Downscaling Satellite-Based Soil Moisture in Heterogeneous Regions Using High-Resolution Remote Sensing Products and Information Theory: A Synthetic Study

Subit Chakrabarti, *Student Member, IEEE*, Tara Bongiovanni, Jasmeet Judge, *Senior Member, IEEE*, Karthik Nagarajan, and Jose C. Principe, *Fellow, IEEE*

***Abstract*—In this study, a novel methodology based upon the information-theoretic measures of entropy and mutual information was implemented to downscale soil moisture (SM) observations from 10 km to 1 km. It included a transformation function that related auxiliary remotely sensed (RS) products at high resolution to *in situ* SM observations to obtain first estimates of SM at 1 km and merging this estimate with SM at coarse resolutions through Principle of Relevant Information (PRI). The PRI-based estimates were evaluated using synthetic observations in NC Florida for heterogeneous agricultural land covers (LC), with two growing seasons of sweet corn and one of cotton, annually. The cumulative density function showed an overall error in SM of < 0.03 cubic meter/cubic meter in the region, with a confidence interval of 95% during the simulation period. The PRI estimates at 1 km were also compared with those from the method based upon Universal Triangle (UT). The spatially averaged root mean square error (RMSE) aggregated over the vegetative LC were 0.01 cubic meter/cubic meter and 0.15 cubic meter/cubic meter using the PRI and UT methods, respectively. The RMSE for downscaled estimates using the UT method increased to 0.28 cubic meter/cubic meter when Laplacian errors are used, while the corresponding RMSE for the PRI remains the same for both Laplacian or Gaussian errors. The Kullback–Liebler divergence (KLD) for estimates using PRI is about 50% lower than those using the method based upon UT indicating that the probability density function (PDF) of the PRI estimate is closer to PDF of the true SM, than the UT method.**

***Index Terms*—Downscaling, entropy, microwave brightness (MB) temperature, mutual information, Observation System Simulation Experiment (OSSE), Soil Moisture Active Passive (SMAP), Soil Moisture and Ocean Salinity (SMOS).**

Manuscript received September 20, 2013; revised December 30, 2013 and January 27, 2014; accepted February 12, 2014. Date of publication July 24, 2014; date of current version August 4, 2014. This work was supported by the National Aeronautics and Space Administration-Terrestrial Hydrology Program under Grant NNX09AK29G and Grant NNX13AD04G.

S. Chakrabarti, T. Bongiovanni, J. Judge, and K. Nagarajan are with the Center for Remote Sensing, Agricultural and Biological Engineering Department, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL 32611-0570 USA (e-mail: subitc@ufl.edu).

J. C. Principe is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6200 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2014.2318699

I. Introduction

HYDROMETEOROLOGICAL models simulate atmospheric and hydrological processes at scales of 1–10 km that are significantly influenced by local and regional distribution of soil moisture (SM) [1]. Microwave observations at frequencies < 10 GHz are highly sensitive to changes in near-surface moisture and have been widely used to retrieve SM information [2]–[11]. While satellite-based SM products from microwave observations are available at spatial resolutions of a few meters, with temporal resolutions of several weeks, passive observations are obtained at coarser resolutions of tens of kilometers but with finer temporal resolutions of subdaily to 2–3 days. The recently launched European Space Agency-Soil Moisture and Ocean Salinity (SMOS) [12] and the near-future National Aeronautics and Space Administration (NASA)-Soil Moisture Active Passive (SMAP) [13] missions have provided SM at unprecedented spatial resolutions of 10–25 km. For these observations to be relevant to hydrometeorological and agricultural studies, the SM values need to be downscaled to several kilometers. This downscaling of SM is also essential to understand the effects of land surface heterogeneity on aggregated microwave signature and to quantify SM under dynamic vegetation conditions.

Typically, SM retrieved from satellite products at coarser scales is spatially downscaled under the assumption that it follows a known hierarchical model, such as power-law [14] or fractal law models [15], [16] that are based upon temporal persistence of SM across scales [1]. Distributed hydrology models have been also used to investigate scaling characteristics of SM at different spatial resolutions [17], [18]. Most of these approaches assume static vegetation conditions and/or spatial homogeneity. Studies involving heterogeneous and dynamic land-cover (LC) conditions are necessary to evaluate the accuracy of downscaling algorithms for realistic applications.

Some studies use other remote sensing (RS) products, such as leaf area index (LAI), land surface temperatures (LSTs), [19] and LC [20] observed at a finer resolution in conjunction with microwave observations at a coarser resolution to retrieve SM at 1 km. In these methods, the high-resolution RS products are either used in land surface models (LSMs) to directly estimate SM or are upscaled to match the resolution of microwave observations [15], [19], [21]. The LSMs require *in situ* or other data

for calibration and validation. *In situ* data are sparsely available at regional scales and in remote regions. Empirical models based upon statistical methods such as regression [4], [22], [23], multiscale tree [24], and geostatistical methods, such as cokriging and block kriging, have been used to extract SM information from diverse RS products. Another class of techniques employs the triangle method [25], [26] using statistical models to extrapolate the dependent data within a hypothetical triangle/trapezoid formed by the observed data, [27]–[30]. Kim and Hogue [30] used the triangle method with a derived soil wetness index at a scale of 1 km as a scaling factor and found that the accuracy of downscaled SM was a strong function of the spatial correlation of LST and soil evaporative efficiency to SM. Merlin *et al.* [27] proposed the DISPATCH algorithm that calculated soil evaporative efficiency [31] as a function of normalized difference vegetation index (NDVI) and LST and used the triangular relationship between the efficiency and SM at the two scales to obtain high-resolution SM estimates. This method was validated using data from the Murrumbidgee River catchment area in Southern Australia during the National Airborne Field Experiment '06 [32]. An improved version of the DISPATCH algorithm [28] used a linear or nonlinear (depending on the application scale) evaporative efficiency relationship and a correction for topographic effects. Piles *et al.* [29] developed a linking model from the triangle formed by the observed data at coarse scale and applied the model to data at finer scale, assuming equivalence at both scales. This was validated using data from three days during a four-week period in January and February 2010 from an experiment in the MurremBridge River catchment area. Lack of NDVI observations from the Moderate Resolution Imaging Spectroradiometer (MODIS) limited their validation study to just three days. Although the annual precipitation (PPT) and LCs in the area were heterogeneous, the validation limited the diversity. The robustness of these methods over different vegetation and PPT conditions remains largely unknown.

All the aforementioned methods, as outlined in [33], use second-order statistics as their optimization tool. Scaling algorithms based upon second-order statistics can potentially lead to significant loss of structural information in the data [34], particularly under highly nonlinear heterogeneous and dynamic conditions. The Bayesian hierarchical framework [35] utilizes information from the probability density functions (PDFs) but is computationally intensive and still relies heavily on Gaussian assumptions. Computationally efficient nonparametric algorithms based upon data descriptors that retain information from higher order moments allow for optimal downscaling, particularly under heterogeneous and dynamic land surface conditions. Information-theoretic learning is a framework where the conventional concepts of second-order statistics are substituted by scalars and functions based upon information theory, such as entropy and mutual information [36], [37]. The PRI [38], which is based upon the measures of entropy and mutual information, provides a hierarchical decomposition of spatial data that is optimal in terms of the transfer of information across spatial scales and may be a better alternative to methods that use second-order statistics. Statistical data descriptors beyond the second-order moment are necessary to capture the complete stochastic structure of data and to preserve maximum information to achieve optimal downscaling, particularly over heterogeneous and dynamic environments.

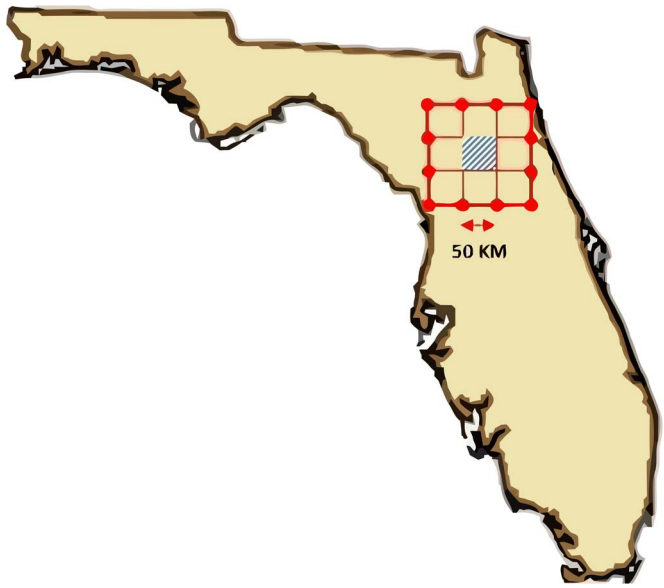

Fig. 1. Study region in North Central Florida. LSP-DSSAT-MB simulations were performed over the shaded 50×50 km^2 region.

The goal of this paper is to implement and evaluate a PRI-based methodology to downscale observations of SM from 10 to 1 km using the observations of LST, LAI, LC, and PPT at 1 km. The coarse resolution is representative of the SM products from the SMAP mission. The downscale methodology presented in this paper can be extended for the SMOS-SM product at 25 km. The primary objectives of this paper are to: 1) downscale observations of SM obtained at 10 km to 1 km using the PRI; and 2) evaluate the PRI-based methodology and compare it with a most recently used method for downscaling for the SMOS-SM product [29].

In the next section, we describe the framework to generate the synthetic data set and the coupled Soil Vegetation Atmospheric Transfer (SVAT)-vegetation-passive microwave model used in this paper.

II. Simulation Framework for Scaling Studies

A simulation framework, similar to [14] was developed to conduct synthetic experiments for heterogeneous agricultural landscapes with dynamic vegetation and to evaluate the proposed downscaling methodology [39]. A 50×50 km^2 region, equivalent to 25 SMAP pixels, was chosen in North Central Florida (see Fig. 1) for the simulations. The region encompassed the University of Florida/Institute of Food and Agricultural Sciences (UF/IFAS) Plant Science Research and Education Unit, Citra, FL, where a series of season-long field experiments, called the Microwave, Water and Energy Balance Experiments (MicroWEXs), have been conducted for various agricultural LCs over the last decade [40]–[42] used in this paper. A spatial resolution of 200 m was chosen for the model simulations, corresponding to the agricultural field site used during the MicroWEXs. Simulated observations of LST, LAI, and T_B were generated at 200 m for a period of one year,

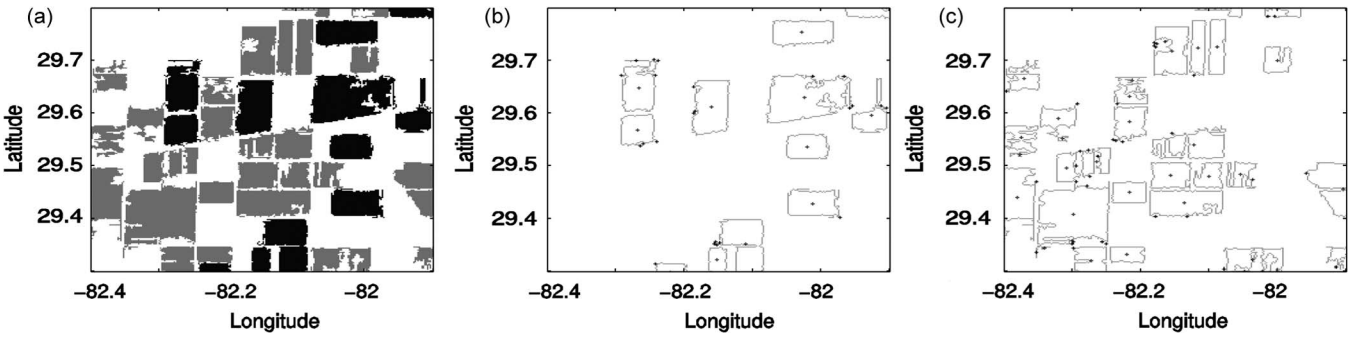

Fig. 2. (a) LC at 200m during cotton and corn seasons. White, gray, and black shades represent bare soil, cotton, and sweet-corn regions, respectively. Homogeneous crop fields along with centers for (b) sweet corn and (c) cotton.

from January 1, 2007 through December 31, 2007. Topographic features, such as slope, were not considered in this paper because the region is typically characterized by flat and smooth terrains with no runoff due to soils with high sand content. The soil properties were assumed constant over the study region. The assumption was based upon sand content estimated in the root zone using soil maps obtained from the U.S. Department of Agriculture. From the soil maps, 95% of the region comprised of soils with over 70% sand by volume in the root zone.

The simulation framework consisted of four components, i.e., a component for simulating meteorological forcings and LC to represent a heterogeneous landscape with dynamic vegetation conditions at 200 m; a coupled SVAT-crop growth model component for simulating SM, soil temperature, and crop growth and development; a forward microwave brightness (MB) model component for simulating L-band brightness temperature; and an upscaling component for generating synthetic T_B, SM, LST, PPT, and LAI observations at 10, and 1 km.

A. Meteorological Forcings and LC at 200 m

The 15-min observations of PPT, relative humidity, air temperature, downwelling solar radiation, and wind speed were obtained from eight Florida Automated Weather Network stations [43] located within the study region (see Fig. 1). The observations were spatially interpolated using splines to generate the meteorological forcings at 200 m. Long-wave solar radiation were estimated following Brutsaert [44].

