scale. McColl and Rigden (2020) provided physical explanations for why the hypothesis stands using a simple model of an idealized atmospheric boundary layer. The approach does not require an explicit parameterization of land surface conditions as it assumes that the turbulent fluxes at the land surface (latent and sensible heat fluxes) are encoded in the near-surface atmospheric states (near surface air temperature and specific humidity, respectively). Using this theory, the Bowen ratio (B) for a given location can be estimated as:

$$B \approx \frac{R_e C_p T_a^2}{\lambda Q_a}$$

where $B$ is Bowen ratio ($= H/\lambda E$) [-], $H$ is the sensible heat flux (W m$^{-2}$), $\lambda$ is the latent heat of vaporization (J kg$^{-1}$), $\lambda E$ is the latent heat flux (W m$^{-2}$), $R_e = 461.5$ is the gas constant for water vapor (J kg$^{-1}$ K$^{-1}$), $C_p = 1005$ is the specific heat capacity of dry air at constant pressure (J kg$^{-1}$ K$^{-1}$), $T_a$ is the screen-level air temperature (K), and $Q_a$ is the screen-level specific humidity of the air (kg kg$^{-1}$). The latent heat flux ($\lambda E$) can then be obtained using the following relation derived based on surface energy balance:

$$\lambda E = \frac{R_a - G}{1 + B}$$

where $R_a$ is the net solar radiation (W m$^{-2}$), and $G$ is the ground heat flux (W m$^{-2}$). More details regarding the calculation of $R_a$ and $G$ are provided in supporting information text S1. SFET based estimates of evapotranspiration have been shown to be remarkably accurate, with prediction errors comparable to errors in the eddy covariance measurements (McColl and Rigden 2020).

2.2. Soil moisture proxy retrieval

Here we retrieve fraction of available water (fAW), a commonly used proxy for soil moisture (Anderson et al 2007a, Hain et al 2009, 2011). fAW is defined as:

$$f_{AW} = \frac{(\theta - \theta_{wp}) \times d}{(\theta_{fc} - \theta_{wp}) \times d}$$

where $\theta$ is the soil moisture content in the root zone soil layer (m$^3$ m$^{-3}$), $\theta_{wp}$ is the soil moisture content at wilting point (m$^3$ m$^{-3}$), and $\theta_{fc}$ is the soil moisture
content at field capacity (m³ m⁻³), and d is the root-zone depth (m). As latent heat flux is dominantly controlled by evapotranspiration of soil moisture, or its proxy, fAW, can be potentially retrieved based on evapotranspiration estimates. For example, LSMs (Wigmosta et al 1994, Wood and Lettenmaier 1996, Panday and Huyakorn 2004, Hanasaki et al 2013, Wang et al 2013, Camporese et al 2014, Ferguson et al 2016) often use a simple soil moisture stress function to relate the simulated available water fraction (fAW) to the ratio of actual evapotranspiration (Eta) and potential evapotranspiration. This ratio, hereafter referred to as the fraction of potential evapotranspiration (fPET), is defined as:

\[ f_{\text{PET}} = \frac{E_T}{E_T} \]  

where \( E_T \) is actual evapotranspiration (mm d⁻¹), and PET is potential evapotranspiration (mm d⁻¹).

Our method obtains \( E_T \) using SFET as described in the previous section. The PET is estimated using the Penman–Monteith equation (Penman 1948, Monteith 1965) assuming the soil moisture conditions are at field capacity (see equation (3) in supporting information text S2). Once fPET is evaluated, the underlying relation between fAW and fPET is used to retrieve fAW.

