Hydrological Sciences Journal

ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage: https://www.tandfonline.com/loi/thsj20

Performance evaluation of soil moisture profile estimation through entropy-based and exponential filter models

Vikalp Mishra, W. Lee Ellenburg, Kel N. Markert & Ashutosh S. Limaye

To cite this article: Vikalp Mishra, W. Lee Ellenburg, Kel N. Markert & Ashutosh S. Limaye (2020) Performance evaluation of soil moisture profile estimation through entropy-based and exponential filter models, Hydrological Sciences Journal, 65:6, 1036-1048, DOI: 10.1080/02626667.2020.1730846

To link to this article: https://doi.org/10.1080/02626667.2020.1730846

Published online: 24 Feb 2020.

HYDROLOGICAL SCIENCES JOURNAL
2020, VOL. 65, NO. 6, 1036–1048
https://doi.org/10.1080/02626667.2020.1730846

Performance evaluation of soil moisture profile estimation through entropy-based and exponential filter models

Vikalp Mishra[a,b], W. Lee Ellenburg[a,b], Kel N. Markert[a,b] and Ashutosh S. Limaye[a]

[a]NASA-SERVIR, Marshall Space Flight Center, Huntsville, Alabama, USA; [b]Earth System Science Center, The University of Alabama in Huntsville, Huntsville, Alabama, USA

ABSTRACT

In this study we analyzed two models commonly used in remote sensing-based root-zone soil moisture (SM) estimations: one utilizing the exponential decaying function and the other derived from the principle of maximum entropy (POME). We used both models to deduce root-zone (0–100 cm) SM conditions at 11 sites located in the southeastern USA for the period 2012–2017 and evaluated the strengths and weaknesses of each approach against ground observations. The results indicate that, temporally, at shallow depths (10 cm), both models performed similarly, with correlation coefficients (*r*) of 0.89 (POME) and 0.88 (exponential). However, with increasing depths, the models start to deviate: at 50 cm the POME resulted in *r* of 0.93 while the exponential filter (EF) model had *r* of 0.58. Similar trends were observed for unbiased root mean square error (ubRMSE) and bias. Vertical profile analysis suggests that, overall, the POME model had nearly 30% less ubRMSE compared to the EF model, indicating that the POME model was relatively better able to distribute the moisture content through the soil column.

ARTICLE HISTORY
Received 10 April 2019
Accepted 6 January 2020

EDITOR
A. Fiori

ASSOCIATE EDITOR
S. Pingale

KEYWORDS
exponential filter; soil moisture; profile; POME; remote sensing

1 Introduction

Soil moisture (SM) is a central component of the global hydrologic budget and often considered as one of the most challenging variables to estimate, particularly at regional and global scales (Robock *et al.* 1995, Lettenmaier *et al.* 2015). Monitoring of SM is traditionally accomplished using ground observations, sophisticated land surface models (LSMs), and/or remote sensing methods. Ground observations despite being relatively more accurate, are often monitored sparsely (Aghakouchak *et al.* 2015) thus cannot be used to assess the spatial dynamics at regional scales.

Further, they are often only active during field campaigns and thus temporally limited and inconsistent (Reichle 2005). However, LSMs can represent SM dynamics consistently in both space and time by capturing the complex physical processes in moisture movement through a soil column using physically based mathematical equations. However, one of the limitations of the LSMs is that they depend on a multitude of input parameters to characterize soil properties which are often not available and therefore parameterized. Thus, though spatially and temporally consistent, LSMs may struggle to represent localized soil characteristics, particularly in data-limited regions.

An alternative to ground observations and LSMs is the use of remotely sensed satellite techniques to estimate SM. Microwave (MW) sensors, such as the Advanced Microwave Scanning Radiometer (AMSR); Soil Moisture and Ocean Salinity (SMOS); and Soil Moisture Active Passive (SMAP), can provide SM measurement over large spatial scales with reasonable accuracy (Njoku *et al.* 2003, Entekhabi *et al.* 2010, Kerr *et al.* 2010). However, the MW-based SM measurements only account for the top few (~0–5) cm of the soil column. There have been multiple studies exploring the assimilation/forcing of the surface SM data from MW sensors into LSMs and crop models with some success in reducing the model errors (Bolten and Crow 2012, Ridler *et al.* 2014, Lievens *et al.* 2015, Yang *et al.* 2016, Baldwin *et al.* 2017, 2019).

However, the general observations from these studies is that the surface SM alone will have only limited success, particularly due to the fact that there is a disconnect between surface and subsurface SM dynamics. This disconnect often leads to increased uncertainties in the model at the root zone. Therefore, a demand for reliable and robust surface to root-zone SM data from satellite sensors has grown over the years.

Root-zone SM is often used in monitoring and evaluating agricultural drought and is a fundamental factor in plant–soil–atmosphere interactions affecting plant growth, sustenance, and evapotranspiration (ET) (Anderson *et al.* 2007, 2012, Bolten and Crow 2012). Furthermore, processes such as infiltration and ET influence SM dynamics at shallow depths (<-100 cm), which can be extremely variable with time and depth (Scott *et al.* 2003, Starks *et al.* 2003). These interdependent hydrological processes affect the daily distribution of SM in the root zone, which is a function of soil characteristics as well as canopy root density and distribution (Mishra *et al.* 2015).

Thus, information pertaining not only to the amount of moisture within a soil column, but rather detailed knowledge of the distribution becomes essential for effective forcing/assimilation of SM into a LSMs or as an independent data source. A multitude of approaches have been proposed to estimate root-zone SM profiles, including traditional methods such as regression, inversion, knowledge-based, and water balance (Kostov

CONTACT Vikalp Mishra vikalp.mishra@nasa.gov 320 Sparkman Dr., Huntsville, AL 35805, USA

and Jackson 1993, Singh 1997, Srivastava *et al.* 1997, van Dam and Feddes 2000, Sadeghi *et al.* 2017), as well as statistical/empirical methods including maximum entropy, neural networks, wavelets or exponential filters (Wagner *et al.* 1999, Al-Hamdan and Cruise 2010, Singh 2010, Kornelsen and Coulibaly 2014, Zhou *et al.* 2016, Pan *et al.* 2017, Qin *et al.* 2018, Sun and Niu 2019). The selection of an approach is usually based on the application, complexities involved and amount of *a priori* information available. For large-scale applications, the number of inputs required becomes paramount, especially in data-sparse regions. Therefore, approaches that require minimum inputs have generated a lot of interest in large-scale applications. In this study, we investigate the applicability and limitations of two such commonly used methods: the exponential filter (EF) and the principle of maximum entropy (POME) model.

The EF-based method to deduce root-zone SM has received a lot of attention in recent years, particularly from the remote sensing community (Albergel *et al.* 2008, Ford *et al.* 2014, Cho *et al.* 2015, Tabatabaeenejad *et al.* 2015, Peterson *et al.* 2016, Tobin *et al.* 2017, Wang *et al.* 2017, Sure and Dikshit 2019). Exponential filters rely on the relationship between the surface and underlying root-zone moisture conditions and assume a hydrologic equilibrium within the soil profile. It has been argued that the EF model tends to provide relatively better results under humid conditions (Mahmood and Hubbard 2007, Tobin *et al.* 2017). However, due to the inherent assumption of no transpiration and constant hydraulic conductivity (Wagner *et al.* 1999, Albergel *et al.* 2008), the EF tends to have reduced sensitivity during prolonged dry periods where plant uptake is the main driving force within root-zone moisture movement. Furthermore, the association of surface to root-zone SM tends to weaken with depth (Carlson *et al.* 1995) for most soil types; thus, the sensitivity of the EF reduces at deeper layers (approx. >50 cm) (Ford *et al.* 2014, Peterson *et al.* 2016). Despite these limitations, the EF method is commonly used among the remote sensing community due to its minimum input requirements; the method only requires surface SM data as input, which is readily available from MW sensors globally.

Another statistical model that is becoming popular is based on the maximization of entropy, i.e. the POME model. It has been argued that the complexities intrinsic to the movement of moisture through a soil column can be best described through the description of its entropy (Mays *et al.* 2002, Pachepsky *et al.* 2006, Al-Hamdan and Cruise 2010, Singh 2010, Castillo *et al.* 2015). The maximization of entropy characterizes the diffusion of moisture through a soil column as a function of time. Unlike traditional approaches, the POME method does not depend on *a priori* information (known or assumed) about the nature or shape of the profile. The method guarantees the minimum variance unbiased profile subject to boundary conditions (Al-Hamdan and Cruise 2010). The POME model has been evaluated extensively under laboratory settings, as well as against ground observations, by Al-Hamdan and Cruise (2010), Singh (2010), and Mishra *et al.* (2015). In addition to the surface SM, The POME model requires two additional inputs: the mean moisture content and the bottom-most layer SM value. Therefore, the POME is relatively better tied to the SM dynamics and physical processes and could potentially represent the SM state more accurately. Like the EF model, the inputs to POME can be obtained from satellite observations; the surface moisture condition from MW sensors, while the mean moisture content can be deduced from thermal infrared data (Crow *et al.* 2008, Hain *et al.* 2009, Mishra *et al.* 2013, 2018). The bottom-most layer (~100–200 cm) typically does not vary temporally and therefore can be parameterized (Mishra *et al.* 2018).