In the 50×50 km^2 region, three primary agricultural LCs, sweet corn, cotton, and bare soil (see Fig. 2) were considered. A highly dynamic landscape was considered, with two seasons of sweet corn per year, the first season starting in March and the second in July at the same location; and one season of cotton planted in June. The sweet-corn season lasts about 78 days, whereas the cotton season is 179 days long, as shown in Table I. Irrigation and crop management schedules were based upon typical management practices from the MicroWEXs and from the farm managers at the PSREU. Bare soil conditions were assumed during the nongrowing season at these locations.

B. Coupled SVAT-Crop Growth Model

The SVAT model used in this paper is the Land Surface Process (LSP) model [45]. The model is forced with micrometeorological parameters, such as air temperature, relative humidity, downwelling solar and longwave radiation, irrigation/PPT, and windspeed. The model has been rigorously tested [46] and extended to wheat-stubble [47] and brome-grass [45] in the Great Plains, prairie wetlands in Florida [48], to tundra in the Arctic [49], and to growing crops [50]. The LSP model was coupled to a vegetation growth model, *viz.*, the decision support system for agrotechnology transfer (DSSAT) to provide estimates of water and radiation fluxes during dynamic vegetation conditions [50]. The DSSAT simulates crop growth and development at a daily step using modules for soil, soil–plant–atmosphere, weather, management, including irrigation and fertilization [51], [52]. The model has been extensively tested in different hydroclimatic regions [52]–[61]. The model was also tested and calibrated for its applicability to North-Central Florida [62] before it was coupled with the LSP model. In the coupled LSP-DSSAT model, the LSP model provides DSSAT with estimates of SM and temperature profiles and evapotranspiration. The DSSAT model provides LSP with vegetation characteristics that influence heat, moisture, and radiation transfer at the land surface and in the vadose zone [50].

TABLE I
PLANTING AND HARVEST DATES FOR SWEET CORN AND COTTON DURING THE 2007 GROWING SEASON

Crop	Planting DoY	Harvest DoY
Sweet Corn	61	139
	183	261
Cotton	153	332

C. MB Model for T_B Estimation

T_B measurements used by [29], were generated using the widely used $\tau - \omega$ model [63]. A vegetated surface is modeled as a single isothermal layer of vegetation with diffuse boundaries [64]. The soil medium is assumed to be a nonisothermal semi-infinite layered dielectric medium with a rough surface at the upper boundary. The SM and temperature profiles, the leaf/ear biomass, vegetation water content, and plant height provided by the LSP-DSSAT model were used by the MB model to estimate brightness temperature at L-band. Using a zeroth-order radiative transfer approach, the total brightness temperature of a terrain (T_B) is the sum of contributions

from soil ($T_{B\text{soil},p}$), vegetation ($T_{B\text{canopy},p}$), and from sky ($T_{B\text{sky},p}$)

$$\begin{aligned} T_{B\text{soil},p} &= (1-r_p)T_{\text{eff}}\exp(-\tau/\mu) \\ T_{B\text{canopy},p} &= T_c\left[1-\exp(-\tau/\mu)\right](1-\omega)\left[1+r_p\exp(-\tau/\mu)\right] \\ T_{B\text{sky},p} &= T_{\text{sky}}r_p\exp(-2\tau/\mu) \end{aligned} \quad (1)$$

where p is the polarization, r_p is the reflectivity of the rough soil surface, T_{eff} is the effective radiating temperature of the soil, $\mu = cos\theta$, where θ is the incidence angle, T_c is the physical temperature of the isothermal canopy, ω is the single-scattering albedo, τ is the optical depth, and T_{sky} is the downwelling sky brightness.

In this paper, the θ was set to 50°, the T_{sky} was set to 5 K [63], r_p was obtained by integrating the bistatic scattering coefficients from the Integral Equation Model (IEM) model [65], the RMS height was 0.62 cm and, the correlation length was 8.72 cm [41], [66]. The soil dielectric properties were obtained from the four-component model described in [67]. The optical depth was calculated from the linear relationship $\tau = b * VWC$ [68], where b is a regression coefficient that was set to 0.2 for both sweet corn and cotton, following [2].

D. Synthetic Observations at Resolutions of 1 and 10 km

The model simulations were performed over each agricultural field rather than all the pixels to reduce computation time. Based upon LC information at 200 m, contiguous homogeneous regions of sweet corn and cotton were identified, as shown in Fig. 2. A realization of the LSP-DSSAT-MB model was used to simulate T_B, LST, LAI, and PPT at the centroid of each homogeneous region, using the corresponding crop module within DSSAT. The model simulations were performed using the 200-m forcings at the centroid, as shown in Fig. 2. Linear averaging is typically sufficient to illustrate the effects of resolution degradation [69]. The model simulations at 200 m were spatially averaged to obtain PPT, LST, LAI, SM, and T_B at 1 and 10 km. The SM obtained at 1 km was used as truth to evaluate the downscaling methodology. To simulate rain-fed systems, all the water input from both PPT and irrigation were combined together, and the "PPT" in this paper represents these combined values.

III. Downscaling Framework

The problem of downscaling SM, i.e., creating a more detailed representation of SM at 1 km from coarser observations at 10 km, is limited by physical constraints dictated by the convolution of the point spread function of the imaging system. It is therefore an ill-defined problem that should be augmented or regularized with additional information relevant to SM at 1 km. This is provided by a nonparametric Bayesian first estimate of the data, i.e., an estimate obtained by a transformation function. Such a probabilistic approach is also robust against data gaps, as it can provide estimates without *a priori* PPT, LC, LAI, or SM information for a few days within the experimental period. This is particularly helpful because satellite products, such as LST and LAI, are affected by cloud cover resulting in sparse data sets. Here, we describe the downscaling framework based on PRI and briefly discuss the implementation of the UT method. Fig. 3(a) compares the two frameworks.

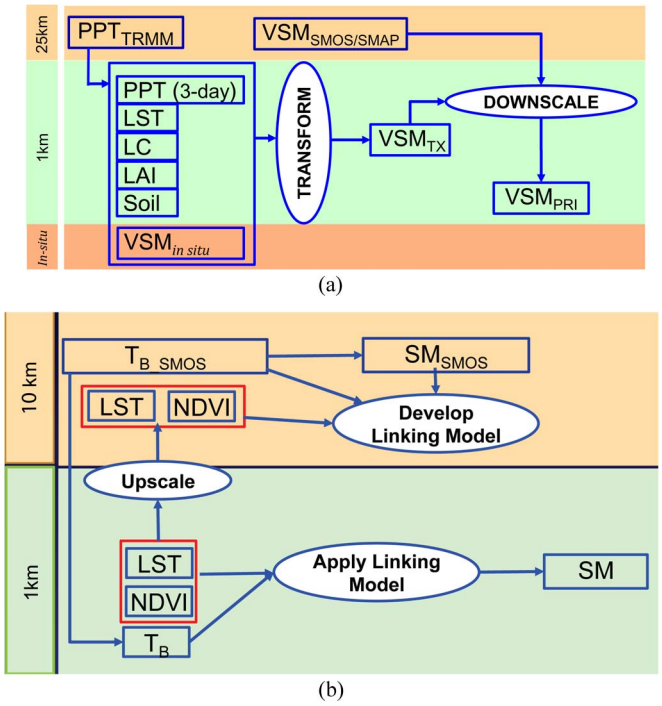

Fig. 3. Flow diagrams of the downscaling steps using the (a) PRI and (b) UT methods.

A. PRI-Based Framework

The downscaling methodology using PRI is implemented in two steps, as shown in Fig. 3(a). In the first step, an initial estimate of SM is obtained at 1 km using transformation functions that relate LST, LAI, PPT, and LC to SM at the same spatial resolution of 1 km. In the second step, this initial estimate is merged with the SM observations at the coarser resolutions of 10 km using PRI to obtain improved estimates of SM at 1 km.

1) Step 1—Transformation of LST, LAI, PPT, and LC at 1 km for Initial Estimates of SM at 1 km: The main objective of the transformation process is to obtain a probabilistic relationship between SM at 1 km using information from LST, LAI, PPT, and LC obtained at the same scale. As mentioned earlier, the PPT values included water input from both PPT and irrigation to simulate rain-fed systems. A discrete formulation of the Bayes rule is used to estimate SM, as given in (2), wherein SM is discretized into k classes, $i \in [1,k]$, LST into k_1 classes LST_{i_1}, $i_1 \in [1,k_1]$, LAI into k_2 classes LAI_{i_2}, $i_2 \in [1,k_2]$, PPT into k_3 classes LC_{i_3}, $i_3 \in [1,k_3]$ and LC into k_4 classes LC_{i_4}, $i_4 \in [1,k_4]$

$$\begin{aligned} &p\left(\text{SM}_i|\text{LST}_{i_1},\text{LAI}_{i_2},\text{PPT}_{i_3},\text{LC}_{i_4}\right) \\ &\quad = \frac{p(\text{LST}_{i_1},\text{LAI}_{i_2},\text{PPT}_{i_3},\text{LC}_{i_4}|\text{SM}_i)p(\text{SM}_i)}{p\left(\text{LST}_{i_1},\text{LAI}_{i_2},\text{PPT}_{i_3},\text{LC}_{i_4}\right)} \\ &\text{SM}_i \\ &\quad = \underset{\text{SM}_i}{\arg\max}\,\frac{p(\text{LST}_{i_1},\text{LAI}_{i_2},\text{PPT}_{i_3},\text{LC}_{i_4}|\text{SM}_i)p(\text{SM}_i)}{p\left(\text{LST}_{i_1},\text{LAI}_{i_2},\text{PPT}_{i_3},\text{LC}_{i_4}\right)} \end{aligned}$$

$$p\left(\mathrm{LST}_{i_1}, \mathrm{LAI}_{i_2}, \mathrm{PPT}_{i_3}, \mathrm{LC}_{i_4}\right) = \sum_{i=1}^{k} p\left(\mathrm{LST}_{i_1}, \mathrm{LAI}_{i_2}, \mathrm{PPT}_{i_3}, \mathrm{LC}_{i_4} | \mathrm{SM}_i\right) p(\mathrm{SM}_i). \quad (2)$$

A nonparametric PDF estimation technique, known as Parzen windows, [70] was used to estimate the PDF, $p(\mathrm{LST}_{i_1}, \mathrm{LAI}_{i_2}, \mathrm{PPT}_{i_3}, \mathrm{LC}_{i_4} | \mathrm{SM}_i)$, in (2). In this technique, symmetric functions, known as kernels, are used to estimate the PDF of a random variable based on a finite set of data samples. The kernels assign weights to each of the samples based on their distance from the point at which the PDF is estimated, with higher weights being assigned to the samples that are closer to the point. In this paper, a zero-mean unity-variance Gaussian function was used as the kernel. With Gaussian kernels, the weight decreases exponentially with the square of the distance, rendering the distant points irrelevant.

Given n training samples $\mathrm{LST}_j, \mathrm{LAI}_j, \mathrm{LC}_j, \mathrm{PPT}_j (j \in [1, n])$, the probability of LST, LAI, LC and PPT belonging to classes LST_{i_1}, LAI_{i_2}, PPT_{i_3} and LC_{i_4}, respectively, for a known SM_i is represented as $p(\mathrm{LST}_{i_1}, \mathrm{LAI}_{i_2}, \mathrm{PPT}_{i_3}, \mathrm{LC}_{i_4} | \mathrm{SM}_i)$ and computed using (3). The PDF is estimated at all combinations of LST, LAI, PPT, and LC classes resulting in a PDF of dimension $k_1 \times k_2 \times k_3 \times k_4$

$$\begin{aligned} &p\left(\mathrm{LST}_{i_1}, \mathrm{LAI}_{i_2}, \mathrm{PPT}_{i_3}, \mathrm{LC}_{i_4} | \mathrm{SM}_i\right) \\ &= \frac{1}{n} \sum_{j=1}^{n} \frac{1}{h_1 h_2 h_3 h_4} \psi\left(\frac{\mathrm{LST}_{i_1} - \mathrm{LST}_j}{h_1}\right) \\ &\quad \times \psi\left(\frac{\mathrm{LAI}_{i_2} - \mathrm{LAI}_j}{h_2}\right) \psi\left(\frac{\mathrm{PPT}_{i_3} - \mathrm{PPT}_j}{h_3}\right) \\ &\quad \times \psi\left(\frac{\mathrm{LC}_{i_4} - \mathrm{LC}_j}{h_1}\right) \end{aligned} \quad (3)$$

where ψ is the Gaussian kernel and $h_i,\ i \in [1, 2, 3, 4]$, is a smoothing parameter, which relates to the variance in ψ for LST, LAI, PPT and LC, respectively. The parameter h_i was determined based on Silverman's rule, which is a standard rule-of-thumb that is used in kernel-based estimation algorithms [71], as given in (4). The PDF $p(\mathrm{SM}_i)$ in (2) is a 1-D PDF that was computed at k points

$$h_i = 1.06 \min\left[\sigma_i, \frac{I_{QR_i}}{1.34}\right] n^{-1/5} \quad (4)$$

where σ_i and $I_{QR_i}, i \in [1, 2, 3, 4]$ are the standard deviation (SD) and interquartile range, respectively, computed from the n samples for LST, LAI, PPT and LC.

