Soil Moisture Model Calibration and Validation: An ARS Watershed on the South Fork Iowa River

EVAN J. COOPERSMITH AND MICHAEL H. COSH
Hydrology and Remote Sensing Laboratory, Agricultural Research Service, USDA, Beltsville, Maryland

WALT A. PETERSEN
NASA Wallops Space Flight Facility, Wallops, Virginia

JOHN PRUEGER
National Laboratory for Agriculture and Environment, Agricultural Research Service, USDA, Ames, Iowa

JAMES J. NIEMEIER
IIHR–Hydroscience and Engineering, The University of Iowa, Iowa City, Iowa

(Manuscript received 25 July 2014, in final form 5 February 2015)

ABSTRACT

Soil moisture monitoring with in situ technology is a time-consuming and costly endeavor for which a method of increasing the resolution of spatial estimates across in situ networks is necessary. Using a simple hydrologic model, the estimation capacity of an in situ watershed network can be increased beyond the station distribution by using available precipitation, soil, and topographic information. A study site was selected on the Iowa River, characterized by homogeneous soil and topographic features, reducing the variables to precipitation only. Using 10-km precipitation estimates from the North American Land Data Assimilation System (NLDAS) for 2013, high-resolution estimates of surface soil moisture were generated in coordination with an in situ network, which was deployed as part of the Iowa Flood Studies (IFloodS). A simple, bucket model for soil moisture at each in situ sensor was calibrated using four precipitation products and subsequently validated at both the sensor for which it was calibrated and other proximal sensors, the latter after a bias correction step. Average RMSE values of 0.031 and 0.045 m$^3$ m$^{-2}$ were obtained for models validated at the sensor for which they were calibrated and at other nearby sensors, respectively.

1. Introduction

Soil moisture estimates are valuable for numerous agricultural and hydrologic applications. At the point scale, estimates of field wetness for enhanced agricultural decision support enable real-time irrigation scheduling in drier climates (Rao et al. 1988) and real-time estimates of potential field trafficability in more humid locations (Coopersmith et al. 2014b). At larger scales, soil moisture is an integral component of hydrologic storage and subsurface flows at the watershed scale (e.g., Grayson et al. 1997). At watershed scales, an increasingly important use for large-scale estimates is satellite remote sensing calibration and validation (Jackson et al. 2010, 2012). These satellites are able to estimate soil moisture on the scale of 25–50 km, using passive microwave remote sensing. Active microwave (radar) remote sensing will be able to decrease the measurement scale to 3 km (Entekhabi et al. 2010). Few watershed networks are able to monitor surface soil moisture adequately at both the large (>25 km) and small (~3 km) scale. In addition, in situ instruments are currently able to measure soil moisture on the scale of 1 m (Ochsner et al. 2013). For this reason, there is significant benefit in developing the relevant framework.

Corresponding author address: Evan Coopersmith, Hydrology and Remote Sensing Laboratory, USDA-ARS, 10300 Baltimore Ave., Bldg. 007, Beltsville, MD 20705.
E-mail: evan.coopersmith@ars.usda.gov

DOI: 10.1175/JHM-D-14-0145.1

© 2015 American Meteorological Society
for upscaling point-scale estimates from sparse networks to averages over a much larger, predefined space.

Remotely sensed estimates of soil moisture have been employed in conjunction with hydrologic models as initial conditions and to update the results generated (e.g., Lin et al. 1994). Some of these hydrologic models adopt grid-based structures to facilitate distribution of model results over a larger watershed (e.g., Mas et al. 1995). More recent works have integrated hydrologic water balance models forced by precipitation and energy flux measured at in situ gauges to produce watershed-scale estimates (Stillman et al. 2014). These analyses are illustrative, but ultimately, the cost of in situ sensors impedes the extension of this approach to the watershed-scale estimates where such sensors are not available. Additionally, recent work analyzing in situ networks has been deployed to validate satellite observations, comparing satellite estimates from the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) and other satellite estimates to the Soil Climate Analysis Network (SCAN) and the U.S. Department of Agriculture (USDA) Agricultural Research Service (ARS) watersheds (Reichle et al. 2011; Liu et al. 2011). Other works compare soil moisture estimates from satellites to those generated by land surface models (e.g., Loew et al. 2013).

Another use of high-resolution soil moisture estimates would be for satellite validation. NASA’s Soil Moisture Active Passive (SMAP) mission, launched on 31 January 2015, will produce soil moisture estimates at 3-, 9-, and 36-km scales. These estimates must be validated against in situ measurements, which are made via in situ watershed and sparse networks. This requires a mathematical framework for transforming soil moisture measurements by sparse networks of in situ sensors into representative area estimates. By leveraging public sources of precipitation data in concert with models calibrated at local, in situ sensors, soil moisture estimates can be upcaled to provide estimates at the desired spatial scales.

To upscale the in situ measurements to these management and satellite validation scales, it is necessary to develop models that can be applied at a variety of scales, given a set of land surface parameters, meteorological data, and an in situ network time series. Land surface parameters that impact surface soil moisture include topography, such as slope and aspect, soil type, and vegetation cover (Jacobs et al. 2004; Joshi and Mohanty 2010). Among meteorological variables, precipitation plays the primary role in forcing soil moisture. Precipitation data can be obtained from a variety of sources, including the NEXRAD network, whose derived products include NEXRAD’s Stage IV precipitation product and the North American Land Data Assimilation System (NLDAS; Mitchell et al. 2004) derived, “quasi-operational” precipitation values rather than directly observed estimates. For this study, it is advantageous to reduce the number of variables influencing soil moisture, as each parameter or variable adds complexity to the modeling structure. This approach requires that, at each location, we employ a model calibrated at a similar location. With every additional dimension employed to define “similar,” it becomes exponentially more difficult to locate such similarity. To this end, a relatively flat, agricultural landscape with homogeneous soil would be the ideal location to host this modeling study.