The relation between fAW and fPET is usually derived using pre-defined functions (Mahrt and Pan 1984, Wetzel and Chang 1987, Stewart and Verma 1992, Anderson et al 2007a, Hain et al 2009). By inter-comparing four different functions that relate fAW and fPET, Hain et al (2009) reported better estimates of fAW when it is derived as a nonlinear function of fPET and \( B_c \). \( B_c \) is the plant factor that captures the effects of stomatal control on the plant transpiration under well-watered conditions (see equation (7) in supporting information text S2). Instead of using a pre-defined function structure (for which the results are shown in figure S2 in supporting information), here we fit a single statistical relation between observed fAW and fPET over all sites using a nonparametric kernel-based regression method (Nadaraya 1964, 1965, Watson 1964). The data used for developing the regression relation is restricted to a training period, and excludes the model evaluation period (more details are in section 2.4). Development of the relation using regional data is expected to moderate the discrepancy between estimated fAW and observed soil moisture, especially those arising from the mismatch in root zone depth (for which fPET is estimated) and the depth to which soil moisture observations are averaged to obtain fAW for the training period. The kernel regression method (see equation (9) in supporting information text S2) has been demonstrated to successfully capture nonlinear relations effectively in numerous studies (Rubin et al 2010, Kannan and Ghosh 2013, Salvi and Ghosh 2013, Raghav et al 2020). Here, the predictors of the kernel regression are fPET and \( B_c \), and the predictand is fAW. The kernel regression algorithm described by Hayfield and Racine (2008) and implemented in the R-package by (R Core Team 2013) is used. Notably, such a relation may be derived using any other observed soil moisture data sets existent within the region, such as SCAN and USCRN (Schafer et al 2007, Diamond et al 2013). In regions where soil moisture data does not exist, as noted earlier, fAW may be obtained using the pre-defined nonlinear functions.

2.3. Retrieving volumetric soil moisture (θ)

fAW (derived in section 2.2) is converted to actual volumetric soil moisture (\( \theta_c \), m⁴ m⁻³) using the field capacity (\( \theta_{fc} \)) and wilting point (\( \theta_{wp} \)) data using:

\[ \theta = (\theta_{fc} - \theta_{wp}) f_{AW} + \theta_{wp}. \]

Here we assume the field capacity to be the volumetric water content at −33 kPa and wilting point to be the water content at −1500 kPa. In absence of site-specific root zone depth data, \( \theta_{fc} \) and \( \theta_{wp} \) used here are the average values within 0–100 cm from the soil surface.

2.4. Data used for model implementation, validation, and inter-comparison

The model is implemented in the state of Oklahoma and the results validated at the Oklahoma Mesonet sites (Brock et al 1995). Data from North American Land Data Assimilation System-Phase 2 (NLDAS-2) (Xia et al 2012a) such as air temperature, specific humidity, wind speed, shortwave downward radiation, longwave downward radiation, and near-surface atmospheric pressure are used to obtain estimates of potential evapotranspiration, \( E_T \), fPET, and fAW. The data has a temporal resolution of 1 h and spatial resolution of 1/8°, and so are the corresponding resolutions of our evapotranspiration estimates. Our method uses two vegetation dependent parameters viz., minimum solar radiation for transpiration (\( R_d \)), and a parameter related to vapor pressure deficit (\( h_d \)) for the calculation of PET (see equation (6) in supporting information S2). The model uses MODIS Global 500 m Collection 5 land cover (Friedl and Sulla-Menashe 2015) to obtain the most prevalent land cover by area within a NLDAS grid. Vegetation dependent land surface parameters corresponding to this land cover is assigned to the model grid based on the Noah model lookup table (Koren et al 2010). Notably, the model does not require any parameter calibration. Soil properties used for retrieving VSM are obtained from the MesoSoil database (Scott et al 2013), which includes physical properties of 13 soil types for 545 individual soil layers across 117 Oklahoma Mesonet sites. MesoSoil provides the data (e.g. sand-silt-clay fraction, volumetric water content at −33 kPa,
moisture products. This includes variable infiltration widely available temporally continuous root zone soil moisture observation data as well. The relation may be developed using any alternative soil sites to ensure the generality of the approach, as such a relation that is then used for fAW estimation at all the sites used in this study. This was obtained over the aforementioned composite periods (2–5 days each due to data coverage gaps caused by clouds). The composite periods included 15–19 Jun 2002, 11–12 May 2003, 28–29 May 2003, 4–5 Jul 2003, 27–31 Jul 2003, 6–7 May 2004, 1–2 Jun 2004, and 1–3 Aug 2004. Although we simulate gap-free daily fAW and VSM estimates and these are used for evaluation against in situ observations, composite estimates are obtained for intercomparison with ALEXI- derived fAW estimates over the aforementioned composite periods. In contrast, as SMAP RZSM product is only available since March, 2015, comparison with it is performed for warm periods of 2015–2019.

