

A Machine Learning-Based Geostatistical Downscaling Method for Coarse-Resolution Soil Moisture Products

Yan Jin , Yong Ge , *Member, IEEE*, Yaojie Liu , Yuehong Chen, Haitao Zhang, and Gerard B. M. Heuvelink 

Abstract—The surface soil moisture (SSM) products derived from microwave remote sensing have a coarse spatial resolution; therefore, downscaling is required to obtain accurate SSM at high spatial resolution. An effective way to handle the stratified heterogeneity is to model for various stratifications; however, the number of samples is often limited under each stratification, influencing the downscaling accuracy. In this study, a machine learning-based geostatistical model, which combines various kinds of ancillary information at fine spatial scale, is developed for spatial downscaling. The proposed support vector area-to-area regression kriging (SVATARK) model incorporates support vector regression and area-to-area kriging by considering the nonlinear relationships among variables for various stratifications. SVATARK also considers the change of support problem in the downscaling interpolation process as well as for solving the small sample size in trend prediction. The SVATARK method is evaluated in the Naqu region on the Tibetan Plateau, China, to downscale the European Space Agency's (ESA) 25-km-resolution SSM product. The 1-km-resolution SSM predictions have been produced every eight days over a six-year period (2010–2015). Compared with five other downscaling methods, the downscaled predictions from the SVATARK method performs the best with *in situ* observations, resulting in a 24.4% reduction in root-mean-square error with $0.08 \text{ m}^3 \cdot \text{m}^{-3}$ and a 8.2% increase in correlation coefficient with 0.72, on average. Additionally, anomalously low SSM values, an indicator of drought, had a record low anomaly in mid-July for 2015, as noted by previous studies, indicating that SVATARK could be utilized for drought monitoring.

Index Terms—Area-to-area kriging, downscaling, soil moisture, support vector regression.

I. INTRODUCTION

SURFACE soil moisture (SSM) is an essential hydroecological parameter for monitoring energy, water, and carbon cycles [1], [2]. Continuous SSM at fine spatial resolutions provides crucial information for hydrological models, precipitation forecasting models, land-atmosphere models, drought and flood forecasting, and vegetation growth monitoring [3]–[5]. In general, soil moisture is acquired by using *in situ* measurements [6], including wireless sensor networks [7] and cosmic-ray soil moisture observing system [8], which have helped overcome the sparse sampling and poor dynamic limitations of traditional *in situ* methods. These ground-based measurement methods require suitable ground conditions and are limited to small monitoring areas. With the development of active and passive microwave remote sensing techniques [9], it becomes possible and more convenient to acquire SSM information dynamically at different spatiotemporal resolutions over large areas. A series of SSM products derived from various satellite-based microwave sensors has been released [10]–[12]. However, with spatial resolutions of tens of kilometers, the current microwave-based SSM products are limited to large-scale monitoring applications. Most hydrological applications need a high-resolution SSM product which would enhance the knowledge of the hydrological processes at local scale [13]. In consideration of the spatial representativeness and overall accuracy, the 1-km spatial resolution has become a common high resolution in many SSM downscaling researches [13], [14].

Many approaches have been developed for downscaling these coarse-scale SSM products. Some of these benefit from ancillary information that captures the variations of SSM at fine resolution, combined with correlated variables [15]. There are two main sources of ancillary variables—active microwave data and visible/infrared data. Change detection-based downscaling algorithms [16], [17] and Bayesian merging methods [18], [19] have been proposed to downscale the coarse SSM by using active microwave data. The active microwave technique is highly sensitive to SSM and can even penetrate clouds; however, it is greatly affected by soil roughness and vegetation. An alternative downscaling approach is to use fine resolution optical/thermal data. A number of downscaling algorithms have been developed

Manuscript received June 11, 2020; revised August 14, 2020 and October 4, 2020; accepted October 23, 2020. Date of publication November 3, 2020; date of current version January 6, 2021. This work was supported in part by the National Natural Science Foundation for Distinguished Young Scholars of China under Grant 41725006, in part by the National Natural Science Foundation of China under Grant 42001332, Grant 42001375, Grant 41531174, and Grant 41531179, in part by the Strategic Priority Research Program of Chinese Academy of Sciences under Grant XDA20030302, and in part by the Scientific Research Fund of Nanjing University of Posts and Telecommunications under NUPTSF Grant NY219035. (*Corresponding author: Yong Ge.*)

Yan Jin and Haitao Zhang are with the School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing 210023, China (e-mail: jinyan@njupt.edu.cn; zhanghaitao@njupt.edu.cn).

