

# Downscaling Microwave Brightness Temperatures Using Self Regularized Regressive Models

Subit Chakrabarti, *Student Member, IEEE*,

Jasmeet Judge, *Senior member, IEEE*, Anand Rangarajan *Member, IEEE*, Sanjay Ranka *Fellow, IEEE*

## Abstract

An novel algorithm is proposed to downscale microwave brightness temperatures ( $T_B$ ), at scales of 10-40 km such as those from the Soil Moisture Active Passive mission to a resolution meaningful for hydrological and agricultural applications. This algorithm, called Self-Regularized Regressive Models (SRRM), uses auxiliary variables correlated to  $T_B$  along-with a limited set of *in-situ* SM observations, which are converted to high resolution  $T_B$  observations using biophysical models. It includes an information-theoretic clustering step based on all auxiliary variables to identify areas of similarity, followed by a kernel regression step that produces downscaled  $T_B$ . This was implemented on a multi-scale synthetic data-set over NC-Florida for one year. An RMSE of 5.76 K with standard deviation of 2.8 k was achieved during the vegetated season and an RMSE of 1.2 K with a standard deviation of 0.9 K during periods of no vegetation.

## Index Terms

Microwave Remote Sensing, Brightness Temperatures, Information Theoretic Clustering, Model Selection, Kernel Regression.

arXiv:1501.07683v1 [cs.CV] 30 Jan 2015

Preprint submitted to the IEEE International Geoscience and Remote Sensing Symposium 2015.

This work was supported in part by the NASA-Terrestrial Hydrology Program (THP)-NNX09AK29G and NNX13AD04G.

S. Chakrabarti and J. Judge are with the Center for Remote Sensing, Agricultural and Biological Engineering Department, Institute of Food and Agricultural Sciences, University of Florida. A. Rangarajan and S. Ranka are with the Department of Computer % Information Science % Engineering, Gainesville, FL-32611, USA E-mail: [subitc@ufl.edu](mailto:subitc@ufl.edu)

## I. INTRODUCTION

Spatio-temporal distribution of soil moisture (SM) heavily influences atmospheric and hydrological processes. Accurate SM at spatial scales of 1-5 km is critical for agricultural applications such as drought monitoring, risk management, and productivity predictions, with major implications for food security and sustainability. Microwave observations at frequencies  $<10$  GHz are very sensitive to SM in the top 5-10 cm, due to large differences in dielectric constants of dry and wet soils and have been widely used to retrieve SM [1], [2].

Most downscaling studies have downscaled SM derived from microwave observations [3], while very few studies have downscaled satellite  $T_B$  observations directly to match model scales. Even for 5 simple soil moisture scenarios with limited variability, and ignoring errors due to model inadequacies, input parameter uncertainties and sensor calibration errors, the resulting error due to sub-footprint variability, was found to be multi-modal and, in some scenarios, had an root mean square error of  $0.09 \text{ m}^3/\text{m}^3$ . Downscaling such a biased product will only increase the final error, even if the downscaling method has built in corrections for physical and meteorological heterogeneity. Thus, downscaling  $T_B$  directly and then assimilating the downscaled product into hydrology models or crop growth models may significantly improve root zone soil moisture and crop yield estimates. Piles et. al. [4] downscaled  $T_B$  directly into SM by applying the class of methods known as Universal Triangle (UT) method to the problem and used a linear regression based linking model to relate low resolution SM to H-Pol and V-Pol  $T_B$  from the SMOS mission, and other high resolution products, aggregated to the resolution of SMOS observations. Subsequently, using the assumption that the same relationship holds at high resolution, they estimated SM at 1 km. The assumption of scale invariance is not theoretically substantiated and has been found to result in high downscaling errors, particularly during heterogeneous land cover (LC) conditions [1]. Das et. al. [5] used the correlations between fluctuations of passive radiometer observations and active radar backscatter using a time series approach to obtain a merged SM product at 9 km. Some studies have used statistical inversion techniques like linear inversion with regularization [6], Singular value decomposition (SVD) [7] and gradient descent in Banach spaces [8]. A major drawback in these approaches is the assumption of static vegetation conditions and in-footprint spatial homogeneity. Scaling algorithms based upon second-order statistics can potentially lead to significant loss of structural information in the data [2], especially under highly non-linear heterogeneous and dynamic conditions. Studies involving heterogeneous and dynamic land cover conditions, such as agricultural regions are necessary to understand the validity of downscaling algorithms over real data. In this study, a novel algorithm is presented that downscales  $T_B$  directly using auxiliary information provided by satellite derived LST, Leaf area index (LAI), precipitation (PPT) and LC at high resolution in conjunction with a very limited set of *in-situ* SM observations. The *in-situ* SM observations are converted into  $T_B$  observations using a biophysical model. In the absence of *in-situ* SM observations, airborne or *in-situ*  $T_B$  observations can also be used for training. In areas without any high resolution data, representative data from other regions with similar topography, slope and soil texture can be used. This algorithm, based on self-regularized regressive models (SRRM) does not make any assumptions about the vegetation conditions and uses higher order statistical descriptors to downscale, and quantify the efficacy of downscaling. The goal of this study is to implement a downscaling algorithm that disaggregates coarse-scale remotely sensed products with auxiliary fine-scale data. The primary objectives are to, 1) estimate  $T_B$  at 1 km using  $T_B$  at 10 km and other spatially correlated variables for a multi-scale synthetic dataset [9] based in North-Central Florida, and 2) to conduct a thorough statistical analysis of the downscaled  $T_B$  in order to evaluate the efficacy of the SRRM downscaling algorithm.