Bias error (BE), mean absolute error (MAE), root mean square difference (RMSD), unbiased root mean square difference (ubRMSD), and Pearson correlation coefficient (R) are used as performance metrics to assess the accuracy of fAW estimates against in situ data. As the standard RMSD is sensitive to biases in the mean or high extreme values (outliers), here we also use the ubRMSD, a metric used by SMAP to quantify the product accuracy (Entekhabi et al. 2010).

3. Results and explanations

3.1. Evaluation of temporal variations in simulated fAW

The retrieved fAW, using the method outlined in section 2, captures more than 48% (R = 0.69) of the variations in the observed FWI (see figure 1(a)). For the same eight composite periods as used in Hain et al. (2009), the model’s performance is better during the composite periods (figure 1(b)) i.e. when clear sky conditions prevail, w.r.t. all periods (figure 1(a)). The possible reason for this is that the estimates of net radiation (and therefore latent heat flux) are generally more accurate during the largely clear sky, non-raining days. This is demonstrated at selected FluxNet sites where observed net short wave radiation data is available for evaluation (see figures S4 and S5 in supporting information). Overall, with respect to the observed, the presented method overestimates drier conditions and underestimates wetter
Figure 1. Comparison between daily observed and simulated fraction of available water (fAW) at selected Oklahoma Mesonet stations (identified in figure S3 in supplement information) during April–September of 2002–2004. The red straight line is the 1:1 line. Figure (a) shows the comparison for all days during April–September, 2002–2004, (b) shows comparisons for the eight composite periods (see section 2.4 for definition of composite periods).

Table 1. Error statistics for soil moisture proxy retrieved by our method and ALEXI during composite days at Mesonet stations. The error statistics for ALEXI shown here, have been obtained from Hain et al. (2009).

<table>
<thead>
<tr>
<th></th>
<th>R</th>
<th>RMSE (% fAW)</th>
<th>BE (% fAW)</th>
<th>MAE (% fAW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our method</td>
<td>0.77</td>
<td>18.86</td>
<td>8.7</td>
<td>14.58</td>
</tr>
<tr>
<td>ALEXI</td>
<td>0.69</td>
<td>21.3</td>
<td>-4.3</td>
<td>17.7</td>
</tr>
</tbody>
</table>

conditions. This is in line with the conclusions in other studies (Allen et al. 1998, Scott et al. 2003, Akuraju et al. 2017), where also it was observed that the relation between fAW and fPET is ineffective when soil moisture conditions are above (below) θfc (θwp).

Comparison of our fAW estimates against those obtained from ALEXI for the composite periods shows that our method, overall, matches or outperforms the ALEXI results (as reported in table 2 of Hain et al. (2009)) in the study area. ALEXI fAW has a larger scatter with R² of 0.48 as compared to a R² of 0.60 or R of 0.77 (see figure 1(b) and table 1) from our method. Our method shows a slightly larger positive bias, with BE of 8.7% as compared to BE of -4.3% for ALEXI model. The RMSE and MAE are found to be 18.86% (21.3%) and 14.58% (17.7%) respectively for our method (ALEXI model). It is to be noted that Hain et al. (2009) used a blended relation between fAW and fPET but we use a nonparametric kernel-based method for the same. To assess if better performance of fAW estimates from our method w.r.t. the ALEXI estimates is due to the use of SFET or the kernel-based method, we regenerate the fAW estimates using the blended relation used in Hain et al. (2009). Results (figures 1(b) and S2(b)) show our model performance (R = 0.77) does not change for the composite periods depending on the use of the nonparametric kernel-based method and the blended relation, and the estimates from either are better than that from ALEXI (R = 0.69, see table 1). Notably, the nonparametric kernel-based method yields better model performance when considering estimates from all the days (see figure 1(a) vs. figure S2(a)). These comparisons indicate better fAW estimates from our method is a result of the use of both SFET to obtain evapotranspiration and the non-parametric kernel-based method used to obtain the relation between fAW and fPET. Given that ALEXI derived fAW estimates have been demonstrated to be more effective than Eta Data Assimilation System (EDAS) for accurate Numerical Weather Prediction (NWP) forecasts (Hain et al. 2009), by extension, it may be claimed that estimates from our method will overperform the EDAS product as well. Notably, our method also provides temporally continuous daily estimates. In contrast, ALEXI estimates which uses TIR data suffer from large data gaps due to the presence of clouds and other satellite operational failures (Liou and Kar 2014).