Computing the aforementioned PDFs from the n training samples can be viewed as a learning step. Given LST, LAI, PPT, and LC information at 1 km, the most probable SM value can be inferred using (2) to produce the initial estimates of SM at 1 km.

2) Step 2—Downscaling SM Using PRI: The PRI provides a hierarchical decomposition of spatial data that is optimal in terms of the transfer of information across scales, by minimizing or maximizing a balance between entropy and the Kullback–Leibler divergence (KLD) of the original data with respect to its upscaled or downscaled versions [70]. While entropy is a redundancy reduction term, the KLD is an information preserving term. If entropy is maximum, while KLD is 0, the observations remain the same, whereas if KLD is high with zero entropy, the data set is represented by a single point, the mean. Depending on the proportion of the two metrics, one can find a range of structures, which can be cast as an information extraction process. The optimization equation for PRI is

$$\max_X J(X) = H(X) + \beta \mathrm{KL}(p_X \| p_S) \quad (5)$$

where $J(X)$ is the cost function, p_S is the PDF of the original data, and p_X is the PDF at each iteration. $H(X)$ is the entropy, and KL is the KL divergence. The β is a user-defined weighting parameter that balances the redundancy and information preservation in $J(X)$. As the value of β increases, the cost function gives more and more emphasis to KL, thus preserving more information about the data at the cost of extremely high redundancy reduction. Equation (5) may be rewritten in a different form by using Renyi's quadratic entropy $(H_2(X))$ and cross-entropy $(H_2(X, S))$, respectively, [72], to obtain a more practical cost function $J(X)$, namely,

$$\max_X J(X) = (1 - \beta) H_2(X) + 2\beta H_2(X, S). \quad (6)$$

An iterative solution $x_k,\ k = [1, 2, \ldots, N]$ can be obtained by taking the derivation of $J(X)$ with respect to x and equating it to zero. The resulting fixed-point rule is given in (7)

$$x_k(t+1) = c \frac{1 - \beta}{\beta} \frac{\sum_{j=1}^{N} \psi\left(x_k(t) - x_j(t)\right) x_j(t)}{\sum_{j=1}^{M} \psi\left(x_k(t) - s_j\right)} + \frac{\sum_{j=1}^{M} \psi\left(x_k(t) - s_j\right) s_j}{\sum_{j=1}^{M} \psi\left(x_k(t) - s_j\right)} \quad (7)$$

where

$$c = \frac{V(X, S) M}{V(X) N}$$

$V(X)$ is the information potential [73] of data points in X, $V(X, S)$ is the cross-information potential between data points in X and S, N, and M are the number of data points in X and S, respectively, t represents the iteration, and ψ is the Gaussian kernel as explained in Section III-A.

To illustrate the functionality of the PRI, a simple example created by a set of Gaussian clusters placed in a circle is shown in Fig. 4(a). Variable X in (6) is initialized to the simulated Gaussian cluster data set (S). Applying the PRI for different values of β reveals various structures relevant to the data set. At one extreme, we have a single point, and at the other extreme, we have the data itself as the solution. As we move from one end to another, the PRI displays the modes and principal curves revealing different hierarchical levels of structure from the data.

The implementation of PRI was slightly modified for downscaling SM. Instead of initializing X to S, the initial estimates of SM at 1 km obtained from Section III-A were used as X

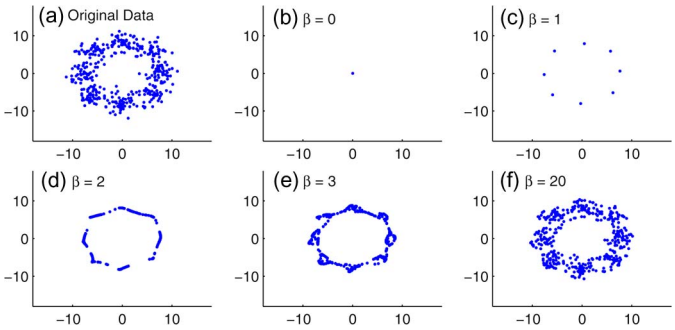

Fig. 4. (a) Conventional implementation of PRI where X is initialized to a sample data set S, producing different nonlinear tranformations in X based on β, (b) implementation of PRI for downscaling by morphing X and S, where X and S are initialized to the observations obtained from the transformation step at 1 km and the coarse-resolution satellite observation at 10 km.

in (6). SM at 10 km were used as S. Under the assumption that X and S represent the same imagery using different measurements, the cross-entropy term $H_2(X, S)$ should be largest when they coincide. This condition is obtained when $\beta = 1$ in (6). However, this case provides the best downscaled X supported by S, in spite of the fact that S represents only an approximate and coarse-level information on SM. In this paper, an intermediate value of $\beta = 2$ was chosen so that the PRI image would approximate the mean level of SM at coarse scales but will also embed the level of detail provided by the initial estimates of SM at 1 km, to obtain morphed estimates of SM at 1 km.

B. Downscaling Framework Based on UT Method

In this paper, we compare the PRI-based method with that based upon the UT. The UT method as used in [29] relates NDVI and LST from MODIS, and SMOS T_B at coarse resolution to SMOS-SM at coarse resolution. The topology for this method is shown in Fig. 3(b). This is achieved in two steps. First, a regression relationship is formulated at coarse resolution as

$$s_m = \sum_{i=0}^{n}\sum_{j=0}^{n}\sum_{k=0}^{n} a_{ijk}(\text{NDVI})^i(\text{LST})^j T_B^k. \tag{8}$$

Parameter n is the order of the model and is typically chosen as two for computational efficiency. NDVI, T_B, and LST are normalized with respect to the range of observations, with values between 0 and 1. In the second step, (8) is applied with regression coefficients a_{ijk}, along with NDVI, T_B, and LST at 1 km to obtain SM at 1 km, assuming that the relationship obtained at coarser resolution still holds true.

IV. Methodology

1) PRI-Based Algorithm: The terrain conditions were simulated from Jan 1 (DoY 1) to December 31 (DoY 365), 2007. This period consisted of two growing seasons of sweet corn and one season of cotton, as shown in Table I. The LST, PPT, and LAI observations at 1 km were obtained by adding noise to account for satellite observation errors, instrument measurement errors, and micrometeorological variability, following [14], [74], [75]. Two error distributions, i.e., Gaussian and Laplacian, were applied. The Gaussian distribution is typically used for most studies, but most measurement noise cannot be fully described with second-order statistics, and it may not be realistic. The Laplacian noise is a much higher order distribution and may be more realistic for highly nonlinear data sets. Implementing two different noise models also allows for testing the robustness of the downscaling methodology. In this paper, the errors with zero mean and standard deviations of 5K, 1 mm, 0.03 m^3/m^3, and 0.25 for LST, PPT, SM, and LAI, respectively, were added to the values at 10 km obtained in Section II-D.

TABLE II
Days Selected for Evaluating PRI Estimates. These Days Capture Variability in PPT/Irrigation and LC

DoY	PPT	LC
39	Dry	Bare
135	Dry, Irrigated	Sweet Corn
156	Wet	Cotton
222	Dry, Irrigated	Sweet Corn and Cotton
354	Wet	Bare

The PRI used LST, PPT, LAI, LC, and SM at 1 km every three days as input to obtain the transformation function. A nonparametric Bayesian classifier was trained with a subset of the aforementioned data with training percentages of 10%, 33%, and 50%. Using the SM observations at 10 km and first estimates of SM at 1 km from the transformation function, downscaled SM estimates using PRI were obtained using (6).

The performance of the PRI method was evaluated using PDFs and cumulative distribution functions (CDFs), as estimated by interpolating the discrete data set using kernel functions. The CDF of the error in downscaling SM, i.e., the absolute difference between the true SM and the downscaled SM estimates, at 1 km aggregated over the simulation period, is analyzed to understand the combined effects of bare and vegetated terrains and of Gaussian and Laplacian noise distributions on PRI performance. This analysis is also conducted for downscaled SM estimates aggregated during the growing seasons of sweet corn and cotton, to understand the effects of vegetation only. The kernel density estimate (KDE) of the true SM and downscaled SM at 1 km with Laplacian error is evaluated for four periods during the simulation year, when the LC in the region is bare, when the LC is a combination of bare soil and sweet corn, when the LC is a combination of bare soil, sweet corn, and cotton, and when the LC is a combination of only bare soil and cotton. During those four periods, the KDE is separately plotted for bare soil and vegetated pixels and for the overall region.

In addition, five days were selected to understand the effects of LC heterogeneity and PPT distribution on the SM estimates at 1 km. Variabilities in PPT, ranging from uniformly wet to uniformly dry, and in LC, ranging from bare soil to vegetated with both cotton and sweetcorn, were used as criteria for selecting the days, as shown in Table II. Quantitative analyses of spatial variations in SM observed under dynamic vegetation

and heterogeneous LC conditions provide an index of dynamic errors that can be expected. The algorithm was evaluated over the simulation period using both KLD and RMSE. For example, a methodology that consistently overestimates SM may have a low KLD value due to the error distribution being unimodal. On the other hand, if the error is multimodal, analyzing a second-order metric, such as RMSE without considering KLD values, can lead to misinterpretation of algorithm behavior, as the distribution of the errors cannot be known *a priori*. The effect of uncertainty in the coarse-resolution observations was investigated using a normally distributed error in SM at 25 km, with standard deviations ranging from 0.02 to 0.16 $\mathrm{m}^3/\mathrm{m}^3$

$$H(e) = -\sum_{i=1}^{n} p(e_i)\log\left(p(e_i)\right) \tag{9}$$

$$\mathrm{KLD(SM)} = \sum_{i=1}^{n} p\left(\mathrm{SM}_i^{\mathrm{est}}\right)\log\frac{p\left(\mathrm{SM}_i^{\mathrm{est}}\right)}{p\left(\mathrm{SM}_i^{\mathrm{truth}}\right)}. \tag{10}$$

The PDF-based measure of KLD [(10)], computed between the PDFs of simulated and estimated SM at 1 km, was used to quantify the efficacy of the downscaling algorithm in preserving the distribution of SM. The effect of the Gaussian and Laplacian noise models are also compared using their KLD measures, which can be also interpreted as the analog of the distance metric between the downscaled and true observation KDE's in some probability space. Second-order error metrics such as RMSE and SD were also included to quantify the errors in the estimates.

Finally, scenarios were simulated where minimal *in situ* observations were available for the transformation function to quantify the performance of the downscaling algorithm in data-sparse regions. The transformation function was generated using only bare soil pixels, and applied to the heterogeneous region, simulating a data poor condition.

2) Comparison of PRI and UT Methods: The UT method uses maximum LST, LAI, T_B, and SM at 1 and 10 km every three days for downscaling. The 10-km observations were obtained by adding noise with two distributions, similar to Section IV-A1. For downscaling with the UT method, first, the regression coefficients in (8) are calculated using normalized versions of LAI, T_B, LST, and SM at 10 km. These regression coefficients are used in linear equations at the fine resolution of 1 km to obtain the downscaled SM.

The PRI estimates were compared with the UT estimates using the CDF of the error and the KDE of downscaled observations, similar to Section IV-A1. The downscaled estimates from the PRI were compared with those obtained from the UT method were then compared with the PRI estimates for the selected five days during the simulation period, representing different micrometeorological and LC conditions.

V. Results and Discussions

A. PRI-Based Downscaling Algorithm

The initial estimate of SM was calculated at 1 km using the probabilistic relationship of PPT, LST, LAI, and LC with SM. The first estimates of SM extracts the structure and hetero-

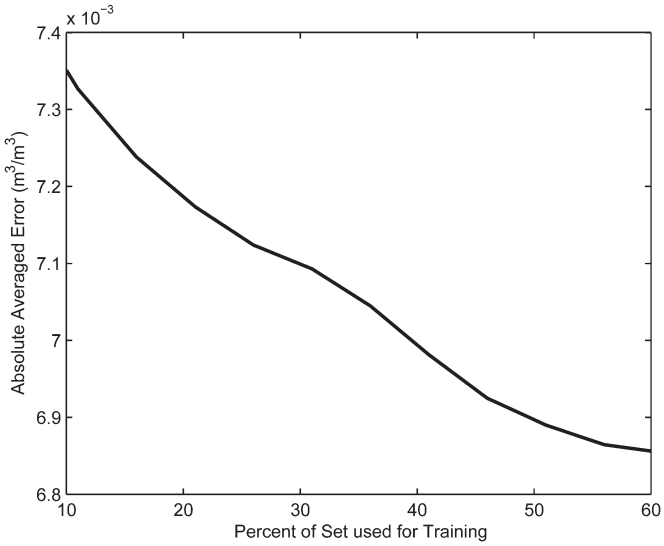

Fig. 5. Absolute averaged error in SM at 1 km over the simulation period as a function of percent of data set used for training during the development of the transformation function.