The EF and POME models are inherently different methods to estimate root-zone moisture content; the EF model tries to mimic the temporal dynamics, whereas the POME attempts to effectively distribute the moisture through a vertical soil column. Although, both models are used in remote sensing applications to deduce root-zone SM content, a detailed comparison of the performance between the two model has not yet been conducted. In this study, we evaluate and analyze the performance of both root-zone SM deduction models against ground observations at various depths. The root-zone SM profiles were developed for multiple years (2012–2017) over 11 sites located in the southeastern USA. The main objective of this study is to evaluate the performance of both models as an algorithm for deducing and developing root-zone SM profiles over a diverse range of soils. As mentioned above, the models are typically applied with satellite remote sensing; however, to reduce the number of uncertainties in this comparison, we evaluate both models under optimum conditions, i.e. forced with *in situ* observational datasets. In addition to quantitatively evaluating both approaches, the expected outcome of this investigation is to understand and help guide the application of these two approaches by the remote sensing community.

2 Method

This study focuses on the southeastern USA, where a relatively dense network of SM probes exists. The climate of the southeast Unite States ranges from subtropical humid to semi-humid, with average rainfall rates varying from 1000 mm to 1500+ mm per year. Though subtropical, many areas in the region experience water-limited conditions throughout the year (Ellenburg *et al.* 2018). *In situ* SM observations were obtained from the US Department of Agriculture Soil Climate Analysis Network (SCAN) sites throughout the region. A total of 27 SCAN sites were available for analysis; however, only 11 sites were selected in this study since they had at least 50% data availability for the study period of 2010–2017 encompassing a range of wet to dry years (Fig. 1). Most of the earlier studies by Wagner *et al.* (1999), Albergel *et al.* (2008), Ford *et al.* (2014), Peterson *et al.* (2016), and Tobin *et al.* (2017), only selected sites that had a significant relationship between surface and bottom layers for analysis. In this study, no such pre-screening of sites was performed based on factors such as lagged autocorrelations; in fact, if these criteria had been followed, only one of the 11 sites could be used. This was done in an attempt to evaluate the performance of both models under data-scarce conditions, where any type of ground observation is at a premium. Moreover, such an approach allows us to assess the full capabilities (or lack thereof) of these models by remaining agnostic to underlying soil characteristics and any relationship between the layers. The 11 stations that were employed represent good distributions across soil properties with two sandy, six loamy sands/sandy loams, and three silty loams (Table 1). Typically, SCAN stations are operated in agricultural fields or close by, and the stations used in this analysis

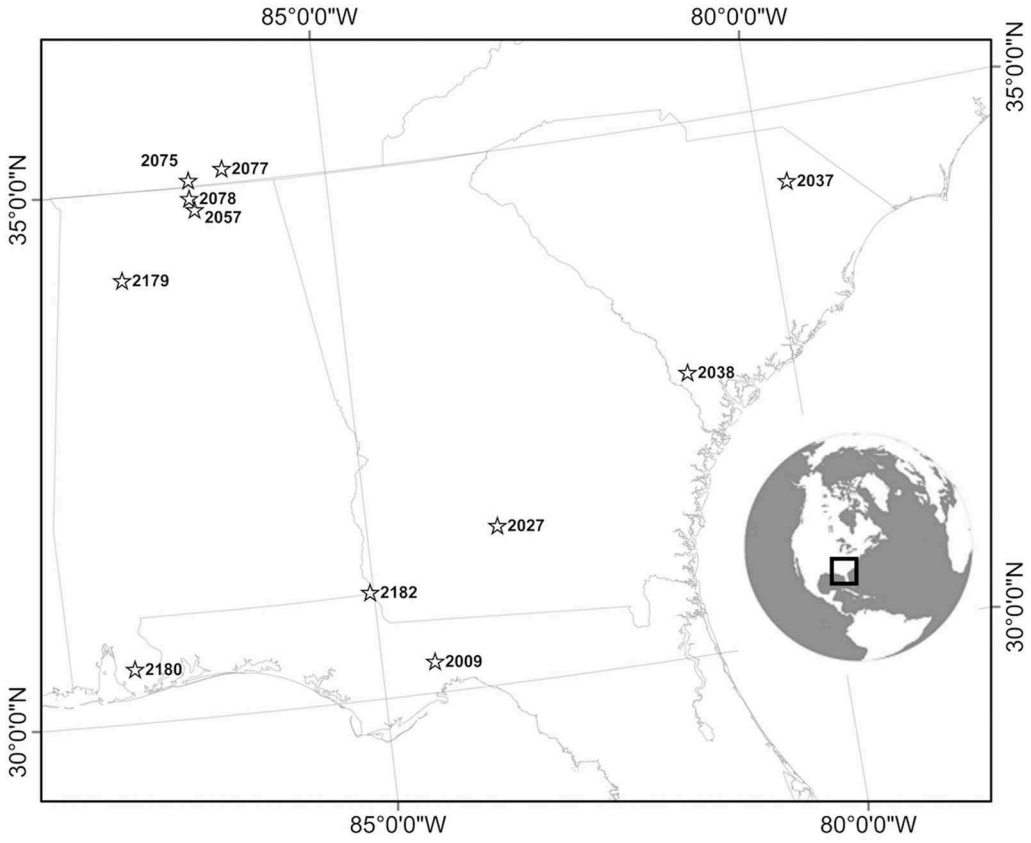

Figure 1. Experimental study area identifying the 11 SCAN stations used.

represent a wide variety of profile distributions (dynamic, wetting and drying). For this analysis, the observational data were used to provide each model with its minimum inputs. All of the methods presented here were implemented in the Python programming language, version 3.6, using the NumPy (van der Walt *et al.* 2011), SciPy (Oliphant 2007, Millman and Aivazis 2011), and Pandas (McKinney 2010) packages. The source code developed and used in the study for the Exponential filter and POME model are publicly available at https://github.com/Vikalp86/soil-moisture-profile.

2.1 Exponential filter

The Exponential function used in this comparative analysis to deduce root-zone moisture content was developed by Wagner *et al.* (1999) and later refined into a recursive form by Albergel *et al.* (2008). Its exponential decay function is based on a time lagged relationship between surface variations and the root-zone moisture content. The recursive formula for deducing root-zone moisture content is given as:

$$SM^*_{rz(t)} = SM^*_{rz(t-1)} + K_t\left(SM^*_{sfc} - SM^*_{rz(t-1)}\right) \quad (1)$$

where SM* is the normalized SM content based on minimum and maximum moisture content over a period of time, rz and sfc refers to root-zone and surface values, respectively; t is the time (in days) and K is the gain function computed as:

$$K_t = \frac{K_{t-1}}{K_{t-1} + e^{\frac{t_n - t_{n-1}}{T}}} \quad (2)$$

where T is the timescale of moisture variations in days, while $t_n - (t_{n-1})$ is the difference in days that the surface moisture content value was available between the current and previous times. At the start of the simulation or when the lag in data availability is greater than 12 days the model is re-initialized with $SM^*_{rz(1)} = SM^*_{sfc(1)}$ and $K(1) = 1$ (Tobin *et al.* 2017). The surface (~5 cm) moisture content input, SM^*_{sfc}, was obtained from the SCAN observations. The only unknown in this function is T which conceptually represents several parameters affecting temporal dynamics of the moisture content within a soil column (Ceballos *et al.* 2005). The accuracy of the exponential function is dependent on T (Albergel *et al.* 2008) and is usually calibrated to individual sites. Albergel *et al.* (2008) found little correlation between soil properties and T along with a weak relationship to

Table 1. NRCS SCAN sites locations and soil texture at different layer depths. [Sand/clay content in percentages]. *Soil texture information from SSURGO.

Site	Lat/Lon	Depth (cm)				
		5	10	20	50	100
2027	31.5/−83.5	88.1/02.8	88.1/02.8	88.1/02.8	88.1/02.8	78.5/11.5
2037*	34.3/−79.7	95.0/04.0	95.0/04.0	95.0/04.0	95.0/04.0	95.0/04.0
2038*	32.7/−81.2	79.0/04.0	79.0/04.0	79.0/04.0	60.0/26.0	60.0/26.0
2057	34.8/−86.6	29.2/22.9	29.2/22.9	29.2/22.9	20.9/41.3	18.6/47.5
2075	35.1/−86.6	14.8/19.1	14.8/19.2	12.9/23.4	12.9/36.0	21.4/47.8
2077*	35.1/−86.2	06.4/26.6	6.4/26.6	06.4/26.6	05.7/37.2	10.8/33.0
2078	34.9/−86.6	12.7/31.1	13.1/36.7	13.1/36.7	07.5/54.1	07.5/54.1
2179*	34.2/−87.5	63.0/11.0	63.0/11.0	46.0/20.0	46.0/20.0	50.0/25.0
2180*	30.5/−87.7	66.0/14.0	66.0/14.0	39.0/24.0	39.0/24.0	39.0/24.0
2182*	31.0/−85.0	85.0/06.0	85.0/06.0	85.0/06.0	85.0/06.0	90.0/05.0

climate (wind speed, ET, season); thus, an optimized T (T_{opt}) is commonly determined by assessing the prediction accuracy using a metric over a calibration period. More specifically, the SM time series is simulated for the calibration period using all realistic values of T, then the objective function is calculated for the simulated time series against the observed data. The T value is optimized for each soil layer and the value that provides the best simulation for each soil layer is selected. This "brute-force" process can be applied to optimize T due to the quick computation time of the exponential filter, the limited number of parameters being optimized, and the constrained dimension space of the parameter – typically 1 to ~60 days (Qiu *et al.* 2014, Tobin *et al.* 2017). For this analysis, a two-year period (2010–2011) was used as spin-up to compute T_{opt} for each test site and each layer separately. The Nash-Sutcliffe efficiency (NSE) coefficient was used as the primary objective function; however, given the sensitivity of NSE to biases, an unbiased-RMSE (ubRMSE) was also used as a secondary objective function when optimizing $T_{opt.}$

2.2 The POME model

Using Shannon entropy (Shannon 1948), the POME model by Al-Hamdan and Cruise (2010) was used in this study to develop vertical SM profiles. The model requires three inputs: boundary conditions (surface and bottom layer moisture content) along with the mean moisture content through the soil column. The model develops a monotonic profile subject to boundary conditions and constraints. The model uses two constraints: (a) the total probability constraint and (b) the mass balance constraint. The constraints are necessary to relate the first moment in probability space to the mean water content of the soil column in physical space. The main POME equation is given as:

$$\Theta(z) = \frac{\ln\left[\exp(\lambda_2\Theta_0) \pm \lambda_2 \exp(1-\lambda_1)\left(Z/L\right)\right]}{\lambda_2} \quad (3)$$

Here, $\Theta(z)$ is the effective moisture content at layer depth z, λ are the Lagrange multipliers, L is the total soil column depth, and Θ_0 is the surface effective SM. The Lagrange multipliers can be solved from application of the constraints and the boundary conditions. For a complete model description and derivation, readers are referred to Al-Hamdan and Cruise (2010).