Yong Ge is with the State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographic Sciences and Natural Resources Research, Beijing 100101, China (e-mail: gey@lreis.ac.cn).

Yaojie Liu is with the International Institute for Earth System Science, Nanjing University, Nanjing 210023, China (e-mail: liuyaojie@nju.edu.cn).

Yuehong Chen is with the School of Earth Sciences and Engineering, Hohai University, Nanjing 210098, China (e-mail: yuehong@hhu.edu.cn).

Gerard B. M. Heuvelink is with Soil Geography and Landscape Group, Wageningen University, 6700 AA Wageningen, The Netherlands (e-mail: gerard.heuvelink@wur.nl).

Digital Object Identifier 10.1109/JSTARS.2020.3035386

to generate fine-resolution SSM, such as Disaggregation based on Physical And Theoretical scale Change (DISPATCH) [20], [21], trapezoid-based methods [22], [23], regression-based approaches [24], [25], and geostatistical methods [26], [27]. For downscaling with optical/thermal data, the statistical correlation between SSM and ancillary variables or physically-based models have been explored [14].

Chauhan *et al.* [28] proposed an empirical polynomial fitting downscaling approach using a polynomial regression at coarse spatial resolution to obtain the fine-spatial-resolution SSM. Since then, further polynomial fitting downscaling methods have been presented by employing multiple data sources or different ancillary parameters [29], [30], such as land surface temperature (LST), vegetation information, brightness temperature, albedo, evapotranspiration, and terrain indices. Meanwhile, geographically weighted regression, which takes into consideration local characteristics [31], and machine learning algorithms have been introduced into downscaling. Machine learning algorithms such as random forest and support vector regression (SVR) perform better in capturing the nonlinear relationships among variables and have already been widely applied to enhance the spatial resolution of coarse-resolution images, including super-resolution land cover mapping [32] and downscaling SSM [33], [34]. Some studies directly combined the fine-resolution trend and coarse-resolution residual to predict the fine-resolution SSM [34], [35]. Interpolation techniques such as bilinear interpolation and kriging interpolation have been generally used in residual analysis for approximating the actual fluctuations [37], [38]. Geostatistical methods with a focus on the spatial correlation between variables have been increasingly applied in downscaling [39], [40]. However, these downscaling approaches ignore the change in supports before and after downscaling. Moreover, due to the stratified heterogeneity [15], the geographical variable would have patch-based variation in the variability, which needs to capture the properties in each stratification (i.e., subregion). The downscaling models for various stratifications established in the scaling process are limited to the smaller samples captured by the model. The SVR approach possesses a strong adaptability, global optimization, and a high generalization ability, and is also suitable for small samples of data [41], [42]. It could provide a solution to the small sample size problem.

Considering all the previous machine learning algorithms, this article proposes a new machine learning-based geostatistical model that integrates SVR and area-to-area kriging (ATAK, which changes the supports before and after the interpolation [43]) to achieve spatial downscaling by fusing various ancillary variables. The proposed support vector area-to-area regression kriging (SVATARK) can tackle the modifiable areal unit problem, as well as model the complex nonlinear relationship among variables in the downscaling process. The downscaling approach was employed to predict 1-km-resolution SSM data by downscaling ESA's 25-km-resolution SSM product, climate change initiative (CCI), with consideration of land cover types. Downscaled SSM predictions were produced every eight days over the Naqu region in the central Tibetan Plateau (TP), and were evaluated using *in situ* SSM measurements. A comparison of the SSM residuals obtained from the ATAK method versus the

residuals from bilinear interpolation and kriging interpolation indicated advantages of the SVATARK downscaling approach, which was also verified by comparing with two other statistical downscaling methods and a physically-based downscaling model.

The remainder of the article is organized as follows. Section II describes the downscaling methodology, including the downscaling strategy during the experiment. Both the study area and the datasets are introduced in Section III. Section IV validates the downscaled predictions and discusses the comparison results. Finally, some conclusions are summarized in Section V.

II. METHODOLOGY

The proposed SVATARK downscaling method mainly consists of both trend and residual models. In this section, we briefly describe the downscaling model components SVR and ATAK and the experimental downscaling scheme.