## II. THEORY

In this study, shown as a flow diagram in Figure 1, a number of models are created dynamically based on generalized proximity regions in the high dimensional correlated data. The membership of a pixels to a model is fuzzy, constrained to a sum of one across the model space. The models itself are trained using a kernel regression based method. The first step of the algorithm, which clusters the study area into proximity regions, is described below. A clustering algorithm which uses information theoretic measures of inter and intra cluster similarity is used [10]. The membership vectors are not discretized after each step to decrease downscaling errors along cluster boundaries. The cost-function is regularized using the weighted Shannon entropy of the membership vector, such

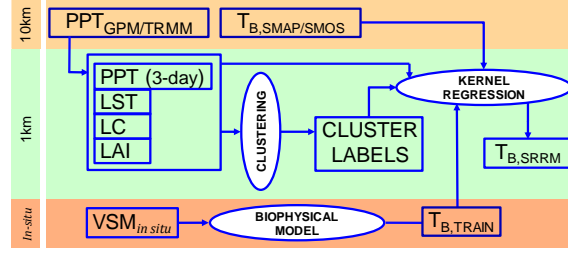


Fig. 1. Flowchart of the SRRM algorithm

that the membership vectors are sufficiently sparse, as follows.

$$\begin{aligned} \hat{j}_{CS}^{REG} = & \frac{\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (1 - \mathbf{m}_i^T \mathbf{m}_j) G_{\sigma\sqrt{2}}(\mathbf{x}_i, \mathbf{x}_j)}{\sqrt{\prod_{k=1}^K \sum_{i=1}^N \sum_{j=1}^N m_{ik} m_{jk} G_{\sigma\sqrt{2}}(\mathbf{x}_i, \mathbf{x}_j)}} \\ & - \mu \sum_{i=1}^N \sum_{k=1}^K m_{ik} \log(m_{ik}) \end{aligned} \quad (1)$$

Getting the correct membership vector then is equivalent to solving this constrained optimization problem:

$$\begin{aligned} \min_{\mathbf{m}_1, \dots, \mathbf{m}_N} & \hat{J}_{CS}^{REG}(\mathbf{m}_1, \dots, \mathbf{m}_N) \\ \text{subject to} & \mathbf{m}_j^T \mathbf{1} - 1 = 0, \quad j = 1, \dots, N \end{aligned} \quad (2)$$

This can be solved using Lagrange multipliers and a stochastic gradient descent scheme to compute optimum values of  $\mathbf{m}$ . In the second step, a kernel based regression technique that uses a training set of pixels and fits a function to it, by minimizing the representational error, is used to generate the downscaled estimates. Ridge regression is a parametric regression technique that adds a scaled regularizing term to the cost function to increase generalization. The cost function for ridge regression is

$$\mathcal{E}(\mathbf{w}, \mathbf{x}) = \frac{1}{2} \sum_i (y_i - \mathbf{w}^T \mathbf{x}_i)^2 + \frac{1}{2} \mu \|\mathbf{w}\|^2 \quad (3)$$