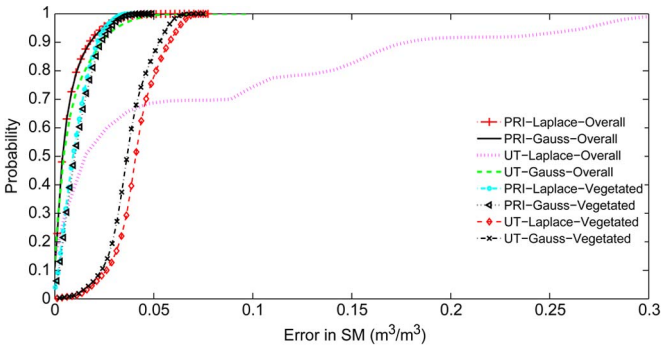

Fig. 6. CDF for SM errors between the true SM at 1 km and those obtained by using PRI and UT methods, with Gaussian and Laplacian noise distributions. The "vegetated" period includes sweet corn crop from DoY 61–139 and cotton crop from DoY 153–332 seasons. Note that the cotton season also includes a second sweet-corn season. The "overall" includes bare soil and vegetated LCs during the simulation period.

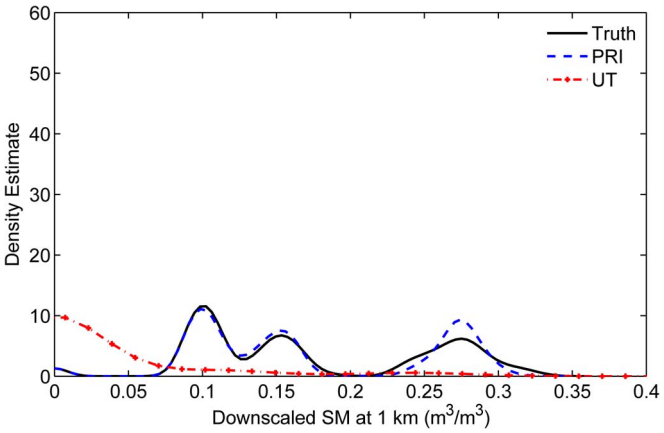

Fig. 7. Comparison of KDE of true SM values at 1 km, downscaled SM estimate using PRI and UT methods at 1 km during the days when the LC in the region consisted of bare soil only, viz., DOY 1–61, 139–153, and 332–365.

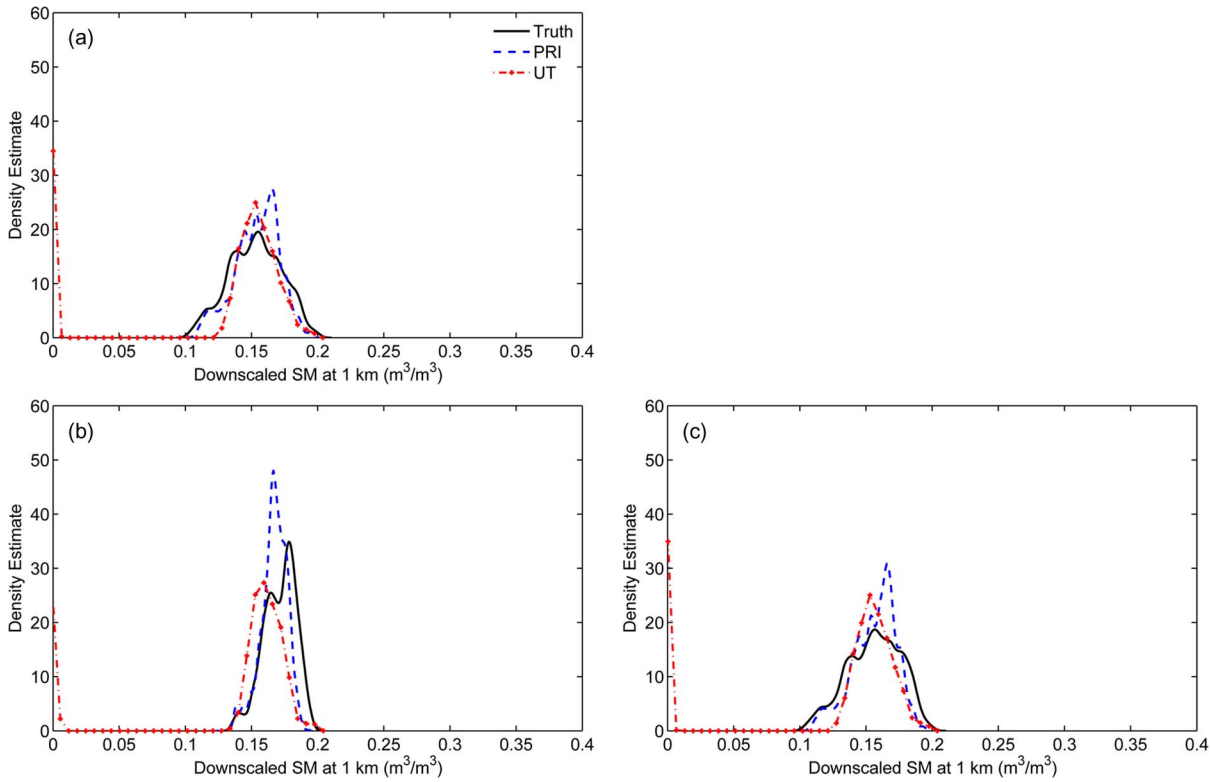

Fig. 8. Comparison of KDE of true SM values at 1 km, downscaled SM estimate using PRI and UT methods at 1 km for days when the LC consisted of sweet corn, and bare soil, DOY 61–139 over (a) only bare soil pixels, (b) only corn pixels, and (c) all pixels.

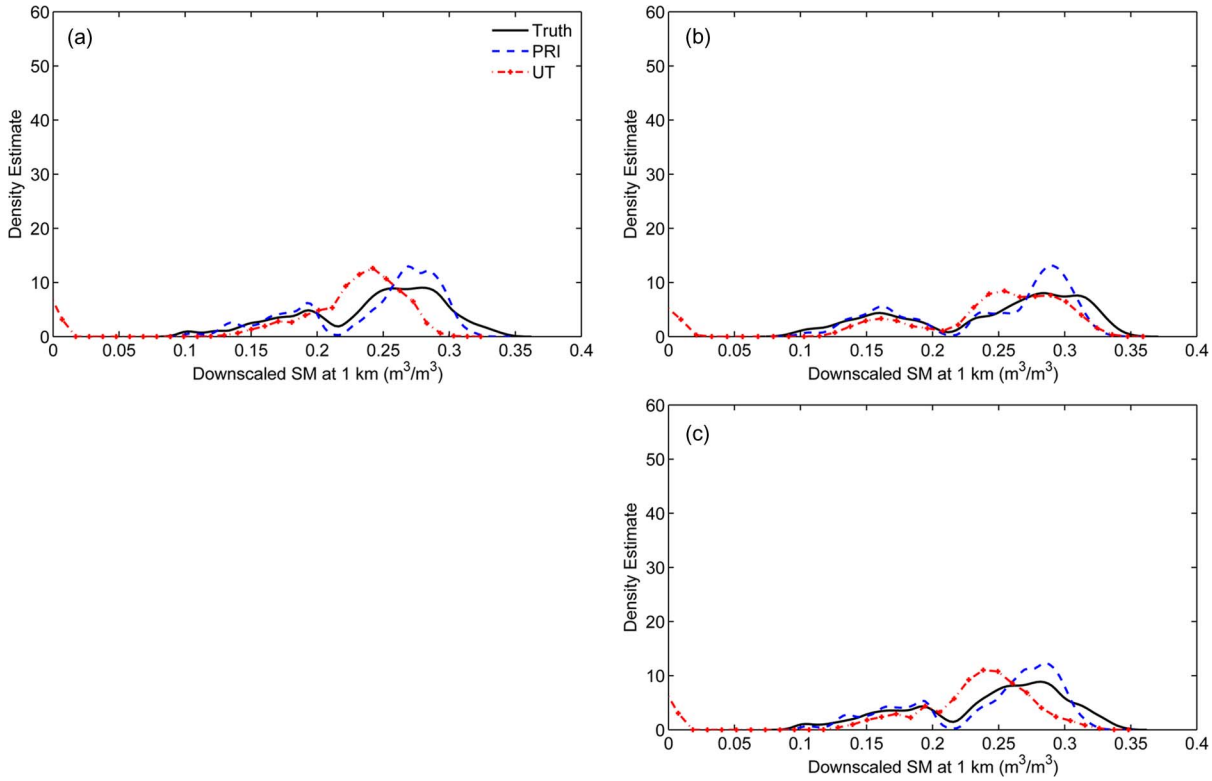

Fig. 9. Comparison of KDE of true SM values at 1 km, downscaled SM estimate using PRI and UT methods at 1 km for days when the LC consisted of cotton and bare soil, DOY 153–183 and 261–332 over (a) only bare soil pixels, (b) only cotton pixels, and (c) all pixels.

geneity in the spatial distribution of SM and has a regionally averaged RMSE of 0.04 m^3/m^3 over the simulation period compared with the true SM at 1 km. As shown in Fig. 5, increasing the size of the training data set above 33% decreases the errors by only 0.005 m^3/m^3 by volume. Thus, in this paper, 33% of the data set was used for training.

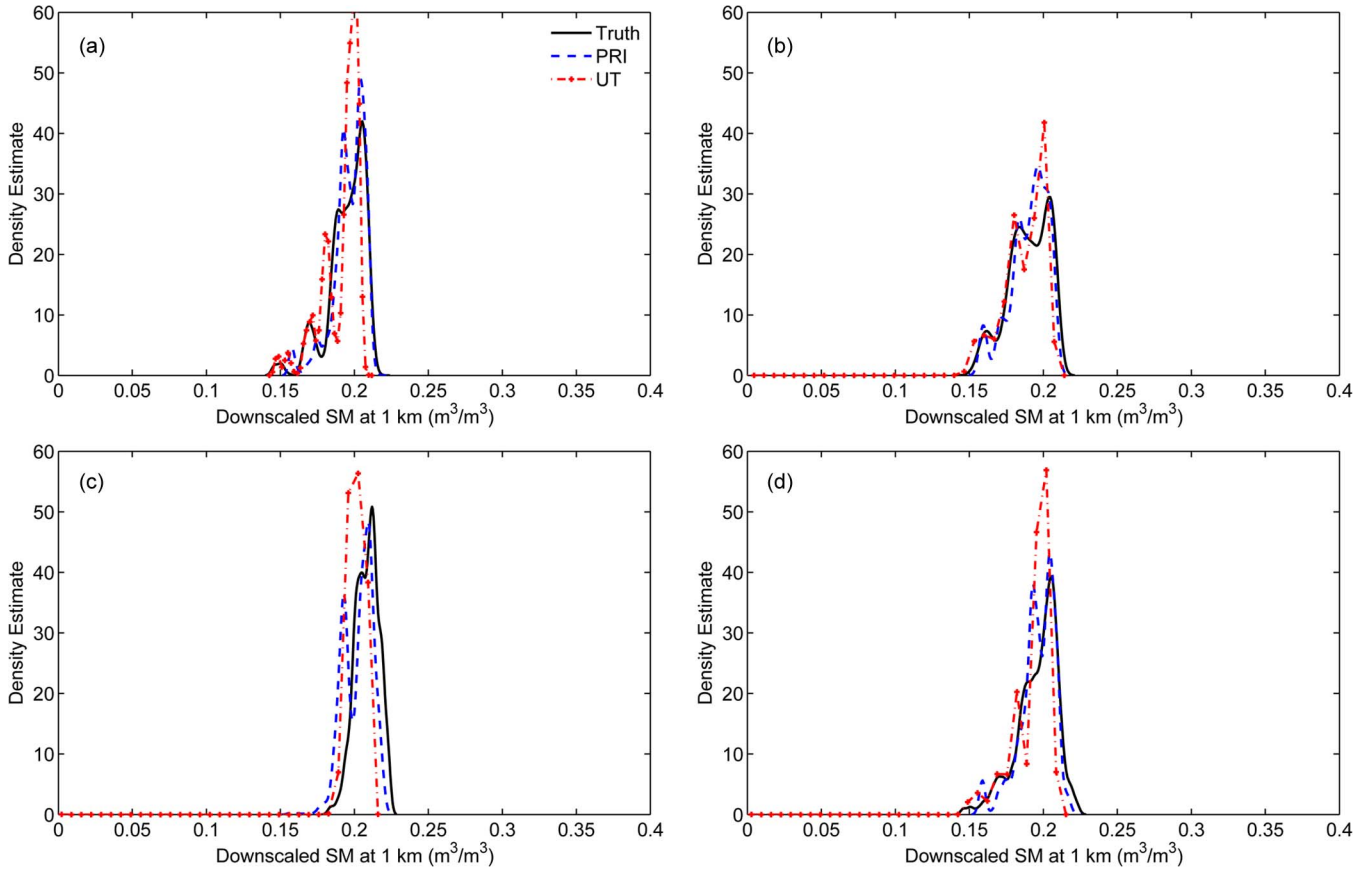

Fig. 10. Comparison of KDE of true SM values at 1 km, downscaled SM estimate using PRI and UT methods at 1 km for days when the LC consisted of sweet corn, cotton, and bare soil, DOY 183–261 over (a) only bare soil pixels, (b) only cotton pixels, (c) only corn pixels, and (d) all pixels.

The CDFs of the error for PRI estimates on days with overall and vegetated LC are shown in Fig. 6. The overall error CDF reaches the maximum when the errors in SM are about 0.05 m^3/m^3, indicating high probability of low error. The CDF for the vegetated days is very similar to that for the overall period, with a maximum difference of 0.003 m^3/m^3, which shows that the PRI is reasonably robust to vegetation heterogeneity. The error CDF for the PRI estimates is similar for both Gaussian and Laplacian noise distributions demonstrating the robustness of the PRI algorithm even for higher moment error models. From the CDF curve, it is clear that the individual pixel errors are less than 0.004 m^3/m^3 by volume, with 95% confidence.