In this study, the model was forced with boundary conditions at 5 and 100 cm depths, using data from the SCAN observations. The profile mean moisture content was also computed from the SCAN observations. These inputs were used to obtain SM profiles at 5-cm depth increments from 0 to 100 cm. The model represents a monotonically increasing (+) or decreasing (–) function to develop dry or wet profiles, respectively. However, not all profiles are monotonic in nature; some tend to have at least one prominent inflection point (the dynamic case) (Mishra *et al.* 2015). The POME model identifies the dynamic case when the mean moisture content is outside the limits of the boundary condition and POME has no solution. The POME model can handle the profiles with inflection point by sub-dividing the profile into two monotonic profiles and applying the model over them separately. The presence of an inflection point poses another challenge in terms of input requirements; the depth as well as SM value at the inflection point are necessary. Inflection point location and value is a function of soil characteristics and moisture inputs from rainfall or irrigation. Mishra *et al.* (2015) suggest that, for a given soil type, the inflection point does not vary much over time. Additionally, their study showed that, for most of the soils (except extremely porous), the inflection point was typically present around 20 cm depth. In terms of SM value at the inflection point, they argued that for large-scale applications, the initial inflection point value can be assumed as field capacity and can be used as a calibration parameter to achieve mass balance in each part of the profile. Therefore, in the spirit of minimum data input requirements, as would be the case within the remote sensing environment, for this study, the inflection point was assumed to be at 20 cm depth and an initial value was set as the field capacity. This value was iteratively reduced until the correct mass balance was achieved.

3 Analysis and discussion

When considering the performance evaluation of these two statistical models, it is important to note the inherent differences in model conceptualization, the temporal aspect of the EF *vs* the vertical distribution of the POME. Hence, most of the earlier studies pertaining to the evaluation of the EF model were based on time series analysis (Albergel *et al.* 2008, Ford *et al.* 2014, Tobin *et al.* 2017), while the POME model analysis was typically restricted to analyzing the vertical distributions (Al-Hamdan and Cruise 2010, Singh 2010, Mishra *et al.* 2015). Thus, such comparisons may not fully encapsulate the strengths (or weaknesses) of the respective models in vertical or temporal space. Although, the models were developed to estimate different aspects of SM dynamics, in this study we analyzed model performance temporally as well as vertically to evaluate the SM dynamics in both spaces.

Both the EF and POME models were employed to developed profiles down to 100 cm depth. The EF model was applied at depths of 10, 20, 50 and 100 cm, while the POME model, being a continuous integral, produced the SM profile at 5 cm intervals in the range 0–100 cm. However, SM content at only three layer depths – 10, 20, and 50 cm corresponding to the sensing depths of the SCAN observations – could be considered for performance evaluation. Surface (5 cm) and bottom (100 cm) layer data were not used for evaluation, since these were supplied as input to one or both of the models. Both models were run for a period of 6 years (2012–2017) at a 1-day time step over 11 SCAN sites. Years 2010 and 2011 were used as a calibration period to initialize the EF model. The SM estimates in this study are represented in effective terms as: $(\theta - \theta_r)/(\eta - \theta_r)$, where η is the soil porosity, θ is volumetric SM, and θ_r is the residual moisture content.

3.1 Temporal analysis

The exponential function is known to perform best in moderately moist soils where the sand or clay content does not make up the majority of the soil texture (Albergel *et al.* 2008, Ford

Table 2. Table listing optimal NSE (T_{opt} in days) at each layer depth for all sites using two-years of spin-up period (2010–2011).

Site	10 cm		20 cm		50 cm		100 cm	
	t	T_{opt}	*t*	T_{opt}	*t*	T_{opt}	*t*	T_{opt}
2009	0.73	1	0.52	2	−0.15	6	−2.51	7
2027	0.9	1	0.75	1	−0.81	8	−35.2	59
2037	0.92	1	0.63	1	−1.31	39	−3.54	59
2038	0.87	1	0.84	1	−35.3	39	−51.7	59
2057	0.87	1	−1.61	5	−9.93	16	−4.57	21
2075	−0.88	1	−0.01	6	0.67	9	−49.9	59
2077	0.82	1	−0.31	4	0.13	17	0.04	47
2078	−0.22	1	−5.52	2	−0.12	39	−0.96	12
2179	0.73	1	−0.25	4	−2.95	7	0.45	8
2180	0.83	1	0.09	2	−12.3	39	−14.2	59
2182	0.87	2	0.80	3	−22.1	8	−4.76	4
Mean	*0.59*	*1*	*−0.37*	*3*	*−7.65*	*21*	*−15.17*	*36*

et al. 2014, Tobin *et al.* 2017). However, in our analysis, soils with higher sand content are found to relate to a more significant T_{opt} in the calibration period, resulting in a better performance overall, especially in the upper two layers. Table 2 summarizes the optimal NSE and T_{opt} at all SCAN sites for all layers found during the calibration period. Only two of the 11 sites had negative optimum NSE (2075 and 2078) sites for depths of 50 and 100 cm. Overall, the mean optimum NSE for all sites was 0.59 at 10 cm, indicating a strong relationship between the surface and the layer immediately adjacent. At 20 cm depth, the mean NSE across all sites was – 0.37, though the mean might be skewed due to a single site (2078) with NSE of – 5.52. Excluding site 2078, the mean NSE was 0.14 for the rest of the sites, indicating that the model was better in simulating the SM condition than assuming the mean. The results of optimum NSE were consistent with the findings of Carlson *et al.* (1995) and Capehart and Carlson (1997), who argued that as the depth increases, the surface and root-zone essentially gets decoupled and exhibit very weak (if any) relationship.

Figure 2 shows the boxplot of the overall temporal statistics across all sites of the POME and EF derived root-zone SM estimates against the SCAN observations by layer. In terms of correlation coefficient (*r*), the POME estimates outperform the EF model at all depths. This can be explained in part by the two additional inputs (mean SM and bottom boundary conditions) that are required by the POME model. However, the median values at 10 cm depth were very similar in both cases, with r = 0.89 and 0.88 for the POME and EF models, respectively. The difference increases with depth: r = 0.91 *vs* 0.83 at 20 cm, and 0.93 *vs* 0.58 at 50 cm for the POME and EF models, respectively. Similar trends were observed for the unbiased RMSE (ubRMSE). In fact, the ubRMSE (effective SM for the POME model) tends to reduce with depth, while the reverse is observed for the EF SM estimates (Table 3). It can be argued that, due to the use of the lower boundary condition as one of the inputs to the POME model, the errors in lower depths are greatly reduced. In the case of bias, the EF SM estimates at 10 cm depth have relatively low mean bias (0.005 in effective SM). With depth, the bias tends to become negative indicating general underestimation of SM by the EF model. The POME, on the other hand, showed relatively low bias consistently at all depths. Detailed statistical results of both models against the SCAN observations are presented in Table 3.

Although, Table 3 details the temporal analysis results for each of the 11 sites by layers, two locations (site 2027 and 2057) were chosen to discuss the model performance in detail. These two sites represent the extremes (with respect to soil properties and model agreement) among all the sites evaluated. Figure 3 showcases the SM time series for year 2016 to demonstrate the temporal dynamics estimated by the models compared with the SCAN observations. Site 2027 is an example where the two models track closely to each other as well as the *in situ* observations, while site 2057 demonstrates where the two models depart dramatically in the lower layers. At Site 2027 located in South Georgia, the soil texture is reported as sandy (0–60 cm), with sand content at 88.1% (Table 1). One can see that both models slightly overestimate the moisture content in each layer

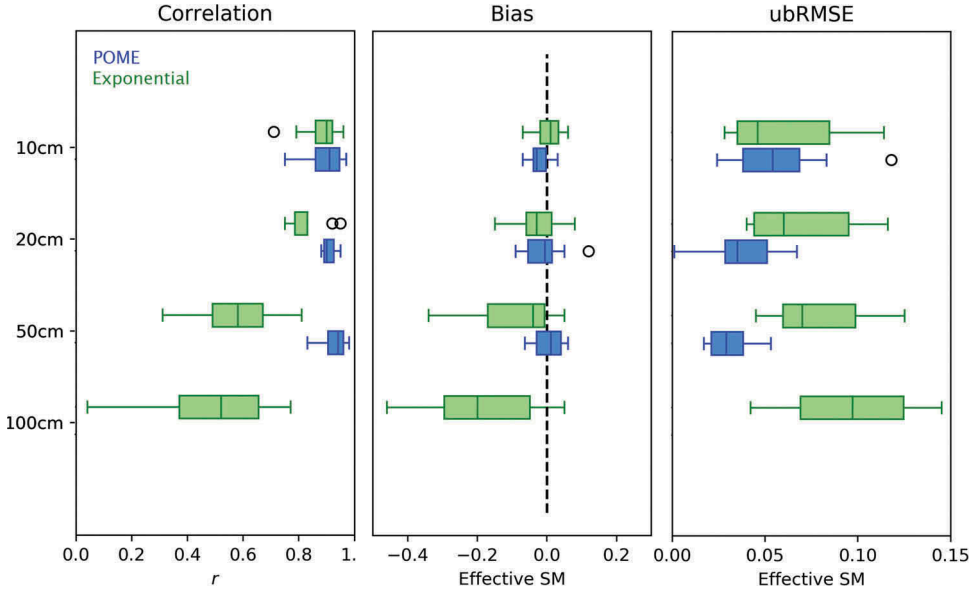

Figure 2. Boxplot of temporal statistics of correlation coefficient (*r*), bias and ubRMSE for the exponential filter and POME model against SCAN observations over all 11 sites.

presented for most of the year (except around November). While some differences exist throughout the time series, the two models differ the most during the drying period (generally, August/September–October) and particularly in the lower layers (Fig. 2(a)). The EF model begins to under predict once the moisture movement is less tied to surface events. Once the surface is dried, its hydraulic connectivity with the subsurface is significantly reduced and the two moisture pools (surface and root-zone) become effectively decoupled (Carlson *et al.* 1995, Capehart and Carlson 1997). Additionally, the EF model begins to smooth out the individual responses to precipitation events at the lower layers, while the POME model shows a clear positive bias despite being given the 100 cm boundary condition.