The weights can be calculated by differentiating the error cost function with respect to the weights and setting it to zero. If this computation was carried out in a Reproducing Kernel Hilbert Space (RKHS), then the inner-products can be replaced with a kernel evaluation. Let  $\mathcal{H}$  be a Hilbert space with an inner-product metric  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ . Then according to the representer theorem, a kernel function  $\kappa(\mathbf{x}, \mathbf{y})$  exists on  $\mathbb{R}^N \times \mathbb{R}^N$  such that  $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{H}} = \kappa(\mathbf{x}, \mathbf{y})$ . Now, if  $\Phi: \mathbb{R}^N \rightarrow \mathbb{R}^N$  is a mapping that transforms the feature vector in the original vector space to  $\mathcal{H}$ , then the weights can be redefined as,

$$\mathbf{w} = (\mu \mathbf{I}_D + \Phi \Phi^T)^{-1} \Phi \mathbf{y} \quad (4)$$

Then, the estimated value of  $y$  for a new data-point  $\mathbf{x}'$  is,

$$\hat{y} = \mathbf{w}^T \Phi(\mathbf{x}') = \mathbf{y} (\mu \mathbf{I}_N + \mathbf{K})^{-1} \kappa(\mathbf{x}, \mathbf{x}') \quad (5)$$

where  $\mathbf{K}$  is the Gram matrix of inner products of all the training data points. This does not address the constant that must be present in the regression. To solve this problem, the feature vector is augmented by adding a constant feature 1 to all samples.

### III. METHODOLOGY

#### A. Multiscale synthetic dataset

The proposed algorithm for disaggregation was tested using data generated by a simulation framework consisting of the Land Surface Process (LSP) model and the Decision Support System for Agrotechnology Transfer (DSSAT)

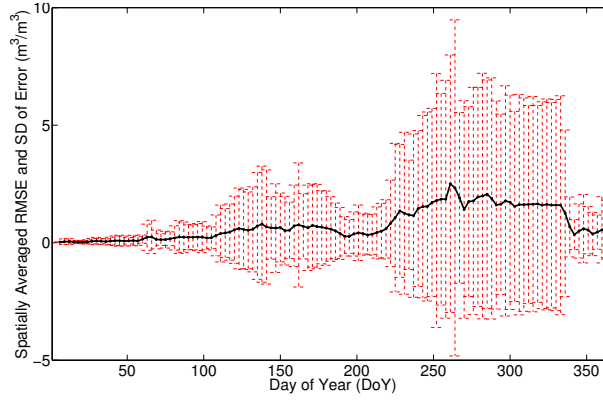


Fig. 2. Root Mean Square Error (RMSE) and Standard Deviation (SD) between true  $T_B$  and downscaled  $T_B$  at 1 km.

model, described in [9]. A  $50 \times 50 \text{ km}^2$  region, equivalent to 25 SMAP pixels, was chosen in North Central Florida for the simulations.

Based upon land cover information at 200 m, contiguous, homogeneous regions of sweet-corn and cotton were identified. A realization of the LSP-DSSAT model was used to simulate 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. Linear averaging is typically sufficient to illustrate the effects of resolution degradation [11]. The model simulations at 200 m were spatially averaged to obtain PPT, LST, LAI, and SM at 1 and 10 km. To simulate rain-fed systems, all the water input from both precipitation and irrigation were combined together, and the “PPT” in this study represents these combined values, representing a rainfed system.

$T_B^{1\text{km}}$  measurements, were generated from  $SM_i^{1\text{km}}$  for the entire area, using the widely used  $\tau - \omega$  model [12]. 10 % of the  $T_B$  measurements were used for training. The rest was used as “truth” for validating the downscaled  $T_B$ . An alternative and better approach is to convert the downscaled  $T_B$  into SM using an inverse model or assimilation [13], since SM is the key variable required at high resolution. However, that detracts from the focus of this study by adding an extra layer of complexity. A vegetated surface is modeled as a single isothermal layer of vegetation with diffuse boundaries [9]. The soil medium is assumed to be non-isothermal, semi-infinite layered dielectric medium with a rough surface at the upper boundary. The soil moisture 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  $T_B$  of a terrain is the sum of contributions from soil, vegetation, and from sky. In this study, similar to [1] the  $\theta$  was set to  $50^\circ$ , the  $T_{sky}$  was set to 5K,  $r_p$  was obtained by integrating the bistatic scattering coefficients from the IEM model, the RMS height was 0.62 cm and, the correlation length was 8.72 cm [12]. Other model parameters are set similar to [9].