Figures 2(b) and (c) shows the spatial variation of temporal correlations between daily fAW estimates for April–September (composite periods) of 2002–2004 from our method and in-situ observations. The correlation is positive at all the sites with the highest correlation of 0.77 (0.99) and the lowest correlation of 0.18 (0.11) at 0.05 significant level (see figures 2(d) and (e)). Between different land covers, the highest correlation is found in Woodland Savanna while the lowest correlation is observed in cropland areas (see figure 2(d)). Relatively poor performance in croplands w.r.t. woodland is often attributed to the heterogeneity introduced by irrigation and is consistent...
with the conclusions of Naeimi et al (2009). Notably, the Bowen ratio obtained from SFET (see equation (1) that is then used here for evaluation of fAW (using equations (4) and S9), is an integrated land-atmosphere feedback response of areas surrounding the Mesonet station. Hence, fAW derived from this method is likely to represent an effective soil moisture within the grid (of spatial resolution 1/8°), and may diverge from point estimates especially if the area experiences significant moisture heterogeneity.

3.2. Evaluation of spatial variations in simulated fAW

The retrieved fAW, overall, captures the spatial gradient of root-zone soil moisture conditions within the study area (figure 3). For example, for the composite period 15–19 June 2002, observed dry soil moisture conditions in extreme western Oklahoma and wet soil moisture conditions in eastern Oklahoma are reflected in the model estimates as well (see figure 3(a)). Similarly, our method also retrieves the dry soil moisture conditions all across Oklahoma during the composite period 27–31 July 2003 (figure 3(e)). Although, the overall spatial variability of soil moisture is captured by our model, there are disagreements. This could be due to a number of factors including scale mismatch between point measurements and our retrieval which is performed using meteorological data over a 0.125° × 0.125° grid. Additional sources of errors may be from varying fAW vs. fPET relations across different land covers, quality of input data, and errors in estimate of ET especially on cloudy days.

Next, the average of daily spatial correlations for all days during April–September of 2002–2004 is obtained. The average spatial correlation using data from all aforementioned days is equal to 0.62 (see figure S6 in supplement information). The corresponding average correlation for the eight composite periods is 0.71. During the eight composite periods, the highest correlation is 0.84 for the composite period 15–19 June 2002 when the soil moisture conditions are wettest (see figure 3(a)) and the lowest correlation is 0.54 during the composite period 27–31 July 2003 when the moisture conditions are driest (see figure 3(e)). The difference in model performance between wet and dry dates is likely due to multiple factors, including the existence of a larger spatial gradient in soil moisture during the wet composite period, and a smaller sensitivity of moisture dynamics on evapotranspiration when the ground is dry.
Figure 3. Spatial maps of fAW over the state of Oklahoma for different composite periods. For each composite period, the left panel shows in-situ observations of FWI at different mesonet sites and the right panel shows retrieved fAW using our method.