2. Methodology

a. Problem overview

To achieve the stated goal of upscaling soil moisture estimates, this analysis adopts a distributed modeling approach at the 500-m scale. Figure 1 presents the South Fork watershed, which can be divided into 500 m × 500 m (0.25 km²) regions, each of which will produce, from its interpolated precipitation time series, a soil moisture estimate. In turn, these squares can be upscaled to produce estimates at the 3- (36 squares), 9- (324 squares), or 36-km (5184 squares) scales, which in turn can be compared against satellite estimates.

This paper focuses its analysis on a Long-Term Agroecosystem Research site in the South Fork Iowa River in central Iowa. The ARS monitors this test watershed with 15 in situ soil moisture and precipitation stations, each of which provides hourly estimates of soil moisture profiles and precipitation from dual gauges. This network was founded to provide ground reference data for the NASA SMAP mission and the Global Precipitation Measurement (GPM) mission. Five additional stations in the domain (labeled 16–20 in the subsequent analysis) are operated by NASA GPM in coordination with the University of Iowa and consist of four Campbell Scientific 655 soil water content reflectometers at depths of 5, 10, 20, and 50 cm and two MetOne 380 precipitation gauges. This domain was the focus of two recent remote sensing campaigns, Soil Moisture Experiments in 2002 and 2005 (SMEX02 and SMEX05; e.g., Bindlish et al. 2006). In addition, this site hosted the Iowa Flood Studies (1FloodS), as a part of the GPM mission’s Ground Validation (GV) program (Petersen and Krajewski 2013) during the months of April–June 2013 in coordination
b. Site description

The South Fork of the Iowa River is characterized by several features that allow useful simplifying assumptions for the purposes of upscaling the in situ measurements from the sparse network of soil moisture sensors. First, this region is primarily homogeneous from a topographic perspective, which will be verified more rigorously in mathematical terms, described within this section. Second, this watershed displays homogeneity in hydroclimatic terms, as the entire test site falls within a temperate, water-limited northern plain characterized by the typical Midwest seasonal precipitation pattern, displaying maximum rainfall and runoff during summer (growing season) months (Coopersmith et al. 2012). These characteristics do not vary over the 36-km square in which this analysis occurs. Finally, this watershed is also considered homogeneous in edaphic terms, characterized by predominately mollisol, loam/silty clay loam with installed tile to compensate for otherwise poor drainage (see www.jecam.org/?/site-description/usa-iowa). Images depicting the soils of Iowa can be found in Figs. A1 and A2. These anthropogenic features compromise the utility of many more complex hydrologic models (e.g., Ye et al. 2012). For these reasons, this work assumes the primary source of spatial variability of soil moisture to be spatial variability of precipitation.

c. Precipitation products

In addition to the in situ gauges themselves, precipitation data are also available through NEXRAD’s Stage IV data product and through NLDAS. NLDAS data used in this analysis were chosen from the ascending (A) series. Each of these precipitation products are gridded to spatial scales more coarse than those necessitated by a 500-m grid. Therefore, via simple spatial interpolation techniques, precipitation values can be subsequently deployed as point estimates at individual sensor locations and elsewhere within the 500-m grid as needed. This paper will explore these precipitation products in more detail to determine which performs optimally for the purposes of soil moisture model scaling. While the primary objective of this analysis is soil moisture modeling and the development

FIG. 1. Illustration of the South Fork watershed with in situ stations labeled. The background image illustrates the clay percentage within the top 10 cm of soil. For the study region, soil textures fall between 23% and 28% clay, with most stations presenting a value of 27%. On the left, an image of a representative soil moisture station is included.
of estimates at the desired 500-m (and ultimately 3 km) scales, some preliminary analysis will reveal which precipitation product is best suited for this task. Unfortunately, with only 1 year of data to consider, with in situ gauges with sporadically missing precipitation values, and similar gaps with respect to NEXRAD data, a more detailed analysis of the biases between these products is not practical. This question is left for future research with richer datasets.

Four precipitation products were examined for use in the larger-scale estimation process. The first was the precipitation gauges themselves, located atop the soil moisture sensors placed throughout the watershed. The second, NEXRAD’s Stage IV data product, is derived from ground-based radar, available anywhere within the continental United States. Each of these first two products, unfortunately, suffers from the issue of missing data or occasional failure to report precipitation. Thus, a third precipitation product was created by selecting the maximum reported precipitation between the gauge and its NEXRAD Stage IV estimate. While this inflates the total quantity of precipitation observed, the model parameter responsible for rate of drainage will adjust accordingly. Finally, the fourth precipitation product interpolated estimates from NLDAS, which, though less granular spatially, are robust to issues of missing data.

d. Selection of valid sensors

The diagnostic soil moisture equation (Pan et al. 2003; Pan 2012) was calibrated at each of the 18 valid sensors using each of the four precipitation products. The resulting performance was then used to determine which precipitation product was suitable for broader precipitation estimation over the larger spatial areas required for SMAP.