A. Support Vector Regression Method

Support vector machines (SVMs) have been widely applied to classification and regression, which minimize both empirical risk and structural risk to seek the best compromise between the complexity and learning capability of a model [41], [42]. For regression, SVR was first introduced by Vapnik *et al.* [44].

Let $\chi = \{x_i, y_i; i = 1, \dots, n\}$ be the training dataset with ancillary vectors x_i and corresponding targets y_i . The input space χ can be mapped into some feature space Φ using the nonlinear function $\varphi = \chi \rightarrow \Phi$. In the feature space Φ , the training data may exhibit linearity, which can be approximated by linear regression. The general form of the nonlinear SVR function can be expressed as

$$f(\omega, b) = \omega \cdot \varphi(x) + b \quad (1)$$

where ω and b are the parameter vectors. The kernel function $K(x_i, x_j) = \langle \varphi(x_i) \cdot \varphi(x_j) \rangle$ can be used to calculate the inner products in the feature space Φ . By introducing α_i and $\bar{\alpha}_i$ in the dual form to solve the optimization problem in SVR, the regression function of the nonlinear SVR allowing the kernel function is expressed as

$$f(x_i) = \sum_{k=1}^n \left(\alpha_k - \bar{\alpha}_k \right) K(x_i, x_k) + b. \quad (2)$$

More details about the nonlinear SVR can be found in Smola and Schölkopf [45]. It is well known that the kernel function and its hyperparameters have a great impact on the performance of nonlinear SVR model. In our study, ϵ -SVR is used with the Gaussian radial basis function as its kernel function. The relevant penalty coefficient (C) and gamma (γ) can be optimized by minimizing the model error. The SVR was implemented in R "e1071" package [46]. The epsilon in SVR models was taken as 0.1, whereas the C ranged from 8 to 64 and γ ranged from 32 to 128. Owing to the stratified heterogeneity, the SVR models are established for different land cover types, considering that different underlying surfaces might influence the relationship among SSM and ancillary variables.

B. Area-to-Area Kriging Method

The area-to-area kriging is a case of areal interpolation, which changes the supports before and after the interpolation [43]. A linear combination of areal data is used to predict other areal values. The target areal value z over a given unit u_α is estimated with the K neighboring observations at units u_i

$$z(u_\alpha) = \sum_{i=1}^K \lambda_i(u_\alpha) \cdot z(u_i) \quad (3)$$

where $\lambda_i(u_\alpha)$ is the weight assigned to $z(u_i)$, which can be calculated by minimizing the prediction error variance. The corresponding kriging system is written as

$$\begin{cases} \sum_{j=1}^K \lambda_j(u_\alpha) \cdot \bar{C}(u_i, u_j) + \mu(u_\alpha) = \bar{C}(u_i, u_\alpha), i = 1, \dots, K \\ \sum_{j=1}^K \lambda_j(u_\alpha) = 1 \end{cases} \quad (4)$$

where $\mu(u_\alpha)$ is the Lagrange multiplier, $\bar{C}(u_i, u_j)$ and $\bar{C}(u_i, u_\alpha)$ are block-to-block covariance terms. The most important step for the implementation of ATAK is to obtain the point support covariance for deriving the covariance terms. A deconvolution procedure can be used to achieve the point support covariance [47]. In our study, 25 nearest neighboring pixels were employed to predict the target area of ATAK.

C. Support Vector Area-to-Area Regression Kriging

The proposed SVATARK is based on SVR for trend prediction and ATAK for residual prediction. The residual is the difference between the observed value and the estimated value, such as, coarse residual is defined as the discrepancy between the original input 25-km SSM and aggregated 25-km SSM from 1-km trend predictions in this article. Let $Z(S_i)$ and $X_k(S_i)$ be the target and k ancillary random variables at coarse pixel S_i . The nonlinear regression model between $Z(S_i)$ and $X_k(S_i)$ can be obtained using (2), denoted by $f_{\text{SVR}}(\cdot)$. Assuming that the statistical relationship among variables is scale-invariant, the trend component of the fine spatial resolution can be estimated by using the coarse regression function

$$m(s_j) = f_{\text{SVR}}(x_k(s_j)) \quad (5)$$

where $x_k(s_j)$ represent k ancillary variables of fine pixel s_j .