Figure 4. Comparison between daily observed and simulated volumetric soil moisture at the Oklahoma Mesonet sites (shown in figure S3 of supplementary information). (a) SM estimates from our method, (b) SM estimates from the VIC model, (c) SM estimates from our method, (d) SM estimates from the VIC model, and (e) SM estimates from SMAP. Top panel shows the comparison during April–September of 2002–2004, the period for which performance statistics of ALEXI based estimate are also available thus allowing intercomparison with it. The bottom panel shows the comparison during April–September of 2015–2019 during which both SMAP RZSM and VIC model data are available. The red straight line is the 1:1 line.

3.3. Retrieving actual soil moisture
The retrieved fAW estimates are converted to VSM using the method outlined in section 2.3. A comparison between VSM estimates from our method and observations is performed (figure 4(a)). The correlation coefficient between the simulated daily soil moisture from our model and observed during the warm period of 2002–2004 is close to 0.90 when considering
all the observation sites into the analysis. The significant increase in correlation for VSM w.r.t. that for fAW (0.9 vs. 0.69 as seen in figures 4(a) and 1(a), respectively) is attributable to spatial heterogeneity in soil properties (see supplement figure S7). The model, overall, captures the temporal dynamics of VSM in diverse land cover types (figure S8 in supplementary information), although just like for fAW (see figure 1) lower (higher) moisture values are overpredicted (underpredicted). We also compare VSM estimates from VIC model (figure 4(b)). In contrast to VSM estimates from our model, the VIC model estimates show a larger scatter with a R of 0.49. Furthermore, the VIC model overestimates the VSM observations with a bias error of 46.46% of the observations (see figure 4(b)). These results are consistent with Xia et al (2015) where it was reported that VIC model overestimated soil moisture. Overall, VSM estimates from our model (VIC model) show correlation, bias, RMSD, MAE, and unbiased RMSD of 0.90 (0.49), −0.78% (46.46%), 14.79% (54.95%), 11.71% (47.59%), and 0.03 m³ m⁻³ (0.06 m³ m⁻³), respectively. We also evaluate the error statistics for VSM retrievals during the eight composite periods (table 2). Results show our method outperformed both ALEXI and VIC model estimates. VIC model shows the largest bias errors among the three methods. Notably, our SM estimates adequately meet the SMAP mission requirement of ubRMSD to be less than 0.04 m³ m⁻³ (Chan et al 2016). Intercomparison of RZSM estimates from the proposed method, VIC, and SMAP L4 RZSM product for warm periods of 2015–2019 further affirm the superior performance of the proposed method (see figures 4(c)–(e)) against these widely available sister products.

### 4. Conclusions and synthesis

This study presented a new method to obtain daily estimates of RZSM proxy and volumetric RZSM. The method was based on SFET, and only needs readily available meteorological data and standard land surface parameterizations to obtain estimates of moisture proxy. In other words, the method is ‘calibration-free’ and thus suitable for predictions in ungauged basins. Soil moisture proxy estimates from the method were in good agreement with the in-situ measurements at the Oklahoma Mesonet sites both temporally and spatially. The estimate of volumetric soil moisture in the root zone adequately met the SMAP soil moisture retrievals requirement of ubRMSD < 0.04 m³ m⁻³.

An intercomparison of our estimate with the ALEXI model, which forms the basis for the next generation of moisture stress measurements (Anderson et al 2016a, Guan et al 2017, Cawse-Nicholson et al 2020, Fisher et al 2020), showed our method matched or outperformed ALEXI derived estimates. Better performance was found to be due to two reasons, a better evapotranspiration estimate using SFET and use of the non-parametric kernel-based method to obtain the relation between fAW and fPET. Another advantage of our method is its gap-free nature. In contrast, ALEXI or other TIR imagery-based retrievals of soil moisture provide estimates only during clear sky days as thermal imagery cannot be collected on cloudy days. Also, unlike our method, TIR-based methods for soil moisture proxy retrievals are dependent on a number of land surface parameters which are difficult to obtain in many cases. Given that ALEXI derived moisture proxy estimates have been demonstrated to be more effective than EDAS for accurate NWP forecasts (Hain et al 2009), by extension, it may be claimed that estimates from our method will overperform the EDAS product as well. Comparison against estimated moisture from VIC, a widely used land surface model, and SMAP L4 RZSM product which is generated by assimilating SMAP L-band brightness temperature observations into the NASA Catchment land surface model, showed our results outperformed them as well. These results indicate the advantage of our method over widely used land surface models, even while they use assimilation.