The relationship between the surface, 20 and 50 cm layers appears to be less explained by the EF at site 2057 (Fig. 2(b)). Site 2057, located in North Alabama, is reported as a mix of loam (0–20 cm, 29% sand), clay loam (20–40 cm, 24.3% sand) and clay (40–100 cm, 20.1% sand). For site 2057, the T_{opt} was 1, 5 and 16 days for layers at 10, 20 and 50 cm, respectively (Table 2). For comparison, the T_{opt} for site 2027 with sandy soil was 1 day at the 10 and 20 cm layers and 8 days at 50 cm. The shorter T_{opt} for site 2027 is consistent with high hydraulic conductivity associated with sandy soil types. The NSE was 0.9 and 0.75 for the top two layers, while at 50 cm depth NSE was –0.8, indicating weakening of the relationship between surface and subsurface SM dynamics. At site 2057, T_{opt} was moderately longer at the lower layers which is typical of loamy soils, however, the NSE was negative in both lower layers (NSE < – 1.6), indicating no clear relationship between surface and lower layers.

Table 3. Time series analysis results from EF and POME models for 2012–2017. *N*: number of days SCAN data was available; NSE: Nash-Sutcliffe efficiency; *r*: correlation coefficient; *E*: ubRMSE; *B*: bias. The ubRMSE and bias are in effective SM.

Site	Depth (cm)	*N**	Exponential filter				POME			
			NSE	*r*	*E* (10^{-2})	*B* (10^{-2})	NSE	*r*	*E* (10^{-2})	*B* (10^{-2})
2009	10	864	0.63	0.85	3.8	−2.1	0.70	0.85	4.1	1.8
	20	1678	0.07	0.75	4.2	4.1	0.83	0.88	2.6	−0.9
	50	1678	−0.88	0.62	4.5	5.4	0.73	0.83	2.9	−1.2
	100	1676	−6.92	0.35	12.6	5.3	–	–	–	–
2027	10	2307	0.84	0.92	2.9	−0.1	0.83	0.95	2.5	−0.7
	20	2307	0.44	0.81	4.6	−3.0	0.76	0.88	3.5	0.5
	50	2306	−0.7	0.48	7.0	−3.0	0.17	0.96	2.2	−6.4
	100	2306	−114.6	0.14	7.5	−34.4	–	–	–	–
2037	10	1492	0.91	0.96	2.8	1.1	0.92	0.97	2.4	−1.8
	20	1492	0.73	0.92	4.0	3.5	0.56	0.95	3.1	−5.5
	50	1492	−1.79	0.51	6.0	−9.1	0.81	0.96	1.8	1.6
	100	1492	−6.1	0.52	5.7	−16.0	–	–	–	–
2038	10	1693	0.84	0.95	3.2	2.5	0.28	0.90	5.4	−7.2
	20	1995	0.61	0.95	4.2	−5.0	0.49	0.90	5.6	−7.2
	50	1995	−92.6	0.50	6.9	−33.2	0.35	0.89	3.7	5.1
	100	1993	−184.9	0.64	4.2	−44.2	–	–	–	–
2057	10	2297	0.68	0.92	7.5	−7.3	0.68	0.87	8.3 4.8	−3.2
	20	2237	−0.20	0.75	11.6	−14.9	0.81	0.90		−0.6
	50	2296	−2.22	0.63	12.5	−24.7	0.01	0.86	2.7	6.3
	100	2277	−1.96	0.55	12.8	−22.5	–	–	–	–
2075	10	1827	0.71	0.90	9.0	5.9	0.74	0.94	6.3	−4.8
	20	1829	0.62	0.83	10.5	−2.8	0.46	0.92	6.7	4.5
	50	1823	0.60	0.78	11.5	−0.4	0.81	0.98	5.3	−4.1
	100	1829	−2.38	0.45	12.2	20.2	–	–	–	–
2077	10	2311	0.73	0.91	4.6	3.8	0.76	0.95	3.7	−3.9
	20	2311	0.67	0.83	6.0	−1.0	0.87	0.94	2.6	0.9
	50	2311	0.34	0.78	7.4	−0.1	0.82	0.94	2.0	1.1
	100	2309	−0.05	0.45	7.1	2.7	–	–	–	–
2078	10	1634	0.39	0.79	11.4	2.0	0.69	0.82	11.8	0.1
	20	1634	−0.42	0.83	8.8	−14.8	−0.82	0.93	0.1	12.3
	50	1634	0.01	0.42	9.1	−0.5	0.70	0.94	3.0	−3.0
	100	1632	0.44	0.67	9.7	−1.5	–	–	–	–
2179	10	1869	0.65	0.87	8.9	5.9	0.74	0.91	6.5	−4.4
	20	1869	0.44	0.83	10.2	7.6	0.47	0.88	5.3	−5.8
	50	1762	0.50	0.81	10.6	−4.4	0.67	0.96	3.9	3.3
	100	1836	−0.07	0.77	14.7	−8.0	–	–	–	–
2180	10	1214	0.44 –	0.90	4.1	−4.5	0.76	0.91	3.9	2.8
	20	1214	0.3	0.78	4.9	−7.1	0.68	0.90	3.3	2.5
	50	1214	−10.4	0.32	5.9	−9.4	0.71	0.97	1.7	−2.8
	100	1214	−14.8	0.05	6.7	−24.5	–	–	–	–
2182	10	1099	0.04	0.71	8.0 6.7	−2.2	0.50	0.75	7.2 4.9	−3.5
	20	1527	−0.14	0.79		−5.2	0.15	0.91		−8.9
	50	1533	−7.9	0.70	4.5	−34.2	−0.3	0.92	4.1	5.1
	100	1532	−15.9	0.72	11.4	−46.1	–	–	–	–
Mean	*10*	*1692*	*0.62*	*0.88*	*6.0 6.9*	*0.5*	*0.69*	*0.89*	*5.6 3.9*	*−2.2*
	20	*1827*	*0.23*	*0.83*		*−3.5*	*0.48*	*0.91*		*−0.7*
	50	*1822*	*−11.9*	*0.58*	*7.8*	*−10.3*	*0.50*	*0.93*	*3.0*	*0.4*
	100	*1827*	*−38.1*	*0.48*	*9.5*	*−19.0*	–	–	–	–