Fig. 7 shows the KDE for the downscaled SM in bare soil pixels during the nongrowing season using the PRI compared with the KDE of the truth. The estimated densities of the PRI matches the best with the true SM in this case, which is expected as with a fully homogeneous LC the 4-D problem becomes essentially a 3-D problem, with one sparse dimension. Figs. 8–10 show the Kernel estimates of the PDF of downscaled SM during the growing seasons when there is only cotton, only corn, and both corn and cotton crops, respectively. A measured theoretic interpretation of the error is the sum of the distances between the true density and the projection of each of the input densities, using the PRI as a mapping function, in a suitable probability space. During bare soil, one of the marginals, the LC, is always degenerate or constant valued and does not contribute to the overall error. Thus, the errors will always be higher for vegetated periods than for the bare soil case. The downscaled SM value in all the vegetated cases is the most accurate when the SM < 0.1 m^3/m^3. The PRI KDE is also identical to the truth KDE during flooded conditions, with SM $>$ m^3/m^3. For the five selected days, the inputs, the first SM estimate, and PRI-downscaled SM are shown in Figs. 11–15. The RMSE, SD, and KLD for these days are shown in Table III. Both DoY 39, shown in Fig. 11 and DoY 354, shown in Fig. 13 are during bare soil LC before and after the growing seasons, respectively. The RMSEs for both days are very low, but due to crop residue and slightly heterogeneous PPT in the region [see Fig. 13(e)], the RMSE for DoY 354 is higher than DoY 39, as shown in Table III. It was found that heterogeneity in any one input, is enough to capture vegetation patterns in the PRI-downscaled estimate, as shown in Figs. 13(a) and 14(a), for corn and cotton, when the LST is fairly uniform across the region, whereas PPT is heterogeneous due to PPT patterns. Due to equal weights given to both coarse resolution and fine resolution, with $\beta = 2$ in (5), PRI tends to average or blur SM across crop boundaries, for example on DoY 222, when there was maximum heterogeneity in LC with corn, cotton, and bare soil, as shown in Fig. 15. This tendency can be altered by a different β, in a real-world scenario. However, the benefit of assigning equal weights is that the algorithm does not over- or underestimate SM by more than 0.005 m^3/m^3 in key areas, to keep the vegetation boundaries sharp. In the most heterogeneous case, on DoY 222 the KLD is the highest, signifying the most difference between the PDF of the downscaled observation and truth. In the other cases, KLD is much lower.

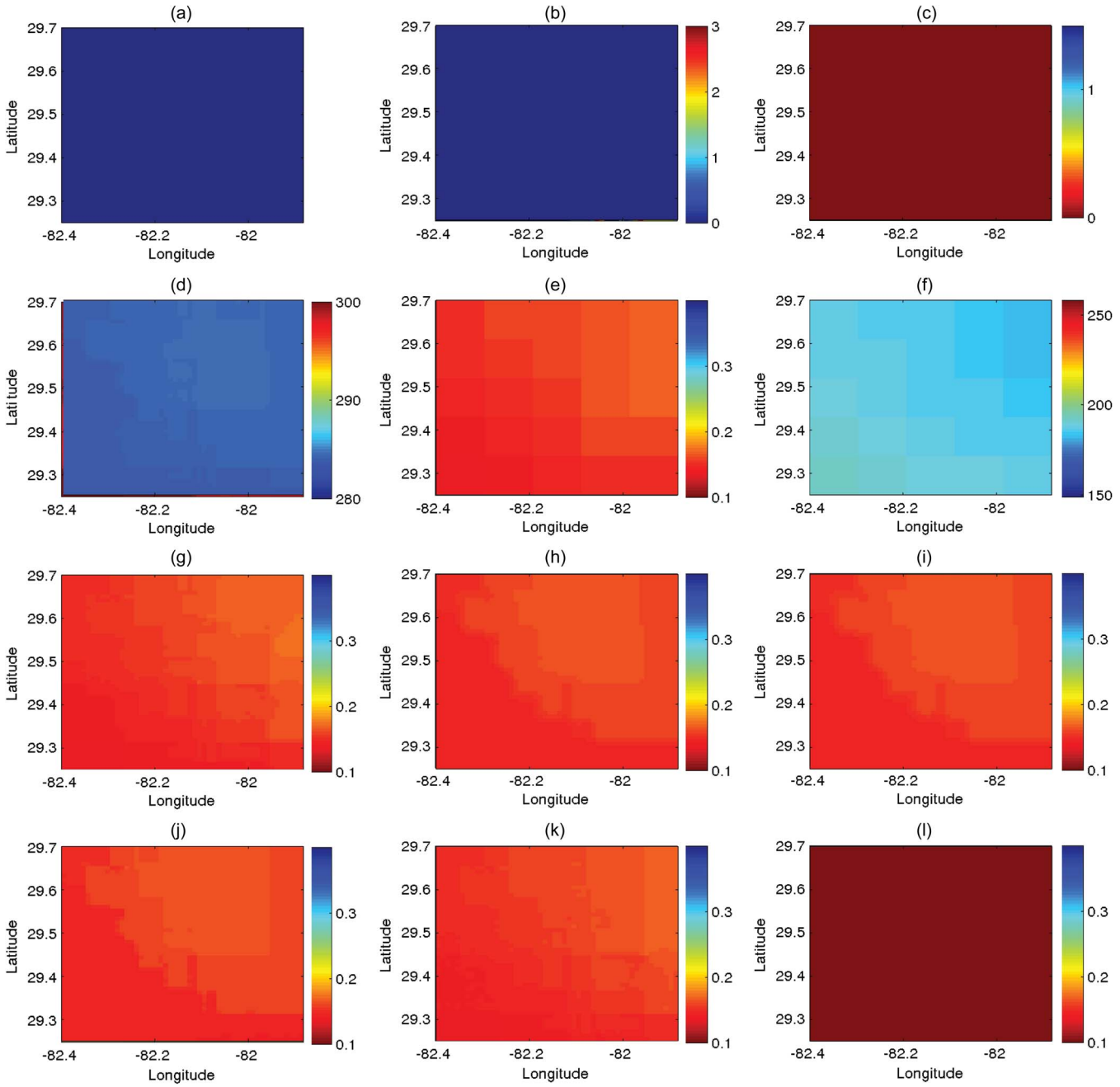

Fig. 11. DoY 39—(a) LC at 1 km (blue represents baresoil), (b) LAI at 1 km, (c) PPT at 1 km, (d) LST at 1 km, (e) SM observations at 10 km, (f) T_B at 10 km, (g) true SM at 1 km, (h) PRI estimate with Gaussian error at 1 km, (i) PRI estimate with Laplacian error at 1 km, (j) transformation function estimate at 1 km, (k) UT estimate with Gaussian error at 1 km, and (l) UT estimate with Laplacian error at 1 km.

Fig. 16 shows the RMSE in downscaled SM at 1 km when normally distributed errors with increasing standard deviations are added spatially to the 10-km observations. The RMSE stays almost constant at 0.004 m^3/m^3 with an increase in SD until the SD is lower than 0.06 m^3/m^3. The RMSE for the bare pixels were slightly lower than those for the vegetated pixels, within 0.05 m^3/m^3. This indicates a very low inherent error in downscaling. As the SD increases from 0.06 to 0.16 m^3/m^3, RMSE increases almost linearly with a slope of 0.16. When this was repeated for different days, the curve was almost same when SD of errors added more than the inherent downscaling error. Table IV shows the impact of the amount of training data on errors in downscaled SM for data-rich and data-poor regions. As expected, the errors increase to 0.025m^3/m^3 by volume from 0.00715m^3/m^3 when 33% of the data was used for training for data poor regions. This value of error is within the uncertainty of satellite observations.

B. Comparison of PRI and UT Methods

The error CDF in downscaled observations using PRI and the UT methods are compared in Fig. 6. The CDF of errors with the UT method using Laplacian error distribution is significantly worse than the CDF with Gaussian errors. The CDF for the bare soil with Laplacian errors is significantly different for UT method because being second order, it cannot effectively deal

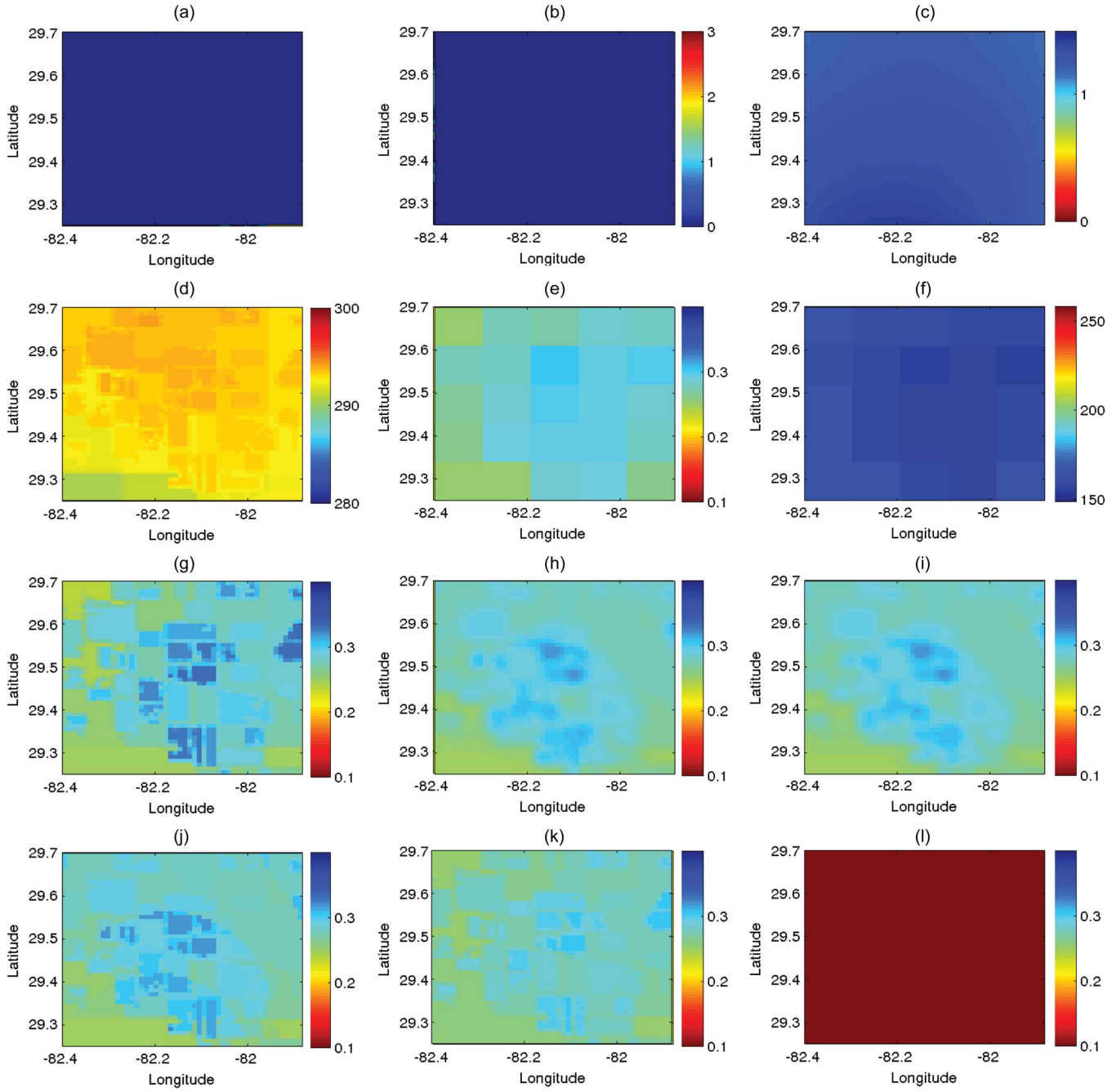

Fig. 12. DoY 354—(a) LC at 1 km (blue represents baresoil), (b) LAI at 1 km, (c) PPT at 1 km, (d) LST at 1 km, (e) SM observations at 10 km, (f) T_B at 10 km, (g) true SM at 1 km, (h) PRI estimate with Gaussian error at 1 km, (i) PRI estimate with Laplacian error at 1 km, (j) transformation function estimate at 1 km, (k) UT estimate with Gaussian error at 1 km, and (l) UT estimate with Laplacian error at 1 km.

with higher order noise and due to convergence issues in some bare soil pixels with the UT method, during the times when LAI was 0 but the SM is heterogeneous due to antecedent PPT; thus, there is some residual SM. Thus, the noise models affect the UT Method, whereas their effect is minimal in the PRI-based method. This indicates that a downscaling method that includes higher order relationships may be more suitable for highly dynamic and heterogeneous LCs such as agricultural fields.

Fig. 7 shows the KDE of downscaled SM for bare pixels using the PRI and UT methods, both using Laplacian error distribution, compared with the PDF of the truth. Even for the straightforward case of only bare soil, the UT method consistently underestimates SM due to convergence issues. For the case of only corn crop, as shown in Fig. 9, or only cotton crop, as shown in Fig. 8, UT method does not provide a reasonable reconstruction of true KDE. This difference is expected because the method since only conserves the first two moments. This is corroborated in the case of most heterogeneous vegetation landscape with both corn and cotton, as shown in Fig. 10, where the differences between the KDE are maximum. During the vegetated days, the UT method over-estimates SM implying a higher probability of more errors, as evident from the error KDE. The PRI estimates are much closer to the KDE of the truth, in all the cases.