The residual component of the fine spatial resolution is estimated using (3), interpolating the coarse residual with I neighboring coarse pixels $e(S_i)$

$$\begin{aligned} e(s_j) &= \sum_{i=1}^I \lambda_i(s_j) \cdot e(S_i) \\ &= \sum_{i=1}^I \lambda_i(s_j) \cdot [Z(S_i) - f_{\text{SVR}}(X_k(S_i))] \end{aligned} \quad (6)$$

where $\lambda_i(s_j)$ are the weights assigned to I neighboring coarse pixels for the prediction at fine resolution.

Combining (5) and (6), the SVATARK downscaling model prediction $z(s_j)$ can be expressed as

$$\begin{aligned} z(s_j) &= m(s_j) + e(s_j) = f_{\text{SVR}}(x_k(s_j)) \\ &+ \sum_{i=1}^I \lambda_i(s_j) \cdot [Z(S_i) - f_{\text{SVR}}(X_k(S_i))] \end{aligned} \quad (7)$$

D. Downscaling Strategy

To illustrate the different performances of various downscaled predictions, five downscaling methods have been adopted to compare with the proposed SVATARK, including support vector regression kriging (SVRK), support vector regression with interpolation (SVRB), K -nearest neighbors (KNN) [24], geologically weighted regression kriging (GWRK) [25], and DISPATCH [20]. The six downscaling methods were used to downscale the CCI SSM product over 36 months (during May–October, 2010–2015). LST, normalized difference vegetation index (NDVI), land cover (LC), blue sky albedo (BSA), digital elevation model (DEM), aspect, and slope were directly used as ancillary variables in five statistical downscaling methods (i.e., SVATARK, SVRK, SVRB, KNN, and GWRK). As a physical-based downscaling method, the DISPATCH relies on an SSM proxy term to model the spatial variability of SSM over the coarse-resolution pixels, where the soil evaporative efficiency (SEE) indirectly estimated from LST and NDVI was employed in this article [13], [48]. Considering the relatively low coverage of daily remotely sensed observations, the eight-day composites of all variables were employed by using average aggregation to maintain stability and representativeness of each variable. To deal with the missing value issue in remote sensing data introduced by cloud cover, sensor failure, and so on, various gap filling methods have been developed, and a spatial–temporal prediction method [49] was adopted to replace the missing values for LST and BSA in this article. Prior to performing the downscaling algorithm, a bias correction step [40] was used for remotely sensed SSM data to reduce the influence of the original SSM product taking into account its discrepancy compared to ground observations. The downscaling procedure is shown in Fig. 1, including the downscaling and validation processes.

The entire implementation process of the downscaling strategy is shown in following steps.

Step 1): Prepare the SSM data and ancillary variables at different spatial resolution. All of the input data are prepared by resampling, aggregation, gap filling, and bias correction, resulting in the 25- and 1-km variables with full spatial coverage for each eight-day period as well as the eight-day *in situ* measurements within 1×1 km grids. The eight-day *in situ* measurements could be obtained by averaging the *in situ* SSM values within eight days, and then would be aggregated into 1×1 km grids.

Step 2): Establish regression models at 25-km spatial resolution for five statistical downscaling methods. Three trend models [i.e., SVR, KNN, and geologically weighted regression (GWR)] are established by employing the 25-km SSM and ancillary variables, in which the SVR models are trained for each land cover type [details in Fig. 2(b)] without two classes (i.e., water bodies and permanent snow and ice).

Step 3): Predict SSM trend at fine resolution for five statistical downscaling methods. The various regression models are applied to predict 1-km SSM trend part by using 1-km ancillary variables.

Step 4): Calculate the coarse residuals for five statistical downscaling methods. The 1-km trend predictions are aggregated