### B. Implementation of Disaggregation Framework based on Self-Regularized Regressive Models

The simulation period, from Jan 1 (DoY 1) to Dec 31 (DoY 365), 2007, consisted of two growing seasons of sweet corn and one season of cotton. The LST, PPT, and LAI observations at 1 km were obtained by adding noise to account for satellite observation errors, instrument measurement errors, and micro-meteorological variability, following [14]. Errors with zero mean and standard deviations of 5K, 1 mm, and 0.1 for LST, PPT, and LAI, respectively, were added to the values at 10 k.

The SRRM method uses LST, 3-day PPT, LAI, LC at 1 km and  $T_B$  at 10 km every 3 days as input. In the first step, the field is clustered using the inputs at 1 km and the x and y coordinates of each pixel scaled to a range of 0 and 1. This step of the algorithm uses two parameters - the number of clusters,  $n$  and a regularization constant,  $\mu$ . Both the number of clusters and the regularization constant is determined by cross-validating against the absolute mean error in  $T_B$  at the end of the second step for each day. The optimal number of iterations that produces a usable clustering result is determined by evaluating the root mean square error (RMSE) for Day 222, characterized by maximum input heterogeneity, in disaggregated  $T_B$  after every iteration, for upto 200 iterations. At the end of this step, each pixel has a vector of  $n$  numbers,  $(m_1, m_2, \dots, m_N)$  that sum upto 1 describing its membership to each of the  $n$

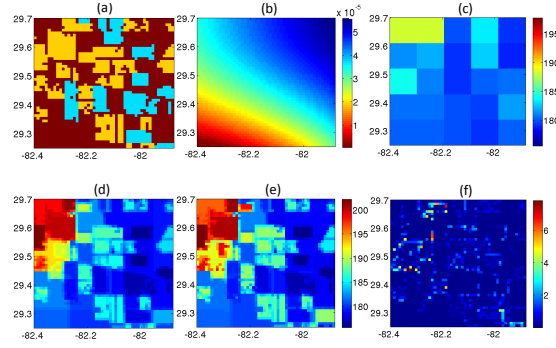


Fig. 3. Day 195 - (a) LC at 1 km (blue represents baresoil, yellow represents sweet-corn and cyan represents cotton), (b) PPT at 1 km (in  $\text{m}^3/\text{m}^3$ ), (c)  $T_B$  at 10 km, (d) True  $T_B$  at 1 km (in K), (e) Downscaled  $T_B$  at 1 km (in K), (f) Absolute Difference Plot between True and Downscaled  $T_B$  (in K). X-axis denotes longitudes ( $^\circ\text{W}$ ) and Y-axis denotes latitudes ( $^\circ\text{N}$ ) for every plot.

clusters. In the second step,  $n$  models,  $\hat{f}_1, \hat{f}_2, \dots, \hat{f}_N$  are developed using LST, 3-day PPT, LAI, LC,  $T_B$  at 1 km and  $T_B$  at 10 km as inputs to the regularized kernel regression algorithm described in Section II, using 10% of the pixels that make up the field. The hard membership of each pixel,  $i$ , for model development purposes is determined by the maximum value in its membership vector,  $\mathbf{m}^i = (m_1^i, m_2^i, \dots, m_N^i)$ . The disaggregated value of  $T_B$  is computed for each point in the test, represented as a vector,  $\mathbf{x}'_i = (\text{LST}_i^{1\text{km}}, \text{PPT}_i^{1\text{km}}, \text{LAI}_i^{1\text{km}}, \text{LC}_i^{1\text{km}}, T_B^{1\text{km}})$  by,

$$T_{B_i}^{1\text{km}} = \mathbf{m}^T \cdot (\hat{f}_1(\mathbf{x}'_i), \hat{f}_2(\mathbf{x}'_i), \dots, \hat{f}_N(\mathbf{x}'_i)) \quad (6)$$

The SRRM method is evaluated by studying the RMSE and standard deviation of the errors over the entire season. Moreover, the downscaled  $T_B$  is plotted versus the true  $T_B$ , for each land-cover. To evaluate how close the density function of the downscaled estimates is to the density function of the true SM, the KL-Divergence (KLD) between the density of the estimated observations and the true  $T_B$  is calculated for different LC's over the season. The PPT, LC, low resolution  $T_B$ , true  $T_B$  and downscaled  $T_B$  are examined for representative days when the crop variability and micro-meteorological variability was highest. Quantitative analyses of spatial variations in  $T_B$  observed under dynamic vegetation and heterogeneous land cover conditions provide an index of dynamic errors that can be expected.