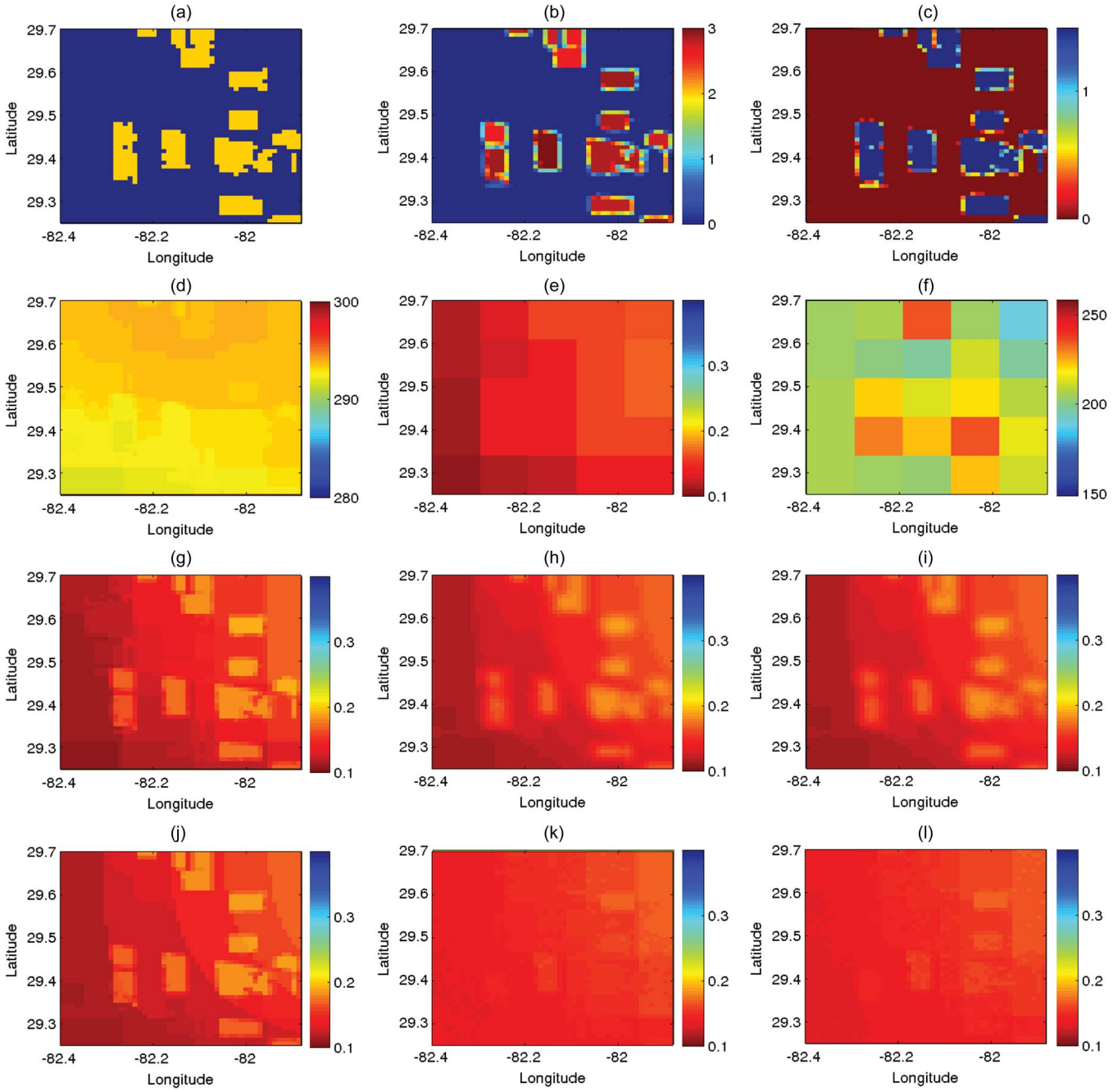

Fig. 13. DoY 135—(a) LC at 1 km (blue represents baresoil and yellow represents sweet-corn), (b) LAI at 1 km, (c) PPT at 1 km, (d) LST at 1 km, (e) SM observations at 10 km, (f) T_B at 10 km, (g) true SM at 1 km, (h) PRI estimate with Gaussian error at 1 km, (i) PRI estimate with Laplacian error at 1 km, (j) transformation function estimate at 1 km, (k) UT estimate with Gaussian error at 1 km, and (l) UT estimate with Laplacian error at 1 km.

DoY 354, as shown in Fig. 12, exhibits the effect of downscaling using PRI on a completely wet day with no vegetation cover. However, as shown in Fig. 12(d), the remnants of crops cause variability in SM, which is accurately reflected in the PRI estimate [see Fig. 12(h)] and also in the UT method estimate [Fig. 12(k)]. The RMSE is 0.5% less for PRI compared than the UT method, even for this case with relatively homogeneous inputs. In constrast, under homogeneous vegetation and PPT conditions, errors remain very low for both PRI and UT method as evident on DoY 39, with bare soil, as shown in Fig. 11. The RMSE's in this case are 0.01 and 0.03 m^3/m^3, for PRI and the UT methods, respectively. The SD are also very low.

On a day with vegetated soil (DoY 135, corn) and heterogeneous PPT [see Fig. 13(c)], the UT Method has 1.3% more RMSE than PRI. The high spatial variability of the inputs (see Fig. 12) is better estimated by PRI as it uses all the modes and moments of the data. Similarly, this effect is also cotton season on DoY 156, as shown in Fig. 14.

During the period when the land is heavily vegetated, the RMSE's increase for both the methods as the data exhibits higher variability. For example, on DoY 222, there is a high variability in LC, LAI, and TB, with two crops and bare soil, highly variable LAI having crops of both corn and cotton [see Fig. 15(b)] and T_B [see Fig. 15(f)]. Interestingly, although PPT

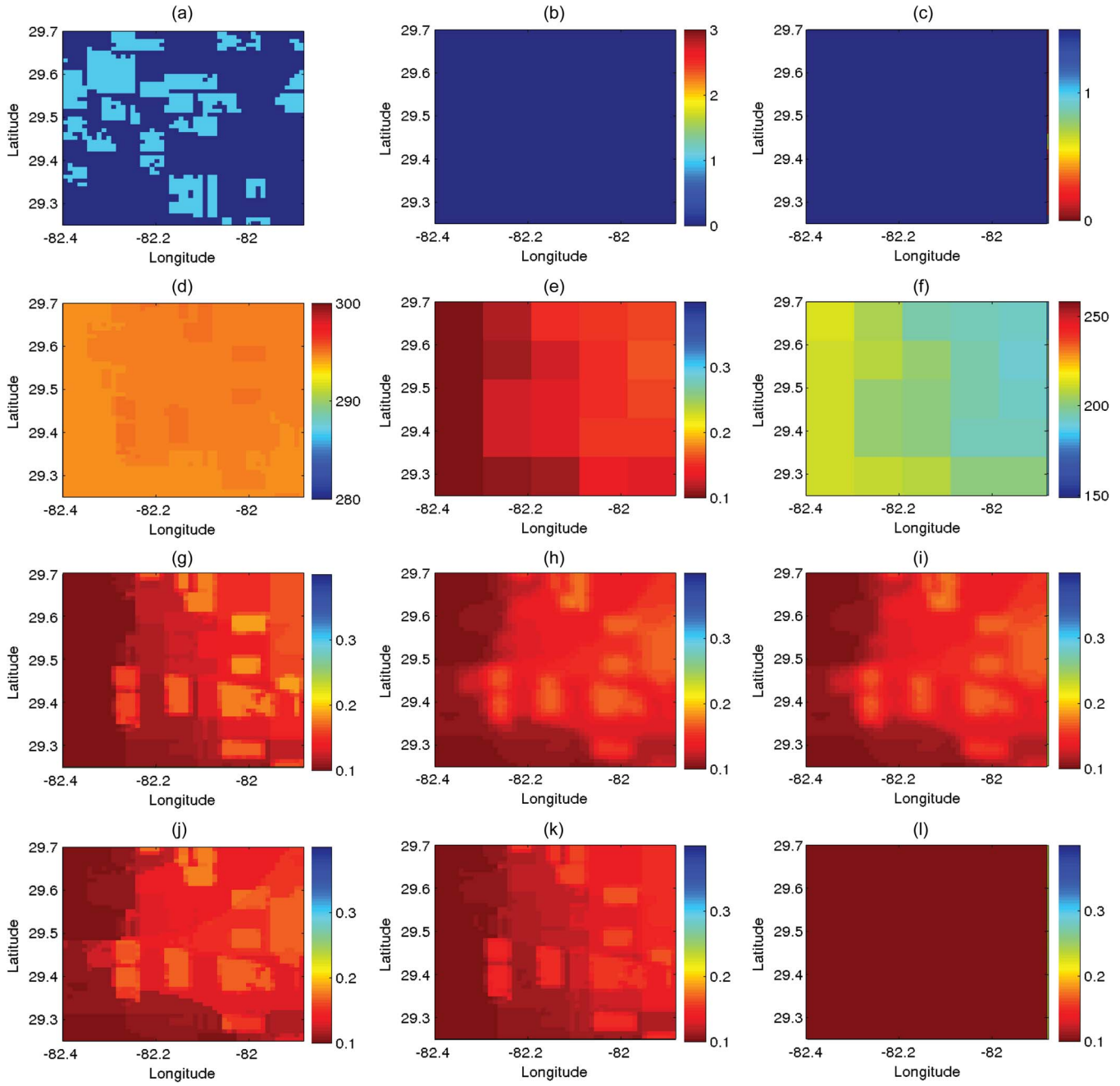

Fig. 14. DoY 156—(a) LC at 1 km (blue represents baresoil and cyan represents cotton), (b) LAI at 1 km, (c) PPT at 1 km, (d) LST at 1 km, (e) SM observations at 10 km, (f) T_B at 10 km, (g) true SM at 1 km, (h) PRI estimate with Gaussian error at 1 km, (i) PRI estimate with Laplacian error at 1 km, (j) transformation function estimate at 1 km, (k) UT estimate with Gaussian error at 1 km, and (l) UT estimate with Laplacian error at 1 km.

is not very heterogeneous, PRI has 1.1% lower RMSE than the UT method. The KLD is > 1 for the UT method, indicating that the downscaled SM has significantly different modes than the true SM at 1 km. For the PRI, the KLD is < 1 on DoY 222 but still considerably higher than the other days.

A second-order regression method, such as the UT method, is expected to behave optimally with Gaussian noise. However, even with no heterogeneity in PPT and LC on DoY 39, the RMSE for the UT method with Laplacian noise increases to 0.16 m^3/m^3, whereas the RMSE for the PRI does not change. This degradation gets worse for vegetated soil. On DoY 156, during cotton season, when PPT is high, the RMSE for the UT method is 5.2%, and PRI increases to 0.89%. The KLD is also high with the Laplacian noise for the vegetated days, for the UTD method, even when the KLD with Gaussian noise is close to 0, as shown in Table III.

VI. Summary and Conclusion

In this paper, we have implemented and evaluated a downscaling methodology based upon PRI that preserves the high variability in SM due to heterogeneous meteorological and vegetation conditions. The PRI preserves heterogeneity by utilizing the data structure rather than the second-order statistics such