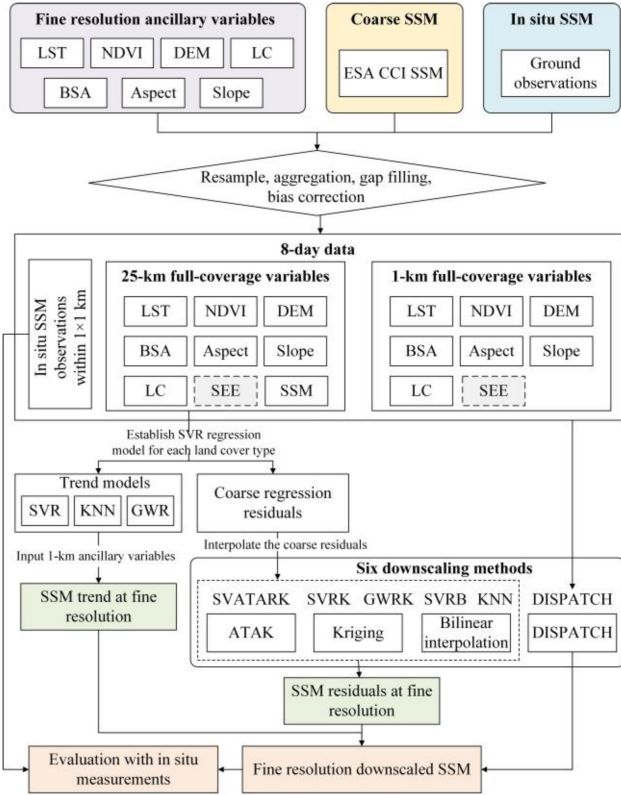


Fig. 1. Flowchart of the downscaling procedure in this study. (SEE was employed in DISPATCH method.)

into 25×25 km, which is compared with coarse SSM values for obtaining the coarse regression residuals.

Step 5): Predict SSM residuals at fine resolution for five statistical downscaling methods. The coarse residuals of SSM are interpolated into 1×1 km regular grids by using ATAK, kriging, and bilinear interpolation for different downscaling methods, which correspond to SVATARK, SVRK, GWRK, SVRB, and KNN, respectively.

Step 6): Generate the 1-km SSM and validate the downscaled results. For five statistical downscaling methods, the downscaled SSM values can be acquired by combining the 1-km trend and residual predictions. For DISPATCH method, the downscaled SSM can be derived from the relationship between SSM and SEE (more details in [13] and [48]). The above-mentioned six downscaled predictions are validated by ground-measured SSM with four classical statistical metrics, including correlation coefficient (R), mean absolute error (MAE) ($m^3 \cdot m^{-3}$), root-mean-square error (RMSE) ($m^3 \cdot m^{-3}$), and slope (SLOP) of linear regression between ground observations and downscaled predictions. In the experiments, SSM values for water bodies and permanent snow and ice were not included.

III. STUDY AREA AND DATA DESCRIPTION

A. Study Area

The study area is $3 \times 3^\circ$ ranging from 30.0° to 33.0° N and 90.5° to 93.5° E in the Naqu region located in the center

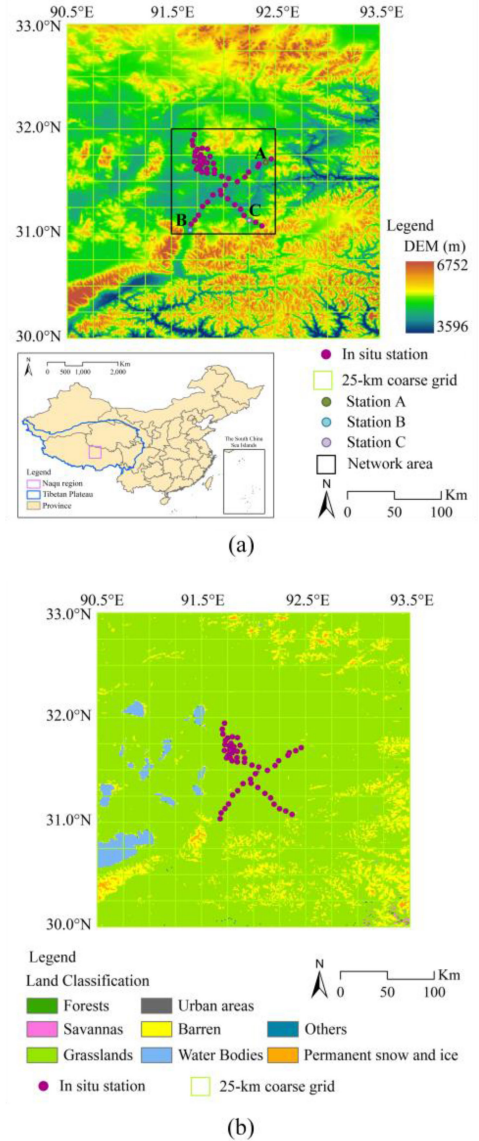


Fig. 2. (a) Location, elevation and (b) land classification of the Naqu region. The spatial distribution of the Naqu network and the coarse grid pixels are also shown. (Network area, Station A, B, and C were used for dynamic analysis.)

of the Tibetan Plateau (TP), China. Due to the influence of the South Asian summer monsoon, the annual precipitation is approximately 500 mm in most of the central TP, with 75% of precipitation events occurring between June and August [50]. Soil thawing and freezing take place around each May and November, respectively.