**N* may be different for each layer, depending upon the availability of SCAN observations.

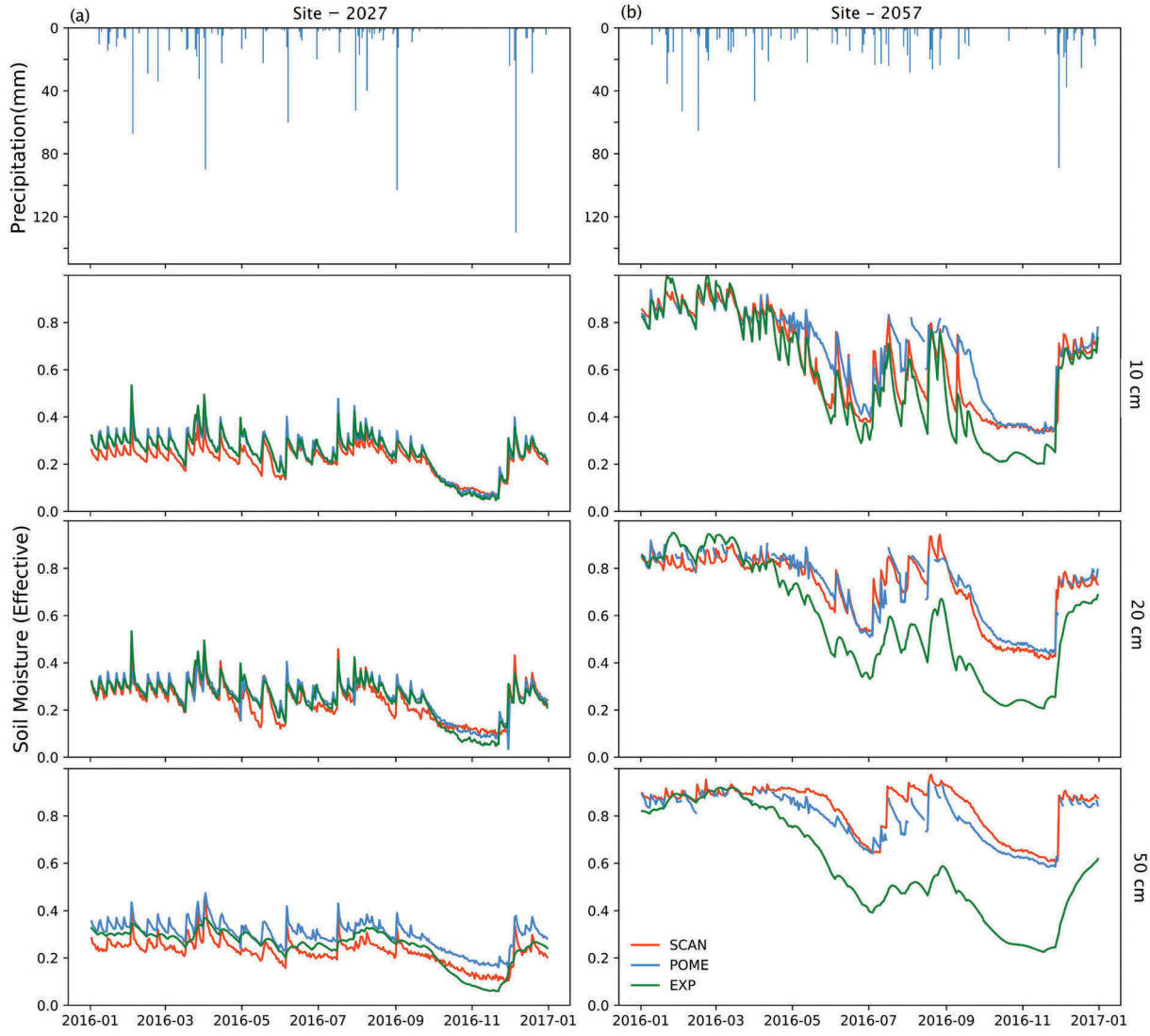

Figure 3. Time series of effective SM at different layer depths (10, 20 and 50 cm) for two sites: 2027 (left) and 2057 (right) along with precipitation data.

For site 2057, the EF actually performs better than POME in the second layer (10 cm) at least in two instances: first around May then around the October time frame. These time periods also seem to coincide with wet/dry transition periods (dynamic moisture profiles) and is most likely the result of the POME model overestimating the inflection point (see profile discussion below). The surface layer moisture fluctuates in the absence of rainfall due to the direct soil evaporation and surface transpiration; however, in the lowest layers (80+ cm), the SM is much less variable and remains rather high. This could be in part due to the presence of higher clay content in the lower layers compared to the surface. Clay soils with low hydraulic conductivity can hold moisture content much longer and are relatively less influenced by loamy surface fluctuations, effectively adding to the already compacted lower layers. In the lower layers the Exponential model begins to divert from the observed values. The EF tends to exaggerate during drying events (i.e. once the growing season starts in April/May) and it is propagated throughout the year. The underestimation during drying months is in tune with the surface moisture drifts (Fig. 3(b)). The POME model, on the other hand, driven by boundary conditions as well as mean moisture content, can better account for these changes with over or under estimation of mean moisture content input.

Overall, at shallow depths (i.e. 10 cm), one can argue that both models performed similarly in terms of time series across all sites. The difference in ubRMSE and bias between the models was less than 10%. This relationship holds for the top two layers in most sandy soils (likely true somewhat beyond the 20 cm mark; however, our analysis is limited to the SCAN observations depths). At 20 cm, loamy soils' difference in ubRMSE is roughly 30% higher. As depth increases the error in the EF model results increased (to 60% higher than the POME at 50 cm depths) indicating the limitations of the model to estimate lower layer moisture content solely from the surface observations. At shallower depths the relationship is more directly related to the surface (soil) evaporation while the lower layers are more related to the vegetative ET that the EF model does not capture effectively. The performance of the EF model tends to be the worst during summertime (March–September) when drying events

induced by lack of rainfall and ET (i.e. root-zone transpiration) decouple the two moisture pools (see Fig. 3(b)).

3.2 Vertical profile analysis

In addition to time series analysis, we further compare the performance of the two models with respect to their vertical SM profile distributions. Figure 4 shows selected SM profiles estimated by both the Exponential and POME models as compared to the SCAN observations. Although shown, surface and bottom layer (5 and 100 cm) data were not considered for this analysis since they were used as inputs. Figure 4 showcases the various types of profiles that are typically observed i.e. the wetting, drying and dynamic front as examples. Figure 4(a) represents a drying profile, the most common profile in non-irrigated systems where the surface soil is relatively dry compared to the lower layers, i.e. a typical stage in a prolonged period (hours or days depending on the soil type) after rainfall or irrigation. The profiles in Fig. 4(c) represents the dynamic case (with at least one prominent inflection point). Such a case can arise a short time after a rainfall event that only influences the surface and has not made it to the lower layers yet, or due to significant differences in soil physical characteristics between the layers. The wetting front that defines the dynamic case can eventually transition into a wetting profile or can revert back to a typical dry profile (in the case of a short rainfall event). Figure 4(b) is an interesting one, where SCAN SM profile shows an inflection point at 20 cm depth, yet the mean (and inflection point moisture content itself) was within the bounds of surface and bottom SM. Therefore, in the POME implementation it was treated as a dry case. Figure 4(d) represents a wetting profile (0–50 cm) that can occur during and just after rainfall events where the profile is progressively more wet toward the wetting front (surface). The four panels are generally representative of all the profiles generated in this comparison and represent three different SCAN sites.

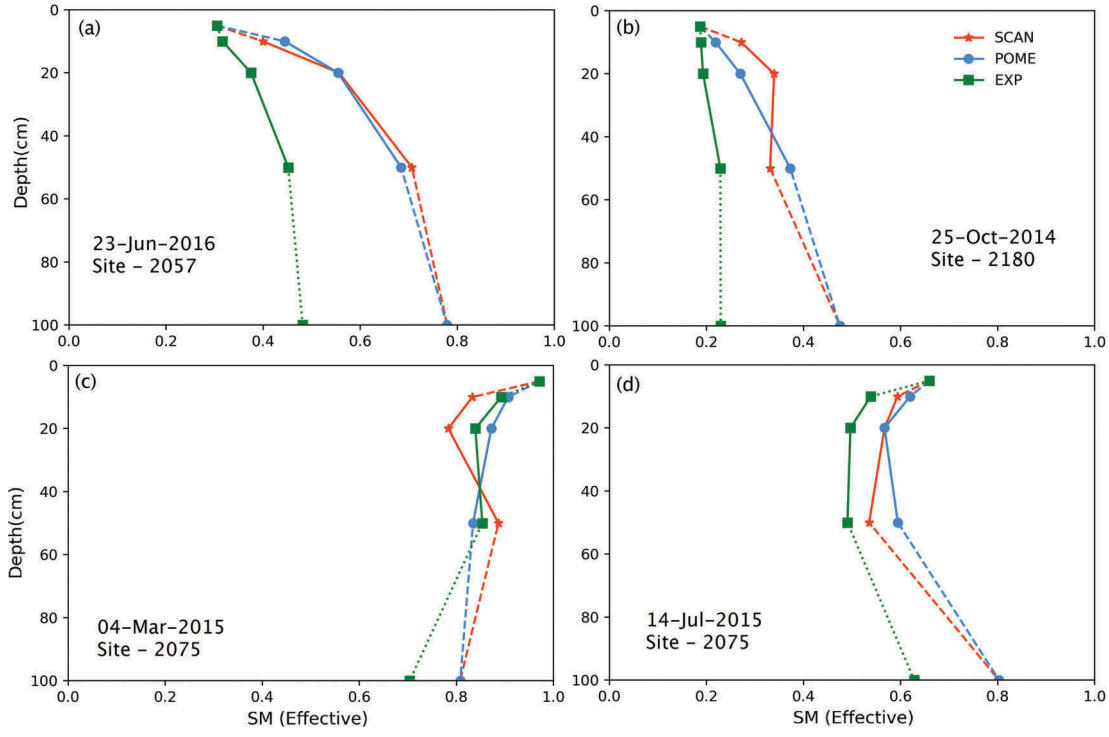

Figure 4. SM profiles for demonstrating the possible shapes for different cases: (a) a dry case; (b) another dry case profile, despite the fact that the SCAN clearly showed a prominent inflection point; (c) a dynamic case; and (d) a wet case. Surface and bottom-most points are not considered in statistics, therefore displayed in dashes.

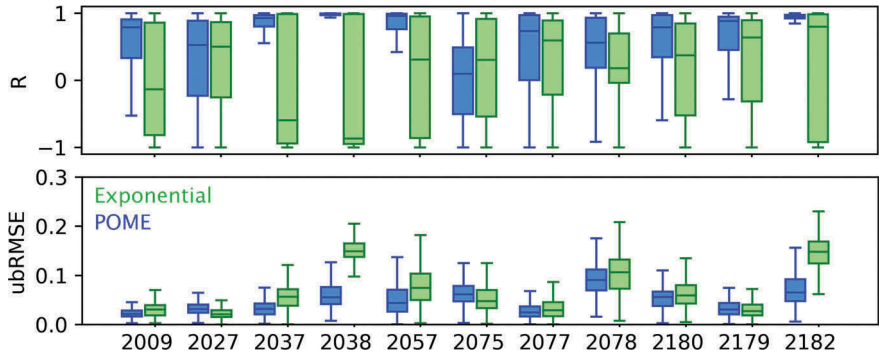

Figure 5. Boxplot of daily statistics of correlation and ubRMSE for each SCAN site.