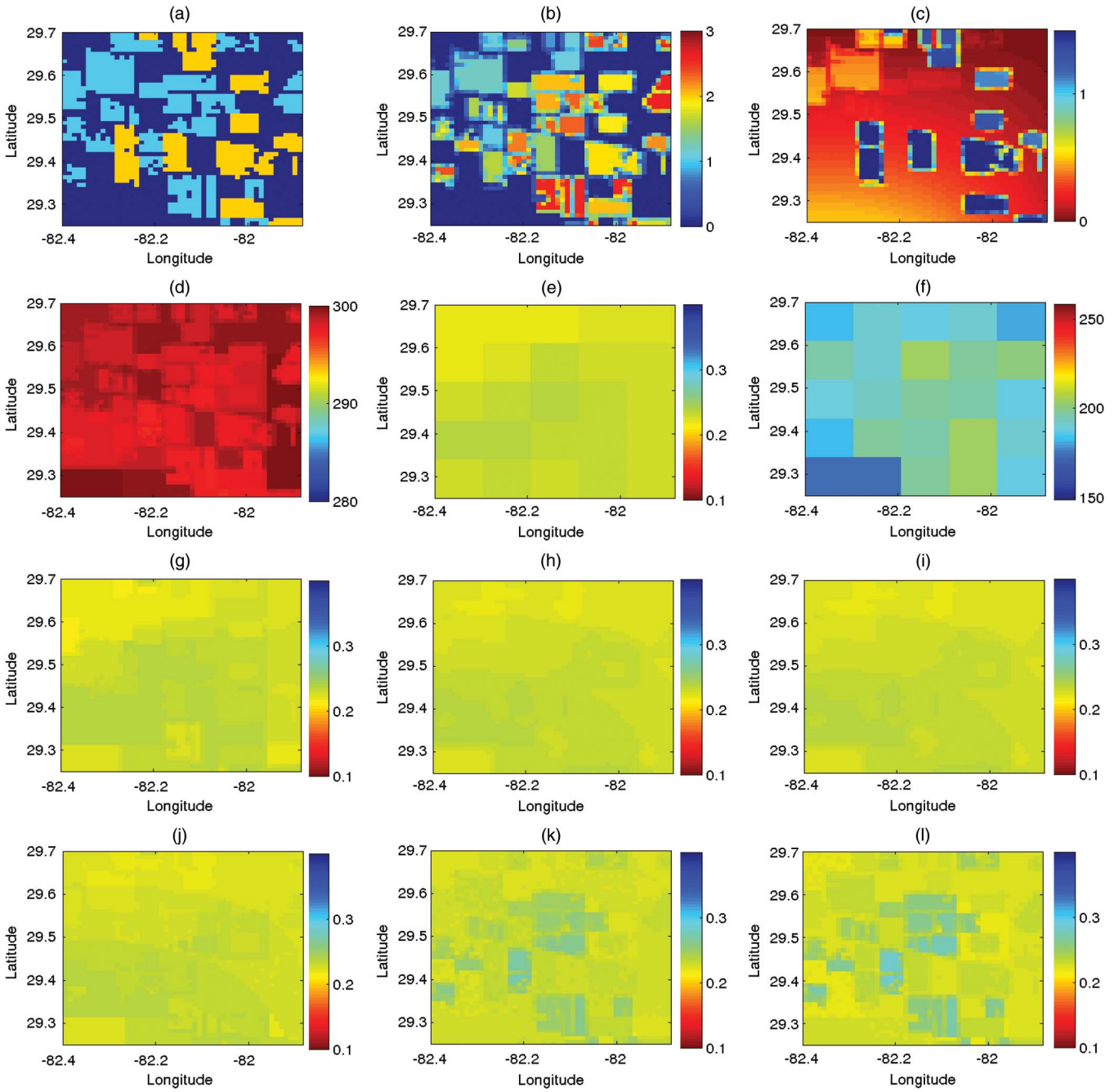

Fig. 15. DoY 222—(a) LC at 1 km (blue represents baresoil, yellow represents sweet-corn, and cyan represents cotton), (b) LAI at 1 km, (c) PPT at 1 km, (d) LST at 1 km, (e) SM observations at 10 km, (f) T_B at 10 km, (g) true SM at 1 km, (h) PRI estimate with Gaussian error at 1 km, (i) PRI estimate with Laplacian error at 1 km, (j) transformation function estimate at 1 km, (k) UT estimate with Gaussian error at 1 km, and (l) UT estimate with Laplacian error at 1 km.

as means and variances. The downscaling methodology was regularized by initial estimates of SM finer spatial resolution by transforming information embedded within RS products, *viz.*, PPT, LST, LAI, and LC, and *in situ* SM using nonparametric PDFs. The RMSE and KLD values in downscaled SM at 1 km for the PRI method, with a Gaussian noise model are both close to zero for bare soil and increases to 0.01 $\mathrm{m}^3/\mathrm{m}^3$ and 0.66, respectively, for vegetated LC. The similarity between the KDE estimates of the true and downscaled SM show that the structure of the PDF of observations is retained while downscaling. In contrast, the UT method has a RMSEs and KLDs of upto 0.15 $\mathrm{m}^3/\mathrm{m}^3$ and 1.07, respectively, when the land is vegetated. These values are close to zero for the bare soil, similar to the PRI. With a Laplacian noise model, the KLD and RMSE values for the PRI method remains almost the same as those for the Gaussian noise model. However, for the UT method, the RMSE's and KLD's increase by almost 100% in some cases, compared with the values with Gaussian noise model.

It is envisioned that the PRI-based method implemented and evaluated in this paper will be applied using satellite-based higher resolution remote sensing data. For example, the PPT data are obtained from the Tropical Rainfall Measurement and

TABLE III
RMSE, SD, AND KLD OVER THE 50×50 km^2 REGION FOR THE DOWNSCALED ESTIMATES OF SM OBTAINED AT 1 km USING THE PRI AND UT METHODS

Estimates	Noise model	DoY	RMSE (m^3/m^3)	SD (m^3/m^3)	KL divergence
PRI	Gaussian	39	0.002	0.002	0.019
		135	0.005	0.018	0.006
		156	0.009	0.021	0.038
		222	0.003	0.021	0.661
		354	0.011	0.021	0.087
	Laplacian	39	0.002	0.002	0.017
		135	0.005	0.018	0.007
		156	0.009	0.021	0.040
		222	0.003	0.021	0.572
		354	0.011	0.021	0.080
UTD	Gaussian	39	0	0.001	0.079
		135	0.149	0.018	0.016
		156	0.007	0.021	0.005
		222	0.015	0.020	1.072
		354	0.010	0.021	0.006
	Laplacian	39	0.157	0.009	0.078
		135	0.014	0.023	0.009
		156	0.130	0.026	0.225
		222	0.023	0.022	1.100
		354	0.281	0.022	0.107

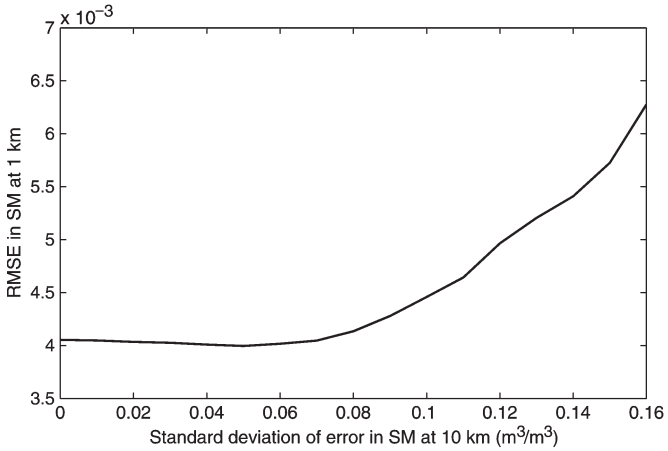

Fig. 16. RMSE in downscaled SM and the truth at 1 km as a function of the SD of errors in the coarse SM at 10 km.

TABLE IV
AVERAGE ABSOLUTE ERRORS (AAE) IN DOWNSCALED SM AS A FUNCTION OF THE PERCENTAGE OF THE DATA SET USED FOR TRAINING IN DATA-RICH AND DATA-SPARSE REGIONS

% of total observations used for training	AAE for Data Rich Region ($\mathrm{m}^3/\mathrm{m}^3$)	AAE for Data Poor Region ($\mathrm{m}^3/\mathrm{m}^3$)
10	7.35×10^{-3}	2.71×10^{-1}
20	7.15×10^{-3}	2.53×10^{-2}
30	7.1×10^{-3}	2.54×10^{-2}
40	7×10^{-3}	2.54×10^{-2}
50	6.85×10^{-3}	2.51×10^{-2}
60	6.85×10^{-3}	2.5×10^{-2}

Global Precipitation Measurement missions and the LAI, LST, and LC products are available from the Moderate Resolution Imaging Spectroradiometer sensor aboard Aqua and Terra satellites.

ACKNOWLEDGMENT

The authors would like to thank computational resources and support provided by the University of Florida High-Performance Computing Center for all the model simulations conducted in this paper.

REFERENCES

[1] A. Loew and W. Mauser, "On the disaggregation of passive microwave soil moisture data using apriori knowledge of temporally persistant soil moisture fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 3, pp. 819–834, Mar. 2008.

[2] T. Jackson and T. Schmugge, "Vegetation effects on the microwave emission of soils," *Remote Sens. Environ.*, vol. 36, no. 3, pp. 203–212, Jun. 1991.

[3] T. Jackson and T. Schmugge, "Surface soil moisture measurement with microwave radiometry," *Acta Astron.*, vol. 35, no. 7, pp. 477–482, Apr. 1995.

[4] S. Ahmad, A. Kalra, and H. Stephen, "Estimating soil moisture using remote sensing data: A machine learning approach," *Adv. Water Res.*, vol. 33, no. 1, pp. 69–80, Jan. 2010.

[5] C. Gruhier *et al.*, "Soil moisture active and passive microwave products: Intercomparison and evaluation over a Sahelian site," *Hydrol. Earth Syst. Sci.*, vol. 14, no. 1, pp. 141–156, 2010.

[6] B. W. Barrett, E. Dwyer, and P. Whelan, "Soil moisture retrieval from active spaceborne microwave observations: An evaluation of current techniques," *Remote sens.*, vol. 1, no. 3, pp. 210–242, Jul. 2009.

[7] J. Qin *et al.*, "Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal," *J. Geophys. Res.*, vol. 114, no. D15, pp. D15103-1–D15103-13, Aug. 2009.

[8] T. Lakhankar, N. Krakauer, and R. Khanbilvardi, "Applications of microwave remote sensing of soil moisture for agricultural applications," *Int. J. Terraspace Sci. Eng.*, vol. 2, no. 1, pp. 81–91, 2009.

[9] K. B. Mao *et al.*, "A method for retrieving soil moisture in Tibet region by utilizing microwave index from TRMM/TMI data," *Int. J. Remote Sens.*, vol. 29, no. 10, pp. 2903–2923, May 2008.

[10] L. Wang, J. J. Qu, S. Zhang, X. Hao, and S. Dasgupta, "Soil moisture retrieval using EOS MODIS and ground measurements in Eastern China," *Int. J. Remote Sens.*, vol. 28, no. 6, pp. 1413–1418, Mar. 2007.

[11] S. Dongsheng, Z. Kai, and G. Zhi, "Advances in research on soil moisture by microwave remote sensing in china," *Chin. Geogr. Sci.*, vol. 17, no. 2, pp. 186–191, Jun. 2007.

[12] Y. Kerr *et al.*, "Soil moisture retrieval from Space: The Soil Moisture and Ocean Salinity (SMOS) Mission," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 8, pp. 1729–1735, Aug. 2001.

[13] *Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond*, National Research Council, Washington, DC, USA, 2007.

[14] W. Crow and E. F. Wood, "The value of coarse-scale soil moisture observations for regional surface energy balance modeling," *J. Hydrometeorol.*, vol. 3, no. 4, pp. 467–482, Aug. 2002.

[15] G. Kim and A. P. Barros, "Space-time characterization of soil moisture from passive microwave remotely sensed imagery and ancillary data," *Remote Sens. Environ.*, vol. 81, no. 2/3, pp. 393–403, Aug. 2002.

[16] R. Bindlish and A. P. Barros, "Subpixel variability of remotely sensed soil moisture: An inter-comparison study of SAR and ESTAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 2, pp. 326–337, Feb. 2002.

[17] M. Gebremichal, R. Rigon, G. Bertoldi, and T. M. Over, "On the scaling characteristics of observed and simulated spatial soil moisture fields," *Nonlin. Process. Geophys.*, vol. 16, pp. 141–150, 2009.

[18] S. Manfreda, M. F. McCabe, M. Fiorentino, I. Rodriguez-Iturbe, and E. F. Wood, "Scaling characteristics of spatial patterns of soil moisture from distributed modeling," *Adv. Water Res.*, vol. 30, no. 10, pp. 2145–2150, Oct. 2007.

[19] O. Merlin, A. G. Chehbouni, Y. H. Kerr, E. G. Njoku, and D. Entekhabi, "A combined modeling and multipectral/multiresolution remote sensing approach for disaggregation of surface soil moisture: Application to SMOS configuration," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 9, pp. 2036–2050, Sep. 2005.

[20] A. Loew, "Impact of surface heterogeneity on surface soil moisture retrievals from passive microwave data at the regional scale: The Upper Danube case," *Remote Sens. Environ.*, vol. 112, no. 1, pp. 231–248, Jan. 2008.

[21] R. H. Reichle, D. Entekhabi, and D. B. McLaughlin, "Downscaling of radio brightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach," *Water Res. Res.*, vol. 37, no. 9, pp. 2353–2364, Sep. 2001.

[22] D. Liu and R. Pu, "Downscaling thermal infrared radiance for subpixel land surface temperature retrieval," *Sensors*, vol. 8, no. 4, pp. 2695–2706, Apr. 2008.

[23] N. S. Chauhan, S. Miller, and P. Ardanuy, "Spaceborne soil moisture estimation at high resolution: A microwaveoptical/IR synergistic approach," *Int. J. Remote Sens.*, vol. 24, no. 22, pp. 4599–4622, Jan. 2003.

[24] P. Kumar, "A multiple scale state-space model for characterizing subgrid scale variability of near-surface soil moisture," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 182–197, Jan. 1999.

[25] R. Gillies, T. Carlson, J. Cui, W. Kustas, and K. Humes, "A verification of the 'triangle' method for obtaining surface soil water content and energy fluxes from remote measurements of the normalized difference vegetation index (ndvi) and surface radiant temperature," *Int. J. Remote Sens.*, vol. 18, no. 15, pp. 3145–3166, Oct. 1997.

[26] T. Carlson, W. Capehart, and R. Gillies, "A new look at the simplified method for remote sensing of daily evapotranspiration," *Remote Sens. Environ.*, vol. 54, no. 2, pp. 161–167, Nov. 1995.

[27] O. Merlin, J. Walker, A. Chehbouni, R. Panciera, and Y. Kerr, "Towards deterministic downscaling of SMOS soil moisture using MODIS derived soil evaporative efficiency," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3935–3946, Oct. 2008.