Figure 2 presents a generalized flowchart of the process by which each individual sensor (marked locations in Fig. 1) is evaluated for its potential to be used to help estimate soil moisture values in a given box. Coopersmith et al. (2014a) demonstrate that for a soil moisture model to be applicable at a location other than the one for which it is calibrated, the new location must display hydroclimatic and edaphic similarity with respect to the original site. Additionally, topographic similarity is also important. For the type of analysis suggested, an a priori assumption is made that any two points within a watershed of the size in question (the box chosen by this analysis is 1880 km² in area) are hydroclimatically similar. The uppermost two diamonds assess the topographic and edaphic similarity between a given 0.25 km² box and the location of the sensor in question. The answers to these questions are always “yes” within the South Fork watershed. Finally, the lowermost diamond determines whether the model calibrated at a given sensor has performed well enough to warrant inclusion in subsequent estimates at other locations. For this purpose, the threshold is a linear correlation coefficient (Pearson’s $r$) greater than 0.9 between in situ measurements and modeled estimates. Correlation is selected, rather that root-mean-square error (RMSE), as a constant bias can be removed (diminishing RMSE), but modeling the shape of the wetting/drying time series accurately (evidenced by high correlations between measurements and modeled estimates) is crucial to larger-scale estimations.

e. Topography

With respect to topographic similarity, the assumption was verified by an analysis of all points within a digital elevation map, at the 20-m scale, of the relevant area. At each coordinate $(x, y, \text{elevation})$, the following four coordinates were gathered: $(x+500 \text{ m}, y, \text{elevation}1)$, $(x+500 \text{ m}, y, \text{elevation}2)$, $(x-500 \text{ m}, y, \text{elevation}3)$, and $(x-500 \text{ m}, y, \text{elevation}4)$. A 2% minimum grade was established for the elevation difference to be considered nontrivial, that is, an elevation change of 10 m over a 500-m horizontal distance. Each of the four coordinates listed above (the points to the north, south, east, and west) were either labeled as a 0 (at least 10 m above point $x, y$), a 1 (within 10 m vertically of point $x, y$), or a 2 (at least 10 m below point $x, y$). Using these four indicators allows for any location to be classified, at the 500-m scale, as a peak (e.g., 2222), slope (e.g., 2020), flat (e.g., 1111), or trough (e.g., 0000). The results, over all
points examined, are shown in Table 1. The results from Table 1 verify that the overwhelming majority of the landscape is flat, even with only a very mild grade required for classification as a slope, peak, or trough. Even many of the locations classified as slopes are quite flat, displaying changes in elevation of 20–25 m over 1 km. Many of the locations characterized as peaks or pits are not produced at scales finer than 500 m, and the landscape is flat, even with only a very mild grade required for classification as a slope, peak, or trough. Even many of the locations characterized as slopes are quite flat, displaying changes in elevation of 20–25 m over 1 km. Many of the locations classified as slopes are quite flat, displaying changes in elevation of 20–25 m over 1 km.

### Table 1. Results of topographic classification.

<table>
<thead>
<tr>
<th>Land surface</th>
<th>Area percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Troughs</td>
<td>1.06%</td>
</tr>
<tr>
<td>Flats</td>
<td>93.82%</td>
</tr>
<tr>
<td>Slopes</td>
<td>4.08%</td>
</tr>
<tr>
<td>Peaks</td>
<td>1.05%</td>
</tr>
</tbody>
</table>

f. Soil moisture modeling

For the purposes of this analysis, the diagnostic soil moisture equation (Pan et al. 2003; Pan 2012) was deployed to calibrate a lumped, surface soil moisture model (depth = 5 cm) at all 20 sensors within the South Fork watershed. The diagnostic soil moisture equation is well suited for such work, as it requires only a precipitation time series and a soil moisture time series at the location for which it is calibrated. At the site at which it is deployed, the parameters calibrated elsewhere can be applied using only a local precipitation time series, without the need for antecedent soil moisture information. Equations (1) and (2) present the diagnostic soil moisture equation from Pan (2012):

\[
\theta_{\text{est}} = \theta_{\text{re}} + (\phi_e - \theta_{\text{re}})(1 - e^{-c_4\beta}) \tag{1}
\]

and

\[
\beta = \sum_{i=2}^{n-1} \left[ \frac{P}{\eta_i} (1 - e^{-\eta_i/z}) e^{-\sum_{i=1}^{j=i-1} (\eta_i/z)} + \frac{P}{\eta_i} (1 - e^{-\eta_i/z}) \right] 
\]

In Eq. (1), \( \theta_{\text{re}} \) denotes residual soil moisture, the value below which the soil moisture levels will not fall even after the prolonged absence of rainfall; \( \phi_e \) represents the soil’s porosity, the maximum soil moisture value attainable; and \( c_4 \) controls the rate of drainage, the rate at which the soil moisture values can decrease with time. In Eq. (2), \( i/j \) denote hourly time stamps, an adaptation of the daily framework of the original model. The parameter \( n \) in Eq. (2) represents the system’s memory, that is, how many hours historically must be considered in generating a soil moisture estimate. It is set by generating a \( \beta \) series for which \( n \) is very large, then generating a \( \beta \) series with increasing values of \( n \) until the values of those \( \beta \) series converge. The values of \( \eta_i \) form a time series describing losses due to deep drainage, surface runoff, and evapotranspiration, a sinusoid with an annual period.