Although our validation results showed an overall satisfactory performance, it is to be noted that the performance is not competent across all soil moisture states and land covers. Our method especially fell short to capture extreme dry soil moisture conditions. Also, the performance was relatively poor in cropland settings. Subpar performance on occasions can be from multiple sources including (a) scale mismatch between the point measurements and model pixel, (b) quality and resolution of input meteorological data, (c) heterogeneity in soil properties, especially when converting moisture proxy to volumetric moisture, (d) absence of strong relation between fractional moisture content and the ratio of actual to potential evapotranspiration for extremely wet and dry moisture states, (e) mismatch between the actual root zone

---

**Table 2. Error statistics for volumetric soil moisture (m³ m⁻³) retrieved by our method, ALEXI, and VIC model during composite days at Mesonet stations. The error statistics for ALEXI shown here, have been taken from Hain et al (2009).**

<table>
<thead>
<tr>
<th>Method</th>
<th>RMSD (m³ m⁻³)</th>
<th>BE (m³ m⁻³)</th>
<th>MAE (m³ m⁻³)</th>
<th>ubRMSD (m³ m⁻³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our method</td>
<td>0.03</td>
<td>−0.005</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>ALEXI</td>
<td>0.06</td>
<td>−0.01</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>VIC</td>
<td>0.11</td>
<td>0.10</td>
<td>0.10</td>
<td>0.05</td>
</tr>
</tbody>
</table>
depth and the depths to which moisture observations are available and/or averaged for comparison, and (f) violation of assumptions that are inherent in SFET. Despite these limitations, this study highlights the advantages of our method over remote-sensing retrievals and land surface model predictions for RZSM retrievals. These advantages make the presented method apt for continuous assimilation of moisture in land surface and numerical weather prediction models. Gap-free and calibration-free moisture estimates from this method can be useful for many applications such as tracking crop stress, monitoring agriculture drought, irrigation management, estimation of groundwater recharge, etc. Future efforts towards improving the retrieved moisture estimates from the method may focus on disaggregation of evapotranspiration and use of alternative meteorological forcing products. To further improve confidence in the applicability of the method for a wider range of settings, forthcoming studies may perform evaluation in other hydroclimatic regions.

Data availability statement

The Oklahoma Mesonet soil moisture data can be obtained upon reasonable request at www.mesonet.org/index.php/past_data/data_request_form. The native NLDAS-2 and SMAP data used in this study are available from the respective data repositories at https://disc.gsfc.nasa.gov/ and https://nsidc.org/data/, respectively. Data that support the findings of this study will be made available upon reasonable request to the authors. The data that support the findings of this study are available upon reasonable request from the authors.

Code availability statement

Codes will be made available upon reasonable request from the authors.

Acknowledgments

This work is partially supported by NSF OIA-2019561, NSF EAR-1920425, and NSF EAR-1856054. We thank Collin Mertz for providing the Oklahoma Mesonet soil moisture and soil properties data which were used for the validation of our results. All analyses were carried out in R www.r-project.org/ and for this, we are very much thankful to the R Core Team (2020).

Conflict of interest

Authors declare no competing interests.

ORCID iDs

Pushpendra Raghav https://orcid.org/0000-0002-2982-069X

Mukesh Kumar https://orcid.org/0000-0001-7114-9978

References

Allen R G et al 1998 Crop evapotranspiration-guidelines for computing crop water requirements-FAQ irrigation and drainage paper 56 Fao, Rome 300 D0559
Anderson M C et al 2011 Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery Hydrol. Earth Syst. Sci. 15 223–39
Baldwin D, Manfreda S, Keller K and Smithwick E A H 2017 Predicting root zone soil moisture with soil properties and satellite near-surface moisture data across the conterminous United States J. Hydrol. 546 393–404
Chauhan N, Miller S and Ardanuy P 2003 Spaceborne soil moisture estimation at high resolution: a