[28] O. Merlin *et al.*, "Self-calibrated evaporation-based disaggregation of SMOS soil moisture: An evaluation study at 3 km and 100 m resolution in Catalunya Spain," *Remote Sens. Environ.*, vol. 130, pp. 25–38, Mar. 2013.

[29] M. Piles *et al.*, "Downscaling SMOS-derived soil moisture using MODIS visible/infrared data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 9, pp. 3156–3166, Sep. 2011.

[30] J. Kim and T. Hogue, "Improving spatial soil moisture representation through integration of AMSR-E and MODIS products," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 446–460, Feb. 2012.

[31] M. Anderson, J. Norman, J. Mecikalski, J. Otkin, and W. Kustas, "A climatological study of evapotranspiration and moisture stress across continental United States based on thermal remote sensing," *J. Geophys. Res.*, vol. 112, no. D10, pp. 2156–2202, May 2007.

[32] O. Merlin *et al.*, "The NAFE'06 data set: Towards soil moisture retrieval at intermediate resolution," *Adv. Water Res.*, vol. 31, no. 5, pp. 1444–1455, Nov. 2008.

[33] C. A. Gotway and L. J. Young, "Combining incompatible spatial data," *J. Amer. Statist. Assoc.*, vol. 97, no. 458, pp. 632–648, Jun. 2002.

[34] R. K. Gupta, T. S. Prasad, P. V. K. Rao, and P. M. B. Manikavelu, "Problems in upscaling of high resolution remote sensing data to coarse spatial resolution over land surface," *Adv. Space Res.*, vol. 26, no. 7, pp. 1111–1121, 2000.

[35] C. K. Wikle and L. M. Berliner, "Combining information across spatial scales," *Technometrics*, vol. 47, no. 1, pp. 80–91, Feb. 2005.

[36] J. Principe, *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*. New York, NY, USA: Springer-Verlag, 2010.

[37] J. Principe, D. Xu, and J. Fisher, *Information theoretic learning, in unsupervised adaptive filtering*. New York, NY, USA: Wiley, 2010.

[38] S. Rao, A. Martins, and J. Principe, "Mean shift: An information theoretic perspective," *Pattern Recog. Lett.*, vol. 30, no. 3, pp. 222–230, Feb. 2009.

[39] K. Nagarajan and J. Judge, "Spatial scaling and variability of soil moisture over heterogenous land cover and dynamic vegetation conditions," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 880–884, Jul. 2013.

[40] T. Bongiovanni *et al.*, "Field observations during the Eighth Microwave, Water, Energy Balance Experiment (MicroWEX-8): From June 16 through August 24, 2009," Center Remote Sens. Univ. Florida, Gainesville, FL, USA, Tech. Rep. [Online]. Available: http://edis.ifas.ufl.edu/ae476

[41] J. Casanova *et al.*, "Field observations during the Fifth Microwave, Water, Energy Balance Experiment (MicroWEX-5): From March 9 through May 2006," Center Remote Sens., Univ. Florida, Gainesville, FL, USA, Tech. Rep., 2006, Circular no. 1514.

[42] T. Lin *et al.*, "Field Observations During the Third Microwave, Water, Energy Balance Experiment (MicroWEX-3): From June 16 Through December 21, 2004," Center Remote Sens., Univ. Florida, Gainesville, FL, USA, Tech. Rep., Circular no. 1481. [Online]. Available: http://edis.ifas.ufl.edu/ae361

[43] Florida Automated Weather Network, 2011. [Online]. Available: http://fawn.ifas.ufl.edu/

[44] W. H. Brutsaert, "On a derivable formula for long-wave radiation from clear skies," *Water Res. Res.*, vol. 11, no. 5, pp. 742–744, Oct. 2010.

[45] J. Judge, A. England, J. Metcalfe, D. McNichol, and B. Goodison, "Calibration of an integrated land surface process and radiobrightness (LSP/R) model during summertime," *Adv. Water Res.*, vol. 31, no. 1, pp. 189–202, Jan. 2008.

[46] J. Judge, L. Abriola, and A. England, "Numerical validation of the land surface process component of an LSP/R model," *Adv. Water Res.*, vol. 26, no. 7, pp. 733–746, Jul. 2003.

[47] J. Judge *et al.*, "A growing season land surface process/radiobrightness model for wheat-stubble in the southern great plains," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2152–2158, Sep. 1999.

[48] B. Whitfield, J. Jacobs, and J. Judge, "Intercomparison study of the land surface process model and the common land model for a prairie wetland in Florida," *J. Hydrometeorol.*, vol. 7, no. 6, pp. 1247–1258, Dec. 2006.

[49] Y. Chung, "A snow-soil-vegetation-atmosphere-transfer/radiobrightness model for wet snow," Ph.D. dissertation, Univ. Michigan, Ann Arbor, MI, USA, 2007.

[50] J. Casanova and J. Judge, "Estimation of energy and moisture fluxes for dynamic vegetation using coupled SVAT and crop-growth models," *Water Res. Res.*, vol. 44, no. 7, pp. W07415-1–W07415-20, Jul. 2008.

[51] C. Jones and J. Kiniry, *CERES-Maize: A Simulation Model of Maize Growth and Development*. College Station, TX, USA: Texas A&M Univ. Press, 1986.

[52] J. Jones *et al.*, "The DSSAT cropping system model," *Eur. J. Agronom.*, vol. 18, no. 3/4, pp. 235–265, Jan. 2003.

[53] K. Boote, J. Jones, G. Hoogenboom, and G. Wilkerson, *Evaluation of the CROPGRO-Soybean Model Over a Wide Range of Experiments*. Boston, MA, USA: Kluwer, 1997.

[54] K. Boote and J. Jones, *Simulation of Crop Growth*. New York, NY, USA: Marcel Dekker, 1998.

[55] S. Jagtap and J. Jones, "Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production," *Agriculture, Ecosyst. Environ.*, vol. 93, no. 1–3, pp. 73–85, Dec. 2002.

[56] T. Mavromatis, K. Boote, W. James, G. Wilkerson, and G. Hoogenboom, "Repeatability of model genetic coefficients derived from soybean performance trials across different states," *Crop Sci.*, vol. 42, no. 1, pp. 76–89, Jan. 2002.

[57] J. Dardanelli *et al.*, "Use of a crop model to evaluate soil impedance and root clumping effects on soil water extraction in three argentine soils," *Trans. ASAE*, vol. 46, no. 4, pp. 1265–1275, 2003.

[58] R. Braga and J. Jones, "Using optimization to estimate soil inputs of crop models for use in site-specific management," *Trans. ASAE*, vol. 47, no. 5, pp. 1821–1831, 2004.

[59] J. Lizaso *et al.*, "Evaluating a leaf-level canopy assimilation model linked to ceres-maize," *J. Agron.*, vol. 97, no. 3, pp. 734–740, May 2005.

[60] G. Alagarswamy, K. Boote, L. Allen, and J. Jones, "Evaluating the cropgro-soybean model ability to simulate photosynthesis response to carbon dioxide levels," *J. Agronom.*, vol. 98, no. 1, pp. 34–42, Jan./Feb. 2006.

[61] J. Timsina, K. Boote, and S. Duffield, "Evaluating the cropgro soybean model for predicting impacts of insect defoliation and depodding," *J. Agronom.*, vol. 99, no. 1, pp. 148–157, Jan. 2007.

[62] J. Casanova, J. Judge, and J. Jones, "Calibration of the ceres-maize model for linkage with a microwave remote sensing model," *Trans. ASAE*, vol. 49, no. 3, pp. 783–792, May/Jun. 2006.

[63] F. Ulaby, R. Moore, and A. Fung, *Microwave Remote Sensing: Active and Passive*, vol. I. Boston, MA, USA: Artech House, 1981.

[64] T. Jackson, T. Schmugge, and J. Wang, "Passive microwave remote sensing of soil moisture under vegetation canopies," *Water Res. Res.*, vol. 18, no. 4, pp. 1137–1142, Aug. 1982.

[65] A. Fung, Z. Li, and K. Chen, "Backscattering from a randomly rough dielectric surface," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 2, pp. 356–369, Mar. 1992.

[66] P. Liu, R. DeRoo, A. England, and J. Judge, "Impact of moisture distribution within the sensing depth on l- and c-band emission in sandy soils," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 6, no. 2, pp. 887–899, Apr. 2013.

[67] F. Ulaby, R. More, and A. Fung, *Microwave Remote Sensing: Active and Passive*, vol. III. Boston, MA, USA: Artech House, 1986.

[68] T. Jackson and P. ONeill, "Attenuation of soil microwave emissivity by corn and soybeans at 1.4 and 5 ghz," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 5, pp. 978–980, Sep. 1990.

[69] W. Crow and E. Wood, "The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using Ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97," *Adv. Water Res.*, vol. 26, no. 2, pp. 137–149, 2003.
[70] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. Hoboken, NJ, USA: Wiley, 2001.
[71] B. Silverman, *Density Estimation for Statistics and Data Analysis*. London, U.K.: Chapman and Hall, 1986.
[72] A. Renyi, "Some fundamental questions of information thoery," *Sel. Papers Alfred Renyi*, vol. 2, no. 174, pp. 526–552, 1976.
[73] D. Erdogmus, "Information theoretic learning: Renyi's entropy and its applications to adaptive system training," Ph.D. dissertation, Univ. Florida, Gainesville, FL, USA, 2002.
[74] C. Huang, X. Li, and L. Lu, "Retrieving soil temperature profile by assimilating MODIS LST products with ensemble Kalman filter," *Remote Sens. Environ.*, vol. 112, no. 4, pp. 1320–1336, Apr. 2008.
[75] J. Privette *et al.*, "Early spatial and temporal validation of MODIS LAI product in the Southern African Kalahari," *Remote Sens. Environ.*, vol. 83, no. 1/2, pp. 232–243, 2002.

Subit Chakrabarti (S'08) received the B.Tech. degree in applied electronics and instrumentation engineering from Haldia Institute of Technology, West Bengal, India. He is currently working toward the Ph.D. degree in Electrical and Computer Engineering at the University of Florida, Gainesville, FL, USA.

His research includes applying machine learning techniques for spatial scaling of satellite-based observations for agricultural applications at the Center for Remote Sensing, University of Florida.

Tara Bongiovanni received the B.S. and M.E. degrees in agricultural and biological engineering from the University of Florida, Gainesville, FL, USA, in 2009 and 2012, respectively.

Her Master's research focused on modeling crop growth and development, and on improving modeled estimates of biomass and yield through assimilating remote sensing observations. She is currently working with the Center of Remote Sensing, University of Florida. Her research interests are in GIS work for drought forecasting models. She also supervises crop vegetation sampling, data collection, sensor monitoring, and data archiving for the Center.

Jasmeet Judge (S'94–M'00–SM'05) received the Ph.D. degree in electrical engineering and atmospheric, oceanic, and space sciences from the University of Michigan, Ann Arbor, MI, USA, in 1999.

She is currently the Director of the Center for Remote Sensing and an Associate Professor with the Agricultural and Biological Engineering Department, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL, USA. Her research interests are in microwave remote sensing applications to terrestrial hydrology for dynamic vegetation; modeling of energy and moisture interactions at the land surface and in the vadose zone; spatial and temporal scaling of remotely sensed observations in heterogenous landscapes; and data assimilation. She is the Chair of the National Academies' Standing Committee on Radio Frequencies and a member of the Frequency Allocations in Remote Sensing Technical Committee in the IEEE-GRSS. She also serves the American Geophysical Union as the Past Chair of the Remote Sensing Technical Committee in the Hydrology Section.

Karthik Nagarajan, photograph and biography not available at the time of publication.

Jose C. Principe (F'00) received the M.S. and Ph.D. degrees in electrical engineering from the University of Porto, Porto, Portugal; University of Florida, Gainesville, FL, USA; and the Honoris Causa degrees from the Universita Mediterranea in Reggio Calabria, Italy, Universidade do Maranhao, Maranhão, Brazil; and Aalto University, Espoo, Finland.

He is a Distinguished Professor of Electrical and Biomedical Engineering with the University of Florida, Gainesville, FL, USA. He is a BellSouth Professor and Founding Director with the University of Florida Computational Neuroengineering Laboratory. In 1987, he joined the University of Florida, after an eight-year appointment as Professor with the University of Aveiro, Aveiro, Portugal.

Dr. Principe is a Fellow of the, AIMBE (2006), IAMBE (2012), and a recipient of the INNS Gabor Award, the IEEE Engineering in Medicine and Biology Society Career Achievement Award, and the IEEE Computational Intelligence Society Neural Network Pioneer Award. In 2004, he served as President of the International Neural Network Society, as Editor in Chief of the IEEE TRANSACTIONS OF BIOMEDICAL ENGINEERING from 2001 to 2007, and as a member of the Advisory Science Board of the FDA from 2001 to 2004. He is currently the Founding Editor in Chief of the IEEE REVIEWS IN BIOMEDICAL ENGINEERING. He has been heavily involved in conference organization and several IEEE society administrative committees. He chaired 78 Ph.D. and 61 Master student committees, and he is the author of more than 600 refereed publications (five books, seven edited books, 19 book chapters, 201 journal papers, and 427 conference proceedings). He holds 22 patents and has submitted seven more.