The full diagnostic soil moisture equation [Eq. (1)], including the \( \beta \) series [Eq. (2)], requires the calibration of six parameters. The first three define the sinusoidal shape of the \( \eta \) series. As \( \eta \) is defined as a sinusoid with a period of 1 year, this requires fitting three parameters (a generalized sinusoid has four, but in this case, the period is already known) to define its vertical shift, its horizontal shift, and its amplitude. These were fit via a genetic algorithm, maximizing the correlation of the \( \beta \) series (using the various sinusoidal-shaped \( \eta \) series generated via the genetic algorithm) with the empirically measured values of \( \theta \). The chosen historical window is 200–250 h at sensors in this watershed. Next, the three remaining parameters from Eq. (1), the porosity of the soil (\( \phi_e \)), the residual soil moisture (\( \theta_{\text{re}} \)), and the rate of drainage (\( c_4 \)), were fit via a second genetic algorithm using the \( \beta \) series (which contains the optimal \( \eta \) series). At each stage, the genetic algorithms are real-coded implementations containing the processes of generation, selection/crossover, mutation, and death, evolving increasingly successful parameter sets until a termination criterion is reached (in this case, 25 and 100 generations without observing a superior solution for the first and second algorithms, respectively). Ultimately, genetic algorithms allow a more computationally efficient exploration of the parametric search space than systematic search or the Monte Carlo approach adopted by Pan (2012), especially given the increased computational demands of an hourly model (rather than the original, daily version). Genetic algorithms have been used to calibrate the diagnostic soil moisture equation’s parameters in Coopersmith et al. (2014a).

At each of the sensors, data begin shortly after installation in early April 2013 and conclude in early October 2013, before freezing and thawing begin to compromise the efficacy of the models employed. Approximately 200 days (4800 h) of precipitation and soil moisture time series are available for the calibration of
This estimate is a function of each sensor’s model. For this reason, models are validated during the same temporal range during which they are calibrated. An example of one such calibration appears in Fig. 3.

Two of the 20 sensors were removed from consideration, regardless of the correlation values obtained when the models were calibrated. The first, sensor 1, experienced significant flooding, reporting soil moisture values above 0.8 m$^3$ m$^{-3}$—figures in excess of the soil’s true porosity and evidence of invalid measurement. A model calibrated at this sensor location would be inapplicable elsewhere. The second, sensor 10, experienced extreme moisture values at the opposite end of the physical range of values, displaying values below 0.03 m$^3$ m$^{-3}$ by volume—below the minimum plausible values at these locations, and, we speculate, the result of damage from gophers and collisions with tractors. The remaining 18 sensors, each of which returned its own set of calibrated parameters, can be considered for application at any 0.25-km$^2$ box within the watershed. The performances of these individual calibrations are presented in the next section, which discusses results, and are ultimately summarized in Figs. 5 and 6. The values of the calibrated parameters are available in the appendix. Equation (3) generalizes the procedure for estimation of soil moisture within a given box $x, y$ at time $t$:

$$
\theta_{x,y,t} = \sum_{i=1}^{n} \left[ \theta_{i,t*,s} - \theta_{i,t,m}(P_{x,y,t}) + \theta_{i,t*,m}(P_{x,y,t}) \right] \times \frac{1/(d_{x,y})^2}{\sum_{j=1}^{n} 1/(d_{x,y})^2},
$$

(3)

where $\theta_{i,t*,s}$ denotes the soil moisture measurement at sensor $i$ at the most recent hour for which a valid reading was recorded ($t^*$), measured in situ ($s$), and $\theta_{i,t,m}$ refers to the soil moisture model estimate for sensor $i$ at hour $t$ using the model $m$ parameters calibrated for that sensor. This estimate is a function of $P_{x,y,t}$, the precipitation product at $x_i, y_i$ (the coordinates of sensor $i$), at hour $t$. The variable $\theta_{i,t,m}$ signifies the soil moisture model $m$ estimate for sensor $i$ at hour $t$ as a function of $P_{x,y,t}$, the precipitation at site $x, y$ at hour $t$. These estimates are weighted by inverse-square distances ($d_{x,y}$)$^2$, the distance from $x, y$ to the location of sensor $i$. Therefore, the estimate for a given box is the inverse distance weighted average of the in situ estimates, added to differences between the soil moisture models’ predictions via the precipitation measured at the box and the precipitation measured at the gauge. Given the assumption that the only reason that two locations should present different soil moisture values is the difference in the precipitation they experience (true if topography, soil, and hydroclimates are held constant), this enables robust soil moisture estimates at any location within the test watershed.

### g. Estimating model performance

One important question to address is the accuracy of each of these models when applied at sites other than those for which their parameters are calibrated. This was achieved by applying calibrated parameters at a sensor other than the one at which they are calibrated. As our approach calls for information from the most proximal sensor to be weighted most heavily, it is illustrative to examine the performance of a set of calibrated parameters at the closest sensor (apart from the calibration site itself):

$$
\theta_{t,t,est} = \theta_{t,t,m}(P_{x,y,t}) + \epsilon.
$$

(4)

In Eq. (4), $\theta_{t,t,est}$ is the soil moisture value at sensor $i$ at time $t$ as an estimate (est) generated by employing $\theta_{t,t,m}$, the soil moisture at site $i$ at time $t$, using model $m$, via the parameters calibrated at site $j$, the closest sensor spatially to sensor $i$. This modeled estimate is a function of the chosen precipitation product at sensor $i$ ($P_{x,y,t}$).