IOP Publishing
Environ. Res. Lett. 16 (2021) 104007
P Raghav and M Kumar
microwave-optical/IR synergistic approach Int. J. Remote Sens. 24 4599–622
Chen X, Kumar M and McGlynn B L 2015 Variations in streamflow response to large hurricane-season storms in a southeastern US watershed J. Hydrometeorol. 16 55–69
Das N et al 2018 SMAP/sentinel-1 L2 radiometer/radar 30-second scene 3 km EASE-grid soil moisture version 2 (NASA National Snow and Ice Data Center DAAC)
Drusch M 2007 Initializing numerical weather prediction models with satellite-derived surface soil moisture: data assimilation experiments with ECMWF’s integrated forecast system and the TMI soil moisture data set J. Geophys. Res. Atmos. 112 D03102
Entekhabi D et al 2010 The soil moisture active passive (SMAP) mission Proc. IEEE 98 704–16
Fried M and Sulla-Menashe D 2015 MCD12C1 MODIS/Terra-aqua land cover type yearly L3 global 0.05°Deg CMG V006 [Data set] NASA EOSDIS Land Processes DAAC (available at: 10.5067/MODIS/MCD12C1(0.006) (Accessed 04 April 2020)
Godfrey C M and Stensrud D J 2008 Soil temperature and moisture errors in operationalEta model analyses J. Hydrometeorol. 9 367–87
Knipper K et al 2019 Evapotranspiration estimates derived using thermal-based satellite remote sensing and data fusion for irrigation management in California vineyards Irrig. Sci. 37 431–40
Liou Y-A and Kar S K 2014 Evapotranspiration estimation with remote sensing and various surface energy balance algorithms—a review Energies 7 2821–49
Li et al 2020 Plant hydraulics accentuates the effect of atmospheric moisture stress on transpiration Nat. Clim. Change 10 691–5
Mahfz A and Pan H 1984 A two-layer model of soil hydrology Bound. Layer Meteorol. 29 1–29
Manabe S 1969 Climate and the ocean circulation: I. The atmospheric circulation and the hydrology of the earth’s surface Mon. Weather Rev. 97 739–74
Nadaraya E A 1964 On estimating regression Theory Probab. Appl. 9 141–45
Nadaraya E A 1965 On non-parametric estimates of density functions and regression curves Theory Probab. Appl. 10 186–90


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


Penman H L 1948 Natural evaporation from open water, bare soil and grass Proc. R. Soc. A 193 120–45


R Core Team 2013 R: a language and environment for statistical computing (Vienna, Austria: R foundation for statistical computing) (available at: http://www.r-project.org)

Raghav P et al 2020 Revamping extended range forecast of Indian summer monsoon Clim. Dyn. 55 3397–411

Riechle R H et al 2017 Assessment of the SMAP level-4 surface and root-zone soil moisture product using in situ measurements J. Hydrometeorol. 18 2621–45

Rubin Y, Chen X, Murakami H and Hahn M 2010 A Bayesian approach for inverse modeling, data assimilation, and conditional simulation of spatial random fields Water Resour. Res. 46 W10523


Salvi K and Ghosh S 2013 High-resolution multisite daily rainfall projections in India with statistical downscaling for climate change impacts assessment J. Geophys. Res. Atmos. 118 3557–78


Stewart J and Verma S 1992 Comparison of surface fluxes and conductances at two contrasting sites within the FIFE area J. Geophys. Res. Atmos. 97 18623–8


Torres R et al 2012 GMES Sentinel-1 mission Remote Sens. Environ. 120 9–24

Trenberth K E, Smith L, Qian T, Dai A and Fasullo J 2007 Estimates of the global water budget and its annual cycle using observational and model data J. Hydrometeorol. 8 758–69


Wagner W et al 2013 The ASCAT soil moisture product: a review of its specifications, validation results, and emerging applications Meteorol. Z. 22 5–33


Watson G S 1964 Smooth regression analysis Sankhya 26 359–72