The results in terms of correlation coefficient and ubRMSE of each profile analysis by each site is shown in Fig. 5. In terms of correlation, similar to the time-series results, the POME model performed better than the EF model for nearly all sites, with an average r of 0.58 compared to 0.14 for the EF model. Overall, the POME model ubRMSE was nearly 30% less than the EF model (0.05 *vs* 0.07), indicating that the POME model was better able to distribute moisture through the soil column compared to the SCAN observations. Due to the constraints of the SCAN observation depths, only three values per profile were available for statistical comparison (10, 20 and 50 cm depths); therefore, these results should be interpreted in light of this knowledge.

The 6 years' worth of data over 11 scan sites gives us a sample size of more than 16,000 profiles. Of all profiles analyzed, 63.4% represented the dry case, followed by the dynamic case profiles (19.3%) and rest of the profiles belonged to the wet case. Figure 6 summarizes the ubRMSE and bias for both model results by each profile type. It should be noted here that the results shown in Fig. 6 for wet and dynamic cases have a smaller sample size compared to the dry case profiles. The POME model results for ubRMSE seem to be relatively stable with a ± 5% (approx.) difference in mean ubRMSE for all three cases. Interestingly, the errors for the monotonic cases (dry and wet) were higher than the dynamic case profiles that were developed by parameterizing the inflection point location. This is in contrast with the finding of earlier studies by Al-Hamdan and Cruise (2010) and Mishra *et al.* (2015), where the monotonic cases had lower errors compared to the dynamic cases. This anomalous result can be due to wrong interpretation of dynamic case profiles as monotonic ones as shown in Fig. 4(b). In terms of bias, a similar trend was observed for all the profile cases. The use of the dynamic case, even with the parameterized inflection point location, seems to perform better in reducing the overall bias for the POME model. In comparison, the EF model seems to perform best under the wet profile scenario. This is consistent with the assumptions inherent in the development of the EF model and results of the prior studies discussed earlier. The dry and dynamic case bias for the EF model is negative indicating a relative underestimation of the moisture content compared to SCAN data. These underestimations can be attributed to the fact that the surface layer dries out much quicker over prolonged drying cycles (due to evaporation) and its effects trickle down to the lower layers, whereas in reality this relationship is essentially decoupled since the lower layers might still hold on to the moisture content. The variation in the EF model performance is higher compared the POME model signifying the importance of the profile type in model performance.

As mentioned earlier, in profiles that exhibited inflection points, the location of the inflection point was assumed to be at the 20 cm depth based on prior studies. However, in reality, the inflection point may develop at different depths at a given time depending on factors such as the soil characteristics, amount and intensity of the precipitation (or irrigation) event, the plant types which govern the root distribution as well as root density. Therefore, the dynamic case profile results are further studied to analyze the impact of inflection point selection over both model performance as well as to evaluate the assumptions made in the POME model implementation for such profiles (see Section 2.2). Figure 7 shows the ubRMSE distribution as a function of percent of dynamic cases for the 20 cm depth layer for both models for each site. As it can be seen, as the percent of dynamic cases increase, the ubRMSE tends to increase for the EF model (with R^2 of 0.7), while the POME model show a slight decline (R^2 of 0.1). Together Figs. 6 and 7, seem to suggest that the performance of the EF model deteriorates as the number of profiles with inflection points increases, whereas the POME model results, with its current mode of implementation, showed a decrease in ubRMSE. This highlights the significant difference between the EF and POME

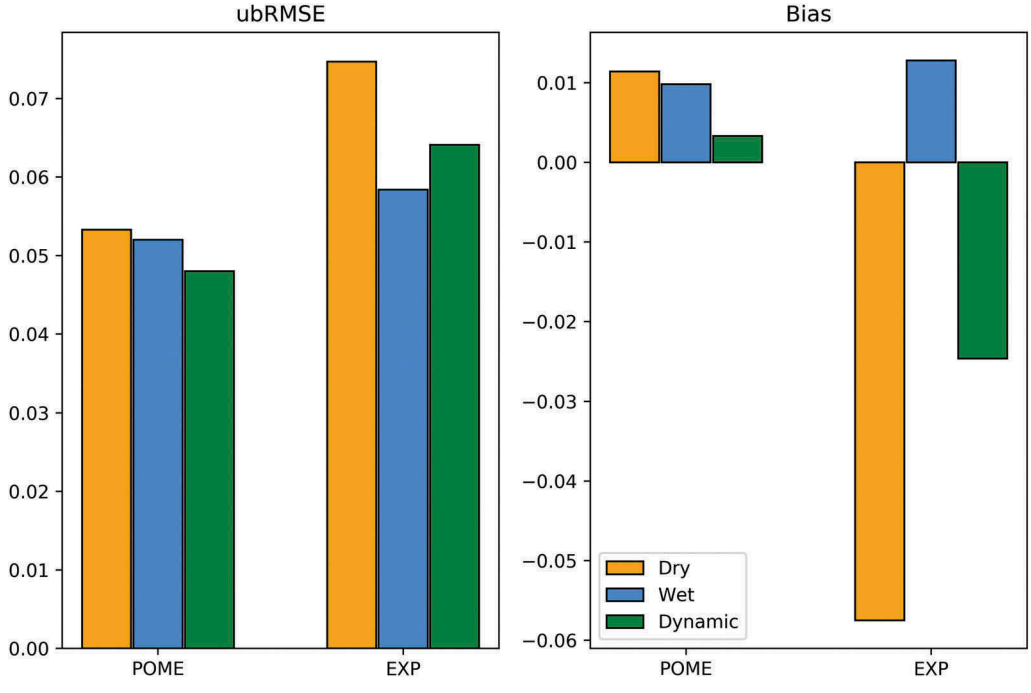

Figure 6. Daily profile statistics by the types: dry, wet and dynamic.

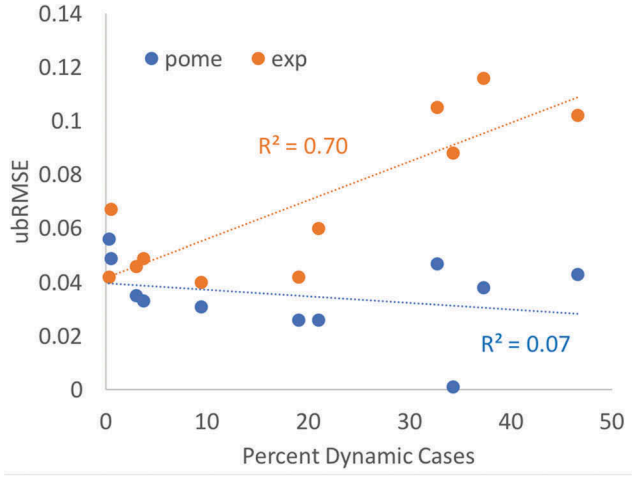

Figure 7. Scatterplot displaying the relation between ubRMSE and the number of profiles with inflection points (dynamic case) for both EF and the POME model for all sites.

models, while implying that the POME model can handle such cases effectively even with parameterized inputs.

This analysis also looked into potential relationships based on seasonality and sub-climate regimes across our study area. On average the performance of both the EF and POME models performed slightly better in the cooler, wetter years of the region (December–March), where wet case profiles are the majority, however, the differences are small and not significant. Overall, the annual spread in ubRMSE was <0.04 and <0.01, respectively, for both temporal and profile comparisons. Though our study is mostly representative of a humid, subtropical climate, there are regional climate differences across the area (i.e. annual rainfall rates vary over 400 mm). When looking at how each site performed with respect to its location, there are no clear trends by site or by location (e.g. latitude, coastal boundaries etc.). These results corroborate well with other studies that found no clear link between the performance of the EF model and climate and/or seasonal variations (Albergel *et al.* 2008, Tobin *et al.* 2017), and it makes sense that the POME results should follow.

Overall, it is observed that the POME model outperforms the EF model in most cases, and it can be seen that the EF model performs worst once drying events start to occur. Wetting fronts have a larger influence on T_{opt} where the relationship between the surface and lower layers is the strongest; hence, the EF model has relatively better characterizations of the vertical SM profiles in these cases. In the drying cases, the relationship of the lower layer SM is not closely (linearly) related to the surface. Once the upper layer(s) start to dry, the hydraulic connectivity with the subsurface is significantly reduced and the error appears to propagate throughout the time series. This explains that though the overall correlation is somewhat maintained, the mass balance significantly altered (≫ bias). This can also be explained, in part, by plant uptake (through ET) in the root-zone. The effect of ET on the profile (via root distributions) adds a bit more complexity to the moisture distribution that cannot be captured by the EF model (in the T_{opt} parameterization). The POME model, on the other hand, due to its reliance on mean moisture content, inclusion of an inflection point, and lower boundary conditions in addition to the surface, tends to perform more reliably throughout the depth. The daily mean moisture inputs take into account the ET and drainage processes unlike the EF model.

The results from the temporal and profile analysis suggest that the SM estimates using the POME model are generally better as compared to the EF model when analyzing against the SCAN observations. However, this is not always the case, and in some scenarios, the single-input EF model can result comparative performances.

3.3 Limitations and implications in large-scale applications

The POME and EF models are statistical estimations of the movement and redistribution of root-zone SM and have certain inherent limitations in terms of model structure, applicability, or input requirements, in addition to the configuration and implementations. For example, the dependence on a calibrated spin-up period can have a negative influence on the EF model performance across scales. T_{opt} is typically calibrated to a specific spin-up period and location and can impart biases on the T parameter used in the exponential filter. To address these biases, in part, the T can be optimized for different seasons; however, we found, as earlier studies did (Albergel *et al.* 2008), that there were no significant differences across seasons and, thus, it was not included in this study. Additionally, as discussed earlier, the inherent assumption of no loss (i.e. ET) limits the overall mass balance of the EF model. With respect to the POME model, the inclusion of an inflection point presents an added benefit of capturing dynamic profiles; however, it carries with it a potential source of error if not properly parameterized. In this study, the inflection point is assumed to be at 20 cm depth, but observations show that it may vary with soil types as well as temporally (Mishra *et al.* 2015).