Some of the discrepancies between soil moisture sensors are simply noise in the calibrations of the sensors themselves. To wit, one sensor may simply report 0.03–0.04 m$^3$ m$^{-3}$ more soil moisture content than another. In
In this case, the results of deploying one sensor’s parameters at its closest neighbor should be evaluated after an appropriate bias correction is introduced:

$$\theta_{i,j,\text{est}} = \theta_{i,j,m,j}(P_{x_i,y_i,t}) + \mu_i - \mu_j + \epsilon.$$  (5)

In Eq. (5), an analogous idea is presented to the approach of Eq. (4), only in this case, the estimated soil moisture is adjusted by the difference in mean values reported by the in situ sensors at the two relevant locations. This bias correction allows for a more appropriate assessment of the value of cross application of calibrated parameters.

It is important to recognize that even these levels of error are conservative. As shown in Eq. (3), every sensor has an opportunity to contribute information to the predictions made at every box within the test watershed. Thus, the individual biases displayed between sensors are likely to be randomly distributed, centered at zero. Consequently, when multiple sensors render estimates at a given box, these multiple sensory estimates see their introduction of noise mitigated by sample size.

### h. Scaling

Equation (3) facilitates the generation of an estimate of average volumetric soil moisture within any chosen 0.25 km$^2$ box, labeled by its center $x$, $y$. In the case of the test site in question, a 47 km $\times$ 40 km area, there are 7520 boxes, each 500 m $\times$ 500 m. Every individual box produces an estimate of soil moisture, as described in Eq. (3). From here, estimating the average soil moisture content over a box of user-specified size was accomplished, as shown in Eq. (6):

$$\Theta_t = \frac{\sum_k \frac{A_{x_k,y_k}}{A_S} \theta_{x_k,y_k,t}}{n},$$  (6)

where $\Theta_t$ denotes the average soil moisture value over the user-defined, larger surface at time $t$, with area $A_S$, as defined by the estimates over the $k$, 0.25-km$^2$ boxes defined by coordinates $x_k$, $y_k$.

### i. Geospatial characteristics

Figure 4 presents a semivariogram for the South Fork watershed’s soil moisture sensors that meet the calibration standards during an afternoon in early August.
These 11 sensors provide some insight into the characteristic length within this test site. Given the shorter east–west dimension (≈40 km) of this test site, as opposed to its longer north–south dimension (≈47 km), pairs of sensors separated by distances exceeding 20 km are not considered. In these cases, longer pairs would characterize the north-to-south spatial variability disproportionately. In Fig. 4, the “sill” appears to occur between 5 and 6 km. Though the precise characteristic length at the South Fork site is not explicitly calculated, it seems clear that this paper’s analysis, performed at the 500-m scale, reflects smaller spatial scales than those of the soil moisture variability at this location as well as the relevant scale of topography. This would seem to be confirmed by the findings of Cosh and Brutsaert (1999), who discovered a characteristic length on the order of 1 km in the Little Washita test site in Oklahoma—a site characterized by considerably more spatial variability in terms of soil type and field usage. Furthermore, after choosing the 11 “best” sensors, the average distance from any sensor to its nearest neighbor is 6.4 km. Thus, in assessing the performance of cross validation over this distance scale, one can infer the overall efficacy of this modeling approach across the test site as a whole.

3. Results

The six parameters, discussed in Eqs. (1) and (2), were fit to calibrate the model at each of the 18 valid, in situ gauges located within the South Fork watershed (see the appendix). Each of the 18 soil moisture time series was calibrated using four separate precipitation products. The results of this analysis are shown in Figs. 5 and 6. Sensors 1 and 10 were excluded, as discussed previously, because of complications from flooding, gopher damage, and tractor collisions. Time stamps during which rain is observed within the previous 4 h were excluded, as temporary ponding immediately following rain events can skew sensor measurements. Moreover, the delay between the arrival of precipitation and its infiltration to the depth of the sensor requires such a conservative approach.

Figure 7 presents the aggregate results from Figs. 5 and 6. The NLDAS precipitation data offer not only the strongest model performances in terms of RMSE, but...
Table 2. Results, including closest sensors, cross application of parameters, and bias correction. The best 11 sensors are shown in boldface in the first column. Sensors from NASA GPM are labeled IFC.