As seen in Fig. 2, most of the study area has a main vegetation type of high elevation alpine grasslands. The period of interest is during the growing season (May 1–October 31) during 2010–2015. In the following dynamic analysis, five ground stations were employed, and three of them were identified as Station A, Station B, and Station C. The network area covers all of the ground sites.

B. In Situ Measurements

The Naqu network was established in July of 2010 for monitoring SSM and soil temperature, and comprises of 57 ground

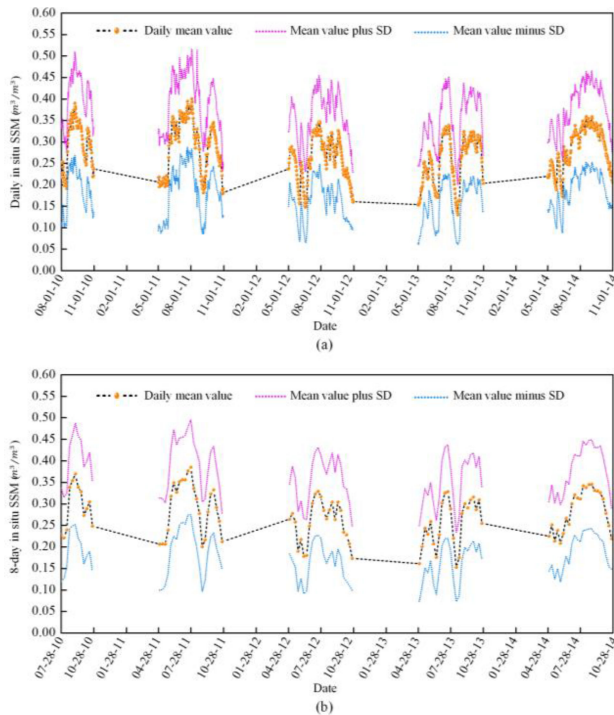


Fig. 3. *In situ* SSM values during the available study period. (a) Daily case. (b) Eight-day case.

stations. The ground stations provide SSM and soil temperature at four different depths of 0–5, 10, 20, and 40 cm, with 30 min and daily sampling intervals. The data are published by the National Tibetan Plateau Data Center¹ from August 1, 2010 to October 31, 2014. The available daily SSM data at depths of 0–5 cm were collected during the period of interest to evaluate the downscaling performances. Not all *in situ* measurements were available during the study period at the 57 ground stations because some stations have been out of operation. The mean and standard deviation (SD) values of *in situ* SSM are shown in Fig. 3 during the available study period for both daily and eight-day cases. For eight-day case, the *in situ* SSM values of every station could be obtained by averaging the corresponding daily SSM measurements within eight days.

C. Coarse-Resolution SSM Product

In 2012, the ESA CCI project for SSM was established to fulfill global long-term SSM monitoring by merging multiple available active and passive microwave-based SSM products [51]. That same year, the first SSM product from the ESA CCI (v0.1) was publicly released. By involving new sensors and improving the merging scheme, the subsequent SSM dataset has been updated over an extended spatiotemporal coverage. The daily SSM product provides a consistent SSM record from 1978 to the present. The latest version (v04.4) of the ESA CCI SSM product at depths of 0.5–5 cm was used in this study, with a spatial resolution of 0.25°.² The SSM data were interpolated and resampled to 25 × 25 km regular grids (see Fig. 2).

¹[Online]. Available: <http://data.tpdc.ac.cn/en/data>

²[Online]. Available: <https://www.esa-soilmoisture-cci.org>

D. Moderate Resolution Imaging Spectroradiometer (MODIS) Products

The MODIS is a key instrument onboard the Terra and Aqua satellites. Fine-resolution ancillary variables LST, NDVI, BSA, and LC information were collected from the Version 6 products of Aqua MODIS.³ The daily LST and 16-day NDVI were provided by MYD11A1 and MYD13A2 at 1-km resolution, while the 16-day albedo and annual LC were provided by MCD43A3 and MCD12Q1 at 500-m resolution. The BSA data were calculated from shortwave radiation of MCD43A3, which uses a linear combination of the black-sky and white-sky albedo data, with weights of 0.34 for the former and of 0.66 for the latter. The SEE data were calculated from the soil and vegetation temperatures which are derived from MODIS LST and NDVI [13], [48]. All MODIS products were reprojected consistently with the ESA CCI product. Missing values were filled using the aforementioned spatiotemporal prediction method to ensure complete coverage. The LST and NDVI data were resampled and aggregated into 1 × 1 km and 25 × 25 km regular grids. The average aggregations of BSA, modal aggregations of LC and SEE were achieved at both fine and coarse grids.