When considering only the minimum inputs required, the EF model clearly is more applicable with only surface observations required. However, this does assume that there is a reasonable dataset available to calibrate T_{opt} and that there is a strong relationship between surface and subsurface layers. On the other hand, the POME model requires three inputs, surface SM, mean moisture, and lower boundary conditions. With respect to their use in large scale applications, both models' minimum inputs can be obtained from alternate sources. For instance, MW sensors can provide the surface observations required for both the models (Albergel *et al.* 2008, Ford *et al.* 2014, Peterson *et al.* 2016, Tobin *et al.* 2017, Mishra *et al.* 2018). The mean moisture content data required for the POME model can be obtained using TIR based moisture estimations (Mishra *et al.* 2013, 2018). The TIR observations can be used to deduce the root-zone mean moisture content indirectly using energy fluxes (Anderson *et al.* 2007, Hain *et al.* 2009). The lower boundary (100–200 cm) are usually well below the crop rooting depth and therefore are generally temporally less dynamic compared to shallow layers (Scott *et al.* 2003) and can either be parameterized or can act as calibration factor. Since one of the goals of developing SM profiles is to assimilate them back into LSMs/crop models for improved representation of the moisture content in data scarce regions, the calibration of the lower

boundary could be an effective approach to tie the remotely sensed profiles into the modeling framework.

Each of the added remote sensing datasets come with their own limitations and uncertainties, such as spatial resolution and sensing depths for MW sensors or cloud constraints for thermal observations (Schmugge *et al.* 1992, 2002, Norman *et al.* 1995). However, this study presented the raw uncertainties of the models themselves to better help understand the opportunities when used with remote sensing. Although it was expected that POME would outperform the exponential model generally, it is important to quantify this improvement given the additional information that must be gleaned from the remote sensing platforms. This study found that the EF model performs similarly to the POME in the upper soil layers (up to, and potentially beyond, 20 cm in sandy soils) in most cases. However, the EF model demonstrated a clear lack of ability (in term of bias and ubRMSE) to consistently estimate SM at the 50 cm mark and beyond. When concerned with SM profile distributions and depths in the lower root-zones the significantly better performance of the POME model is a compelling case for the added minimum inputs. This scenario could be the case in agricultural applications, where time-dependent root-zone SM can be critical for crops. However, in the case of broader land surface applications, where shallow (~0–10 cm) SM information is needed, the EF model may suffice.

4 Concluding remarks

The purpose of this study was to evaluate and compare the simulation of SM profiles using two models commonly used in data limited environment, the EF and POME models. These models typically use remote sensing datasets as inputs to simulate SM conditions. For this study, we used observation as inputs from 11 SCAN sites for 2012–2017 over the southeastern USA, allowing for comparison of the two models under optimum conditions. We found that the POME model generally outperformed the EF model at estimating SM profiles, especially in the lower layers of the soil column. However, it was found that both models had comparable performance at shallow depths (□10 cm) in terms of time series across all sites and potentially beyond 20 cm for sandy soils. However, as depth increased the error in the Exponential model results increased indicating the limitations of the model to estimate lower layer moisture content solely from the surface observations. This was particularly identified during prolonged drying events where the surface and root-zone moisture pools become uncoupled. The POME model, with three data inputs (SM boundary conditions and mean column SM), could be necessary when profile distribution and time specific root-zone moisture availability is important (i.e. agriculture applications) To summarize the implications of this study:

- Satellite remote sensing data can derive two out of three inputs needed for the POME model and one (and only) input(s) required for the exponential filter model.
- Remote sensing of surface SM estimates are more reliable than root-zone SM, making the inputs into the EF model more robust than the POME model.
- The T parameter for the EF is calibrated and not dynamic thus can be influenced by the calibration period chosen for T_{opt}.
- The EF model fails to distribute the moisture properly during dry conditions where the hydraulic connectivity between the surface and root-zone be577 come uncoupled and the resulting error is propagated in time.
- The EF model does not include transpiration in the SM estimates, whereas the POME model infers transpiration from the mean root-zone SM inputs and mass balance calculations.
- The POME model is scalable in space and depth because of the soil agnostic parameterizations, whereas the EF needs T_{opt} to be optimized for a particular soil column thus limiting use to other geographic areas and layers with calibration data.
- The POME model includes the specification of inflection points, i.e, dynamic SM profiles, although requiring the use of some assumptions. The EF does not consider inflection points and only assumes a relationship of the soil layer to the surface layer.

In applications where time dependent root-zone SM can be critical (i.e. agriculture), the significantly better performance of the POME model could justify the added minimum inputs required.

The results of this study are meant as a guide for the application of the models by users to better select an approach given the assumptions and data inputs. When using any model, it is important that users understand the limitations and best uses of the models. The comparison of the two models characterizes the limitations and caveats, guiding their further use for applications. Further research is needed to understand and compare the results of the EF and POME models using inputs from actual remote sensing data inputs rather than idealized inputs; however, the results of this study provide a starting point for practitioners in choosing the best available model for estimating SM profiles from remote sensing datasets.

Acknowledgements

We would like to thank Dr James F. Cruise, Professor Emeritus, Civil and Environmental Engineering Department, The University of Alabama in Huntsville for his guidance and valuable comments on the manuscript. The authors would also like to acknowledge and thank the NRCS for making soil moisture observation data publicly available. Lastly, we thank the anonymous reviewers for their invaluable comments that helped in improving the quality of the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Support for this work was provided through the joint US Agency for International Development (USAID) and National Aeronautics and Space Administration (NASA) initiative SERVIR, particularly through the NASA Applied Sciences Capacity Building Program, NASA Cooperative Agreement NNM11AA01A.

References

Aghakouchak, A., *et al*., 2015. Remote sensing of drought: progress, challenges and opportunities. *Reviews of Geophysics*, 53, 452–480. doi:10.1002/2014RG000456

Albergel, C., *et al*., 2008. From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations. *Hydrology and Earth System Sciences*, 12, 1323–1337. doi:10.5194/hess-12-1323-2008

Al-Hamdan, O.Z. and Cruise, J.F., 2010. Soil moisture profile development from surface observations by principle of maximum entropy. *Journal of Hydrologic Engineering*, 15 (5), 327–337. doi:10.1061/(ASCE)HE.1943-5584.0000196

Anderson, M.C., *et al*., 2012. Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. *Remote Sensing of Environment*, 122, 50–65. doi:10.1016/j.rse.2011.08.025

Anderson, M.C., *et al*., 2007. A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation. *Journal of Geophysical Research*, 112 (D10117). doi:10.1029/2006JD007506

Baldwin, D., *et al*., 2017. Predicting root zone soil moisture with soil properties and satellite near-surface moisture data across the conterminous United States. *Journal of Hydrology*, 546, 393–404. doi:10.1016/j.jhydrol.2017.01.020

Baldwin, D., *et al*., 2019. Estimating root zone soil moisture across the Eastern United States with passive microwave satellite data and a simple hydrologic model. *Remote Sensing*, 11 (17), 2013. doi:10.3390/rs11172013

Bolten, J.D. and Crow, W.T., 2012. Improved prediction of quasi-global vegetation conditions using remotely-sensed surface soil moisture. *Geophysical Research Letters*, 39 (19), L19406. doi:10.1029/2012GL053470

Capehart, W.J. and Carlson, T.N., 1997. Decoupling of surface and near-surface soil water content: A remote sensing perspective. *Water Resources Research*, 33 (6), 1383–1395. doi:10.1029/97WR00617

Carlson, T.N., Gillies, R.R., and Schmugge, T.J., 1995. An interpretation of methodologies for indirect measurement of soil water content. *Agricultural and Forest Meteorology*, 77 (3), 191–205. doi:10.1016/0168-1923(95)02261-U

Castillo, A., Castelli, F., and Entekhabi, D., 2015. An entropy-based measure of hydrologic complexity and its applications. *Water Resources Research*, 51, 5145–5160. doi:10.1002/2014WR016035

Ceballos, A., *et al*., 2005. Validation of ERS scatterometer-derived soil moisture data in the central part of the Duero Basin, Spain. *Hydrological Processes*, 19 (8), 1549–1566. doi:10.1002/hyp.5585

Cho, E., Choi, M., and Wagner, W., 2015. An assessment of remotely sensed surface and root zone soil moisture through active and passive sensors in northeast Asia. *Remote Sensing of Environment*, 160, 166–179. doi:10.1016/j.rse.2015.01.013

Crow, W.T., Kustas, W.P., and Prueger, J.H., 2008. Monitoring root-zone soil moisture through the assimilation of a thermal remote sensing-based soil moisture proxy into a water balance model. *Remote Sensing of Environment*, 112 (4), 1268–1281. doi:10.1016/j.rse.2006.11.033

Ellenburg, W.L., Cruise, J.F., and Singh, V.P., 2018. The role of evapotranspiration in streamflow modeling – an analysis using entropy. *Journal of Hydrology*, 567 (9), 290–304. doi:10.1016/j.jhydrol.2018.09.048

Entekhabi, D., *et al*., 2010. The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE*, 98 (5), 704–716. doi:10.1109/JPROC.2010.2043918

Ford, T.W., Harris, E., and Quiring, S.M., 2014. Estimating root zone soil moisture using near-surface observations from SMOS. *Hydrology and Earth System Sciences*, 18 (1), 139–154. doi:10.5194/hess-18-139-2014