<table>
<thead>
<tr>
<th>Sensor number</th>
<th>Original RMSE (using NLDAS)</th>
<th>RMSE (of best 11, using NLDAS)</th>
<th>Closest sensor (min Euclidean distance)</th>
<th>RMSE using closest sensor’s parameters and local NLDAS precipitation</th>
<th>RMSE using closest sensor’s parameters and local NLDAS precipitation, bias corrected</th>
<th>Closest (of the best 11)</th>
<th>RMSE using closest of best 11 sensors’ parameters and local NLDAS precipitation, bias corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.036</td>
<td>0.036</td>
<td>3</td>
<td>0.065</td>
<td>0.053</td>
<td>3</td>
<td>0.053</td>
</tr>
<tr>
<td>3</td>
<td>0.050</td>
<td>0.050</td>
<td>2</td>
<td>0.071</td>
<td>0.067</td>
<td>2</td>
<td>0.067</td>
</tr>
<tr>
<td>4</td>
<td>0.027</td>
<td>0.027</td>
<td>7</td>
<td>0.063</td>
<td>0.039</td>
<td>2</td>
<td>0.028</td>
</tr>
<tr>
<td>5</td>
<td>0.032</td>
<td>0.032</td>
<td>6</td>
<td>0.044</td>
<td>0.059</td>
<td>3</td>
<td>0.055</td>
</tr>
<tr>
<td>6</td>
<td>0.046</td>
<td>—</td>
<td>5</td>
<td>0.052</td>
<td>0.059</td>
<td>5</td>
<td>0.059</td>
</tr>
<tr>
<td>7</td>
<td>0.029</td>
<td>—</td>
<td>8</td>
<td>0.047</td>
<td>0.035</td>
<td>4</td>
<td>0.045</td>
</tr>
<tr>
<td>8</td>
<td>0.033</td>
<td>—</td>
<td>7</td>
<td>0.050</td>
<td>0.035</td>
<td>9</td>
<td>0.040</td>
</tr>
<tr>
<td>9</td>
<td>0.031</td>
<td>0.031</td>
<td>8</td>
<td>0.065</td>
<td>0.041</td>
<td>17</td>
<td>0.032</td>
</tr>
<tr>
<td>11</td>
<td>0.038</td>
<td>0.038</td>
<td>13</td>
<td>0.055</td>
<td>0.050</td>
<td>9</td>
<td>0.038</td>
</tr>
<tr>
<td>12</td>
<td>0.056</td>
<td>—</td>
<td>14</td>
<td>0.072</td>
<td>0.069</td>
<td>11</td>
<td>0.060</td>
</tr>
<tr>
<td>13</td>
<td>0.029</td>
<td>—</td>
<td>11</td>
<td>0.051</td>
<td>0.034</td>
<td>11</td>
<td>0.034</td>
</tr>
<tr>
<td>14</td>
<td>0.038</td>
<td>—</td>
<td>12</td>
<td>0.059</td>
<td>0.059</td>
<td>20</td>
<td>0.041</td>
</tr>
<tr>
<td>15</td>
<td>0.064</td>
<td>—</td>
<td>20</td>
<td>0.077</td>
<td>0.077</td>
<td>20</td>
<td>0.077</td>
</tr>
<tr>
<td>16 (IFC)</td>
<td>0.023</td>
<td>0.023</td>
<td>18</td>
<td>0.029</td>
<td>0.028</td>
<td>18</td>
<td>0.028</td>
</tr>
<tr>
<td>17 (IFC)</td>
<td>0.016</td>
<td>0.016</td>
<td>6</td>
<td>0.037</td>
<td>0.057</td>
<td>9</td>
<td>0.020</td>
</tr>
<tr>
<td>18 (IFC)</td>
<td>0.033</td>
<td>0.033</td>
<td>16</td>
<td>0.043</td>
<td>0.042</td>
<td>16</td>
<td>0.042</td>
</tr>
<tr>
<td>19 (IFC)</td>
<td>0.022</td>
<td>0.022</td>
<td>7</td>
<td>0.036</td>
<td>0.020</td>
<td>4</td>
<td>0.051</td>
</tr>
<tr>
<td>20 (IFC)</td>
<td>0.030</td>
<td>0.030</td>
<td>15</td>
<td>0.057</td>
<td>0.055</td>
<td>11</td>
<td>0.042</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0351</td>
<td>0.0307</td>
<td>—</td>
<td>0.0541</td>
<td>0.0487</td>
<td>—</td>
<td>0.0451</td>
</tr>
</tbody>
</table>
also the most ubiquitous source of data. Correlation had been the criterion of choice when assessing a model’s viability for use in cross validation (applying calibrated parameters at a new location) because bias correction can occur during the cross-validation process if one sensor consistently reports higher values than another. However, when selecting a precipitation product for calibration of a soil moisture model’s parameters, RMSE is the chosen criterion because no further bias correction is possible, as after those parameters are set, we already have produced mean-zero errors by definition. When predictions are required at individual boxes without gauges therein, local precipitation must either be made via interpolation/extrapolation of gauge data (difficult beyond the convex hull of the network) or via a product that is available more broadly. NEXRAD data are widely available, but NLDAS data yield slightly stronger model calibrations. Thus, NLDAS data will be employed for all subsequent analyses presented in this section.

a. Cross application

The results of this analysis are shown in Table 2. In Table 2, we observe that, by using the nearest sensor’s model parameters, performance decreases (fifth column versus second column). Compared with the average RMSE for models calibrated with NLDAS data at the sensors themselves (0.0351 m$^3$/m$^2$; Fig. 7, right), the RMSE values here are 0.019 m$^3$/m$^2$ higher. This suggests that, when producing a watershed-scale average, the overall RMSE is likely to fall between these two extremes.

As discussed in the methodology section, the results of calibrating a model at one location and applying it elsewhere can be mitigated, in part, by the bias correction [Eq. (5)] that adjusts for the differing calibrations between two sensor locations. The results of this analysis are shown in the sixth column of Table 2. By correcting for these biases, the average RMSE value decreases from 0.0541 m$^3$/m$^2$ (fifth column) to 0.0487 m$^3$/m$^2$ (sixth column).

The flowchart presented in Fig. 2 states, in its lowermost diamond, that only those sensors with sufficiently high-performing models will be considered. The discussed threshold was a linear correlation coefficient $r$ of 0.9 or greater. In this vein, there are only 11 sensor locations to be considered (2, 3, 4, 5, 9, 11, 16, 17, 18, 19, and 20), shown in boldface in the first column of Table 2. Thus, we can repeat the analyses presented in the sixth column, using only the best 11 sensors as candidates for the “closest” sensor chosen for cross application.

The third column, an RMSE of 0.0307 m$^3$/m$^2$ as compared with the 0.0351 m$^3$/m$^2$ RMSE average (second column) over all 18 sensors, is an improvement. This is unsurprising, as we have specifically selected the 11 calibrated models with the highest correlations between modeled estimates and in situ measurements. However, the eighth column, an RMSE of 0.0451 m$^3$/m$^2$ over all 18 sensors, is also an improvement over the corresponding value of 0.0487 m$^3$/m$^2$ from the sixth column. This improvement occurs despite the fact that the average Euclidean distance between each sensor and its closest neighbor who is a member of the best 11 (seventh column) is, by definition, larger than the average Euclidean distance between each sensor and its nearest neighbor over the full 18 (fourth column). Stated differently, making estimates at all locations using only the best 11 sensors is an improvement on using all 18.