#### IV. RESULTS

The RMSE and SD of error between true and downscaled  $T_B$  at 1 km is shown in Figure 2. The standard deviation increases from below 3 K to almost 7 K during the vegetated season. This is due to a large difference in errors with most of the errors originating from pixels on the field boundaries. A Z-test was performed which showed that the errors lie under 10 K overall for more than 95% of the pixels. They were found to be the highest during the growing season for sweet-corn and cotton, and the late season baresoil land cover after harvest, due to residual vegetation after harvest. Figure 4 shows the errors for each land-cover. The errors are high, during the vegetated season, for baresoil land-cover due to sub-pixel heterogeneities. Pixels at 1 km are classified as baresoil when less than half of the area is cultivated. These impure pixels contribute to the high error. This is also evident in Figure 3, which shows the LC, PPT,  $T_B$  at 10 km, true  $T_B$ , downscaled  $T_B$ , and the absolute difference between true and downscaled  $T_B$  at 1 km for Day 195. The absolute difference plot indicates that the maximum errors occur at the edges of the fields which confirms the impure pixel hypothesis. The true and downscaled  $T_B$  are very similar which shows that the algorithm can perform well during heterogeneous micro-meteorological conditions.

The RMSE, SD and KL-Divergence (KLD) over the whole season is shown in Table I. Baresoil pixels at the end of the season affected by remnant crops and impure pixels have a higher KLD, but very close to 0. Vegetated pixels at 1 km have a higher KLD as well. Vegetated pixels with baresoil impurities contribute to these errors. As expected the RMSE is lowest during early season baresoil, when the downscaled estimates are almost perfect. The RMSE's and SD's increase during the growing season, which is expected due to the complexity introduced by different land-cover types.

TABLE I

RMSE, SD, AND KL DIVERGENCE OVER THE  $50 \times 50 \text{ km}^2$  REGION FOR THE DOWNSCALED ESTIMATES OF SM OBTAINED AT 1 KM USING THE PRI AND UT METHODS. (A - BARESOIL PIXELS WITH VEGETATED SUB-PIXELS AT 250 M TILL DOY 332, B - BARESOIL PIXELS AFTER DOY 332 AND C - BARESOIL PIXELS WITHOUT ANY VEGETATED SUB-PIXELS AT 250 M TILL DOY 332)

Land Cover	$\text{KLD}_{\text{SRRM}}$	$\text{RMSE}_{\text{SRRM}}$	$\text{SD}_{\text{SRRM}}$
Corn	$1.8615 \times 10^{-17}$	4.1 K	2.5 K
Cotton	$2.4828 \times 10^{-04}$	3.7 K	3.4 K
Baresoil <sup>A</sup>	$5.6222 \times 10^{-5}$	7.7 K	1.51 K
Baresoil <sup>B</sup>	$5.628 \times 10^{-6}$	8.1 K	1.8 K
Baresoil <sup>C</sup>	$2.5948 \times 10^{-6}$	1.2 K	0.9 K

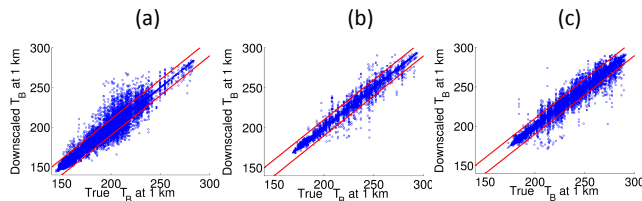


Fig. 4. Disaggregated  $T_B$  vs. True  $T_B$  at 1 km during the whole season for (a) baresoil pixels (b) corn pixels, and (c) cotton pixels. Lines corresponding to 10 K error in  $T_B$  are shown for each plot.