Hain, C.R., Mecikalski, J.R., and Anderson, M.C., 2009. Retrieval of an available water-based soil moisture proxy from thermal infrared remote sensing. *Part I: Methodology and Validation*, 10, 665–683. doi:10.1175/2008JHM1024.1

Kerr, Y.H., *et al*., 2010. The SMOS mission: new tool for monitoring key elements of the global water cycle. *Proceedings of the IEEE*, 98 (5), 666–687. doi:10.1109/JPROC.2010.2043032

Kornelsen, K.C. and Coulibaly, P., 2014. Root-zone soil moisture estimation using data driven methods. *Water Resources Research*, 50, 2946–2962. doi:10.1002/2013WR014127

Kostov, K.G. and Jackson, T.J., 1993. Estimating profile soil moisture from surface-layer measurements: a review. *In*: H.N. Nasr, ed. *Optical engineering and photonics in aerospace sensing*. International Society for Optics and Photonics, 125–136. doi:10.1117/12.154681

Lettenmaier, D.P., *et al*., 2015. Inroads of remote sensing into hydrologic science during the WRR era. *Water Resources Research*, 51, 7309–7342. doi:10.1002/2015WR017616

Lievens, H., *et al*., 2015. SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia. *Remote Sensing of Environment*, 168, 146–162. doi:10.1016/j.rse.2015.06.025

Mahmood, R. and Hubbard, K.G., 2007. Relationship between soil moisture of near surface and multiple depths of the root zone under heterogeneous land uses and varying hydroclimatic conditions. *Hydrological Processes*, 21 (25), 3449–3462. doi:10.1002/hyp.6578

Mays, D.C., Faybishenko, B.A., and Finsterle, S., 2002. Information entropy to measure temporal and spatial complexity of unsaturated flow in heterogeneous media. *Water Resources Research*, 38 (12), 1313. doi:10.1029/2001WR001185

McKinney, W., 2010. Data structures for statistical computing in python. *In*: S. van der Walt and J. Millman eds. *Proceedings of the 9th python in science conference*. Austin, Texas: SciPy, 51–56.

Millman, K.J. and Aivazis, M., 2011. Python for scientists and engineers. *Computing in Science Engineering*, 13 (2), 9–12. doi:10.1109/MCSE.2011.36

Mishra, V., *et al*., 2018. Development of soil moisture profiles through coupled microwave-thermal infrared observations in the southeastern United States. *Hydrology and Earth System Sciences*, 22, 4935–4957. doi:10.5194/hess-22-4935-2018

Mishra, V., *et al*., 2013. A remote-sensing driven tool for estimating crop stress and yields. *Remote Sensing*, 5 (7), 3331–3356. doi:10.3390/rs5073331

Mishra, V., *et al*., 2015. Modeling soil moisture profiles in irrigated fields by the principle of maximum entropy. *Entropy*, 17 (6), 4454–4484. doi:10.3390/e17064454

Njoku, E.G., *et al*., 2003. Soil moisture retrieval from AMSR-E. *IEEE Transactions on Geoscience and Remote Sensing*, 41 (2), 215–229. doi:10.1109/TGRS.2002.808243

Norman, J., Kustas, W., and Humes, K., 1995. Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. *Agricultural and Forest Meteorology*, 77 (3–4), 263–293. doi:10.1016/0168-1923(95)02265-Y

Oliphant, T.E., 2007. Python for scientific computing. *Computing in Science Engineering*, 9 (3), 10–20. doi:10.1109/MCSE.2007.58

Pachepsky, Y., Rawls, W., and Lin, H., 2006. Hydropedology and pedotransfer funtions. *Geoderma*, 131 (3–4), 308–316. doi:10.1016/j.geoderma.2005.03.012

Pan, X., Kornelsen, K.C., and Coulibaly, P., 2017. Estimating root zone soil moisture at continental scale using neural networks. *JAWRA Journal of the American Water Resources Association*, 53 (1), 220–237. doi:10.1111/1752-1688.12491

Peterson, A.M., Helgason, W.D., and Ireson, A.M., 2016. Estimating field-scale root zone soil moisture using the cosmic-ray neutron probe. *Hydrology and Earth System Sciences*, 20 (4), 1373–1385. doi:10.5194/hess-20-1373-2016

Qin, M., Gim´enez, D., and Miskewitz, R., 2018. Temporal dynamics of subsurface soil water content estimated from surface measurement using wavelet transform. *Journal of Hydrology*, 563, 834–850. doi:10.1016/j.jhydrol.2018.06.023

Qiu, J., Crow, W.T., and Nearing, G.S., 2014. The impact of vertical measurement depth on the information content of soil moisture times series data. *Geophysical Research Letters*, 41, 4997–5004. doi:10.1002/2014GL060017

Reichle, R.H., 2005. Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. *Geophysical Research Letters*, 32 (2), L02404. doi:10.1029/2004GL021700

Ridler, M.-E., *et al*., 2014. Assimilation of SMOS-derived soil moisture in a fully integrated hydrological and soil vegetation-atmosphere transfer model in Western Denmark. *Water Resources Research*, 50 (11), 8962–8981. doi:10.1002/2014WR015392

Robock, A., *et al*., 1995. Use of midlatitude soil moisture and meteorological observations to validate soil moisture simulations with biosphere and bucket models. *Journal of Climate*, 8, 15–35. doi:10.1175/1520-0442(1995)008<0015:UOMSMA>2.0.CO;2

Sadeghi, M., *et al*., 2017. Advancing NASA's AirMOSS p-band radar root zone soil moisture retrieval algorithm via incorporation of richards' equation. *Remote Sensing*, 9 (1), 17. doi:10.3390/rs9010017

Schmugge, T., *et al*., 1992. Passive microwave remote sensing of soil moisture: results from HAPEX, FIFE and MONSOON 90. *ISPRS Journal of Photogrammetry and Remote Sensing*, 47 (2–3), 127–143. doi:10.1016/0924-2716(92)90029-9

Schmugge, T.J., *et al*., 2002. Remote sensing in hydrology. *Advances in Water Resources*, 25 (8–12), 1367–1385. doi:10.1016/S0309-1708(02)00065-9

Scott, C.A., Bastiaanssen, W.G.M., and Ahmad, M., 2003. Mapping root zone soil moisture using remotely sensed optical imagery. *Journal of Irrigation and Drainage Engineering*, 129 (5), 326–335. doi:10.1061/(ASCE)0733-9437(2003)129:5(326)

Shannon, C.E., 1948. A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x

Singh, V., 1997. The use of entropy in hydrology and water resources. *Hydrological Processes*, 11 (6), 587–626. doi:10.1002/(SICI)1099-1085-(199705)11:6<587::AIDHYP479>3.0.CO;2-P

Singh, V.P., 2010. Entropy theory for movement of moisture in soils. *Water Resources Research*, 46, W03516. doi:10.1029/2009WR008288

Srivastava, S., *et al*., 1997. On the relationship between ERS-1 SAR/backscatter and surface/sub-surface soil moisture variations in vertisols. *Acta Astronautica*, 40 (10), 693–699. doi:10.1016/S0094-5765(97)00125-2

Starks, P.J., *et al*., 2003. Use of limited soil property data and modeling to estimate root zone soil water content. *Journal of Hydrology*, 272 (1–4), 131–147. doi:10.1016/S0022-1694(02)00260-3

Sun, Y. and Niu, J., 2019. Regionalization of daily soil moisture dynamics using wavelet-based multiscale entropy and principal component analysis. *Entropy*, 21 (6), 548–560. doi:10.3390/e21060548

Sure, A. and Dikshit, O., 2019. Estimation of root zone soil moisture using passive microwave remote sensing: A case study for rice and wheat crops for three states in the Indo-Gangetic basin. *Journal of Environmental Management*, 234, 75–89. doi:10.1016/j.jenvman.2018.12.109

Tabatabaeenejad, A., *et al*., 2015. P-band radar retrieval of subsurface soil moisture profile as a second-order polynomial: first AirMOSS results. *IEEE Transactions on Geoscience and Remote Sensing*, 53 (2), 645–658. doi:10.1109/TGRS.2014.2326839

Tobin, K.J., *et al*., 2017. Multi-decadal analysis of root-zone soil moisture applying the exponential filter across CONUS. *Hydrology and Earth System Sciences*, 21 (9), 4403–4417. doi:10.5194/hess-21-4403-2017

van Dam, J. and Feddes, R., 2000. Numerical simulation of infiltration, evaporation an shallow groundwater levels with the richards equation. *Journal of Hydrology*, 233 (1), 72–85. doi:10.1016/S0022-1694(00)00227-4

van der Walt, S., Colbert, S.C., and Varoquaux, G., 2011. The numpy array: A structure for efficient numerical computation. *Computing in Science Engineering*, 13 (2), 22–30. doi:10.1109/MCSE.2011.37

Wagner, W., Lemoine, G., and Rott, H., 1999. A method for estimating soil moisture from ERS Scatterometer and soil data. *Remote Sensing of Environment*, 70 (2), 191–207. doi:10.1016/S0034-4257(99)00036-X

Wang, T., *et al*., 2017. Evaluating controls of soil properties and climatic conditions on the use of an exponential filter for converting near surface to root zone soil moisture contents. *Journal of Hydrology*, 548, 683–696. doi:10.1016/j.jhydrol.2017.03.055

Yang, K., *et al*., 2016. Land surface model calibration through microwave data assimilation for improving soil moisture simulations. *Journal of Hydrology*, 533, 266–276. doi:10.1016/j.jhydrol.2015.12.018

Zhou, X., Lei, W., and Ma, J., 2016. Entropy base estimation of moisture content of the top 10-m unsaturated soil for the Badain Jaran Desert in northwestern China. *Entropy*, 18 (9), 323–338. doi:10.3390/e18090323