Table 2 suggests that, for any point within the watershed, if that location is very close to one of the best 11 sensors, performances of 0.0307 m$^3$/m$^2$ can be expected, and if said location is farther away, the associated error will be 0.0451 m$^3$/m$^2$. The results are summarized in Fig. 8.

b. Scaling

The set of images in Fig. 9 presents spatial estimates of soil moisture estimates at various temporal snapshots during the summer of 2013. In each duo of images, on the left, soil moisture estimates from the distributed modeling approach outlined in Eq. (3) were performed at every 500 m x 500 m (0.25 km$^2$) square and colored accordingly. On the right, these same individual squares are aggregated into 6 x 6 squares of 3000 m x 3000 m to produce the analogous, less granular images. While the 500-m-scale images occasionally display localized circular patterns near sensors whose values deviate slightly from its neighbors, the 3000-m aggregation mitigates these impacts in many cases.

Figure 9 (top) presents the wet soil after a spring storm, where soil moisture levels are in excess of...
0.30 m$^3$ m$^{-3}$ throughout the watershed and over 0.40 m$^3$ m$^{-3}$ in several locations. Figure 9 (middle) illustrates mixed conditions in mid-July, during a drydown period, in which some sensors are largely dry, while others still have retained a significant portion of their moisture. Finally, Fig. 9 (bottom) presents paired images during an extremely dry period late in the summer, after an extended period without precipitation.

Among the advantages of this simple lumped model, distributed over a 500-m grid, is the ease with which these estimates can be transformed into larger-scale average estimates to be used subsequently to validate larger-scale satellite estimates. It is important to note that prior to the launch of the SMAP satellite on 31 January 2015, there was not an in situ data-driven soil moisture product at the 3-km scale with which to compare these upscaled estimates. Once a remotely sensed product from a satellite is developed, these larger estimates can be more rigorously validated.

c. Comparison with spatial interpolation

Having generated 3-km estimates using a model that integrates precipitation data and the values provided from in situ gauges, it is appropriate to demonstrate the
value of such an approach over simple spatial interpolation between sensors. Figures 10a and 10b present two sets of spatial estimates, separated by 1 h during August 2013. In the southeastern corner of the grid, a localized rain event has occurred. Figures 10a and 10b present the results of a model that estimates soil moisture exclusively via a spatial interpolation, in this case, weighting all sensors via their inverse-square distance to the desired location. In this case, the southeastern-most triad of green sensors (those deemed among the 11 best) does not reflect the precipitation event that has arrived. As a result, the average soil moisture value within the same dotted region remains relatively unchanged (from 0.143 to 0.147 m$^3$ m$^{-3}$). In contrast, Figs. 10c and 10d present the same time steps, but the results integrate the modeled output that integrates precipitation data. A storm has wet the soil in the dotted region, but it has not yet dropped precipitation over any of the network’s in situ sensors. By employing the diagnostic soil moisture equation to produce estimates at those locations that reflect this new precipitation, the average soil moisture value estimated within the 9 km $\times$ 9 km region within the dotted box increases from 0.155 to 0.201 m$^3$ m$^{-3}$. In this case, fully exploiting both the available in situ resources and precipitation estimates accounts for a difference of roughly 0.05 m$^3$ m$^{-3}$ that a simple spatial interpolation would overlook. Other available model approaches that employ local precipitation data would respond to such a localized rain event, but would not utilize the in situ resources that such a test site offers.

4. Discussion

The diagnostic soil moisture equation, though an appropriate selection for this type of modeling,
room for improvement through the incorporation of some of the features it chooses to omit. By utilizing additional information, it is possible for the general approach outlined in the previous sections to produce even more nuanced (and hopefully accurate) estimates of soil moisture.

First, the $\eta$ series, modeling the processes of evapotranspiration and losses due to deep drainage, was presumed to be sinusoidal over a period of 1 year. This was a suitable assumption for a model that was originally designed for daily usage. However, for an hourly model, this simplification can be improved upon given the diurnal cycle to which all soil moisture estimates are subjected. A simple, diurnal cycle was superimposed upon the “eta series” presented first in the methodology section. The specific implementation involved a sinusoid whose peak occurs midway between the listed times of sunrise and sunset (timeanddate.com) and whose nadir occurs at all times between sunset and sunrise (a minimum value is assigned for all results below this threshold). While the model’s lack of improvement from this feature is fully explained by only 2 of the 18 sensor locations (the remaining 16 display a small but significant improvement), it does not seem appropriate to remove an additional two sensors simply to accommodate an additional layer of model complexity.

Second, the $\eta$ series, in subsequent iterations, could be improved by the incorporation of humidity estimates with respect to potential evapotranspiration. Moreover, with the ubiquity of weather stations throughout the country made available through the National Climatic Data Center, one can envision subsequent research modeling soil moisture using a much more detailed $\eta$ series than the straightforward version presented in this paper. While the computational expense would be markedly greater along with increased data demands, the benefits could conceivably allow for soil moisture estimates that more precisely track impacts from storm events.