## V. CONCLUSION

In this study, a downscaling methodology for  $T_B$ , based upon SRRM models, was implemented and evaluated to work with an RMSE of 5.76 K with standard deviation of 2.8 k, during the vegetated season, and an RMSE of 1.2 K with a standard deviation of 0.9 K during periods of no vegetation. Thus this algorithm will be particularly useful during heterogeneous LC and micro-meteorological conditions, with limited set of training observations available. It is envisioned that this algorithm will be applied using the SM observations available from the soon to be launched NASA-Soil Moisture Active Passive (SMAP) mission in conjunction with remotely sensed LST, LAI and PPT data from other satellites in the Earth Observing System. Future scope of research includes data-fusion of SMOS observations of  $T_B$  at multiple angles with SMAP  $T_B$  to increase the information content in the low resolution product for better downscaling accuracies, and using variational data-assimilation to calculate SM from the downscaled  $T_B$ .

## REFERENCES

- [1] S. Chakrabarti, T. Bongiovanni, J. Judge, K. Nagarajan, and J. C. Principe, "Downscaling satellite-based soil moisture in heterogeneous regions using high-resolution remote sensing products and information theory: A synthetic study," *IEEE Trans. Geosci. Remote Sensing*, vol. 53, no. 1, pp. 85–101, 2014.
- [2] S. Chakrabarti, T. Bongiovanni, J. Judge, L. Zotarelli, and C. Bayer, "Assimilation of smos soil moisture for quantifying drought impacts on crop yield in agricultural regions," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 7, no. 9, pp. 3867–3879, 2013.
- [3] O. Merlin, M. Escorihuela, M. Mayoral, O. Hagolle, A. A. Bitar, and Y. Kerr, "Self-calibrated evaporation-based disaggregation of smos soil moisture: an evaluation study at 3 km and 100 m resolution in catalunya spain," *Remote Sens. Env.*, vol. 130, no. 0, pp. 25–38, 2013.
- [4] M. Piles, M. Vall-llossera, L. Laguna, and A. Camps, "A downscaling approach to combine smos multi-angular and full-polarimetric observations with mdis vis/ir data into high resolution soil moisture maps," in *IEEE Int. Geosc. and Rem. Sens. Symposium*, vol. 1. Proc. IGARSS 2012, 2012, pp. 1247–1250.
- [5] N. Das, D. Entekhabi, and E. Njoku, "Tests of the smap combined radar and radiometer algorithm using airborne field campaign observations and simulated data," *IEEE Trans. Geosci. Remote Sensing*, vol. 52, no. 4, pp. 2018–2028, 2014.
- [6] Y. Wang, J. Shi, L. Jiang, J. Du, and B. Tian, "The development of an algorithm to enhance and match the resolution of satellite measurements from amsr-e," *Science China: Earth Sciences*, vol. 54, no. 3, pp. 410–419, 2011.
- [7] A. Gambardella and M. Migliaccio, "On the superresolution of microwave scanning radiometer measurements," *IEEE Geosci. and Remote Sensing Letters*, vol. 5, no. 4, pp. 796–800, 2008.
- [8] F. Lenti, F. Nunziata, C. Estatico, and M. Migliaccio, "On the spatial resolution enhancement of microwave radiometer data in banach spaces," *IEEE Trans. Geosci. Remote Sensing*, vol. 52, no. 3, pp. 1834–1842, 2014.
- [9] K. Nagarajan and J. Judge, "Spatial scaling and variability of soil moisture over heterogenous land cover and dynamic vegetation conditions," *IEEE Geosci. and Remote Sensing Letters*, vol. 10, no. 4, pp. 880–884, 2013.

- [10] R. Jenssen, D. Erdogmus, K. Hild, J. Principe, and T. Eltoft, "Optimizing the cauchy-schwarz pdf distance for information theoretic, non-parametric clustering," in *Proceedings of the 5th international conference on Energy Minimization Methods in Computer Vision and Pattern Recognition*, vol. 1. Proc. EMMCVPR 2005, 2005, pp. 34–45.
- [11] 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. in Water Res.*, vol. 26, no. 2, pp. 137–149, 2003.
- [12] 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, 2013.
- [13] 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. Research*, vol. 37, no. 9, pp. 2353–2364, 2001.
- [14] C. Huang, X. Li, and L. Lu, "Retrieving soil temperature profile by assimilating MODIS LST products with ensemble Kalman filter," *Remote Sens. Env.*, vol. 112, pp. 1320–1336, 2008.