E. DEM Products

The DEM at 90-m resolution provided by the NASA Shuttle Radar Topographic Mission (SRTM) within the study area was employed. The void-filled DEM product was downloaded.⁴ The DEM data were resampled into 1 × 1 km and 25 × 25 km regular grids by using average aggregations. The basic terrain factors at 1 km and 25 km, including aspect and slope, were calculated from the DEM information.

IV. RESULTS AND DISCUSSION

A. Downscaled 1-km SSM

Fig. 4 displays the 25-km SSM images in comparison with the 1-km downscaled SSM predictions by six different models (i.e., SVATARK, SVRK, SVRB, KNN, GWRK, and DISPATCH) for May 1 of 2011, July 20 of 2013, and September 22 of 2015. It can be inferred that the 1-km downscaled results provide more detailed information and variations of the SSM spatial distribution within each 25 × 25 km grid. The SSM data at fine spatial resolution can improve the characterization of the spatial variability of the SSM, which are useful for filling the gap between low-spatial-resolution SSM satellite observations and the needs of catchment-based or regional hydroecological studies. In the downscaled SSM images, the maximum and minimum values of SSM are shown in blue and red, respectively. The blue areas are near the water bodies and in areas with low elevation. Besides surface water, the negative correlation with elevation is another primary factor affecting the spatial distribution of SSM. The results from the proposed SVATARK method showed spatial patterns that were similar to those of the

³[Online]. Available: <https://lpdaac.usgs.gov/>

⁴[Online]. Available: <https://www.usgs.gov/centers/eros>

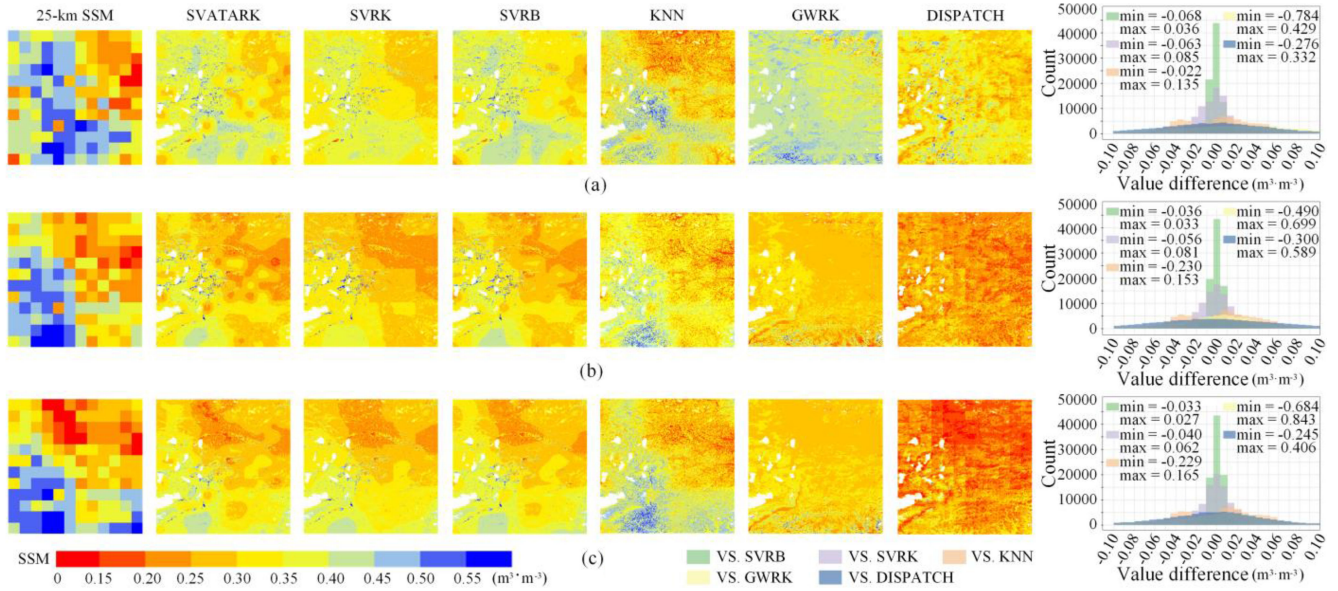


Fig. 4. Comparison of 25-km SSM and downscaled SSM by six models for (a) May 1, 2011, (b) July 20, 2013, and (c) September 22, 2015. The green, purple, pink, yellow, and steelblue histograms show the differences of downscaled results between SVATARK versus SVRB, SVRK, KNN, GWRK, and DISPATCH.