The South Fork watershed represented an ideal location to investigate this approach, as any two locations can safely be assumed to be similar in topographic, edaphic, and hydroclimatic terms. While most watersheds where such an approach might be applied for the purposes of the SMAP mission will be hydroclimatically similar at the 36-km scale, there are many in which topography and soil texture may differ significantly from one location in the test site to another. In these

Table A1. Calibrated parameters from diagnostic soil moisture equation. Boldface rows denote the best 11 sensors, from which the spatial estimates are produced.

<table>
<thead>
<tr>
<th>Site</th>
<th>System memory $n$</th>
<th>$\eta$ series values</th>
<th>Diagnostic soil moisture equation parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Vertical shift</td>
<td>Amplitude</td>
</tr>
<tr>
<td>SF02</td>
<td>250</td>
<td>0.042</td>
<td>0.038</td>
</tr>
<tr>
<td>SF03</td>
<td>250</td>
<td>0.005</td>
<td>0.023</td>
</tr>
<tr>
<td>SF04</td>
<td>250</td>
<td>0.049</td>
<td>0.029</td>
</tr>
<tr>
<td>SF05</td>
<td>250</td>
<td>0.048</td>
<td>0.004</td>
</tr>
<tr>
<td>SF06</td>
<td>250</td>
<td>0.049</td>
<td>0.046</td>
</tr>
<tr>
<td>SF07</td>
<td>250</td>
<td>0.013</td>
<td>0.023</td>
</tr>
<tr>
<td>SF08</td>
<td>250</td>
<td>0.043</td>
<td>0.041</td>
</tr>
<tr>
<td>SF09</td>
<td>250</td>
<td>0.016</td>
<td>0.025</td>
</tr>
<tr>
<td>SF11</td>
<td>200</td>
<td>0.013</td>
<td>0.030</td>
</tr>
<tr>
<td>SF12</td>
<td>200</td>
<td>0.023</td>
<td>0.038</td>
</tr>
<tr>
<td>SF13</td>
<td>200</td>
<td>0.020</td>
<td>0.020</td>
</tr>
<tr>
<td>SF14</td>
<td>200</td>
<td>0.045</td>
<td>0.045</td>
</tr>
<tr>
<td>SF15</td>
<td>200</td>
<td>0.034</td>
<td>0.03</td>
</tr>
<tr>
<td>SF16</td>
<td>200</td>
<td>0.023</td>
<td>0.027</td>
</tr>
<tr>
<td>SF17</td>
<td>250</td>
<td>0.018</td>
<td>0.038</td>
</tr>
<tr>
<td>SF18</td>
<td>200</td>
<td>0.028</td>
<td>0.035</td>
</tr>
<tr>
<td>SF19</td>
<td>250</td>
<td>0.032</td>
<td>0.032</td>
</tr>
<tr>
<td>SF20</td>
<td>200</td>
<td>0.041</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Fig. A1. U.S. Soil Orders Map (www.nrcs.usda.gov/Internet/FSE_MEDIA/stelprdb1237749.pdf). Dark green soils are mollisols. The approximate location of the South Fork sensors is indicated by the black circle.
locations, rather than permit any site within the watershed to offer an inverse-square distance weighted estimate for each box, we can permit estimates only from sites deemed to be topographically similar to the predominant topographic composition of the box in question and sites that are also of similar soil textural properties. In this manner, provided that every “box” within the watershed has at least one sensor that is similar both topographically and edaphically, appropriate predictions can be delivered.

5. Conclusions

Ultimately, the purpose of this analysis was to demonstrate a method for flexibly scaled soil moisture estimates. The South Fork test watershed represents an ideal location to investigate this approach, as the area can be safely assumed to be homogeneous topographically, edaphically, and hydroclimatically. By choosing the best-performing precipitation product and evaluating variable soil moisture in terms of precipitation variability only, estimates can be produced at any point within the test watershed and subsequently aggregated to produce larger-scale estimates.

Using NLDAS precipitation data and subsequently selecting the sensors for which the calibrated models performed best, models applied at the sites for which they were calibrated produced RMSE values of 0.0307 m$^3$ m$^{-3}$, and even when all 18 sensors’ soil moisture values are estimated using the nearest of those eleven sensors, RMSE values of 0.0451 m$^3$ m$^{-3}$ were obtained. This suggests that at a random point located within this watershed, the accuracy of the estimate will likely fall between 0.0307 and 0.0451 m$^3$ m$^{-3}$, as a function of proximity to one of these high-performing 11 sensors. As a watershed with a large range of soil moisture values (minimums of 0.07–0.08 m$^3$ m$^{-3}$, maximums in excess of 0.5 m$^3$ m$^{-3}$), it is reasonable to expect that in watersheds with lower soil moisture variability, the RMSE values would be correspondingly lower.

Moving forward, these techniques can be applied to a wider array of watersheds (those characterized by nontrivial topography and nontrivial soil texture variability) and over larger spatial scales (at which hydroclimatology is no longer a constant). More immediately, such scaling methodology facilitates the usage of test watersheds for calibration and validation of SMAP estimates following its launch on 31 January 2015.

Acknowledgments. USDA is an equal opportunity provider and employer. This work is supported by the NASA Terrestrial Hydrology Program (NNH10ZDA001N-THP).

APPENDIX

Calibrated Parameters and Soil Classifications

The values obtained for the diagnostic soil moisture equation’s parameters of each of the 20 sensors’ calibrations appear in Table A1. The best-performing 11 sensors are presented in boldface. Sensors from NASA GPM are listed as IFC. Soil classifications and textural information are illustrated in Figs. A1 and A2.
REFERENCES