25-km SSM, while SVRK and SVRB produced smoother down-scaled results, KNN displayed more areas with low and high values, GWRK presented a more uniform distribution of SSM with few extreme values, and DISPATCH generated blocking artifacts in the boundaries of 25-km pixels and displayed more areas with high values. The coherence of the ATAK predictions ensures that the average of the disaggregated predictions is equal to the original areal data, and confers the downscaled SSM of SVATARK a continuous pattern. Visual comparison of the downscaled SSM products confirmed that the failure of SVRK to predict extreme SSM and the failure of SVRB to properly capture high SSM were influenced by the kriging and bilinear interpolations. The poor performances of KNN and GWRK might be explained by the use of the nearby important point values when predicting and the limitation of linear regression, respectively. The linear model linked to SSM and SEE might have influenced the performance of DISPATCH, where its block effect was caused by the same coarse SSM value used in the inner high-resolution pixels. SVATARK exhibits better results in modeling extreme SSM (both maximum and minimum) than the other two downscaling methods.

The difference histograms of the downscaled results between SVATARK versus other five methods are also shown in Fig. 4. The histograms of SVR-based methods do not indicate large differences, while it shows obvious distinctions in KNN, GWRK, and DISPATCH cases, perhaps because the same trend model was employed in the first three downscaling methods. For the SVRB case, the minimum differences range from -0.068 to $-0.033 \text{ m}^3 \cdot \text{m}^{-3}$ and the maximums range from 0.027 to $0.036 \text{ m}^3 \cdot \text{m}^{-3}$, while the differences for SVRK, KNN, GWRK, and DISPATCH span from -0.063 to $0.085 \text{ m}^3 \cdot \text{m}^{-3}$, from -0.229 to $0.165 \text{ m}^3 \cdot \text{m}^{-3}$, from -0.784 to $0.843 \text{ m}^3 \cdot \text{m}^{-3}$, and from -0.300 to $0.589 \text{ m}^3 \cdot \text{m}^{-3}$, respectively. Although only three days of downscaled SSM predictions are presented in Fig. 4, it is evident that the three SVR-based downscaling approaches

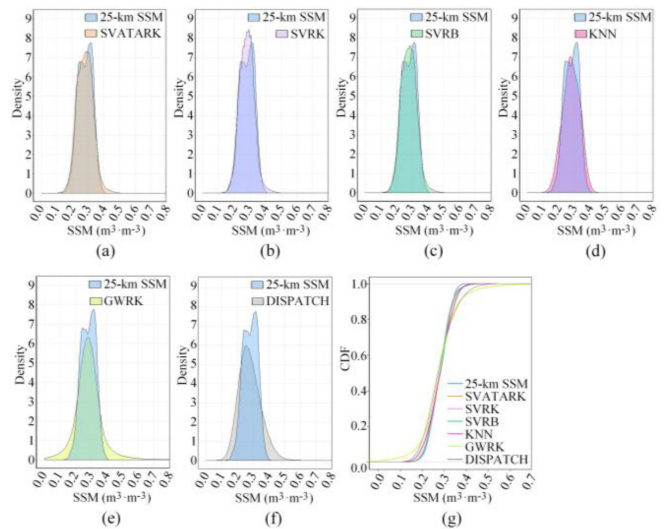


Fig. 5. Density plots of (a) SVATARK, (b) SVRK, (c) SVRB, (d) KNN, (e) GWRK, and (f) DISPATCH, and (g) CDF of six downscaled predictions versus 25-km SSM during the study period.

generate fine-resolution predictions with similar numerical distribution, while the KNN, GWRK, and DISPATCH cases present discrepancies, which are reflected in the similar or dissimilar performances of the cumulative distribution functions (CDFs) derived from 36-month data (see Fig. 5). In addition, as found in the visual comparison, the CDFs and density plots of the six downscaled predictions match well with those of the 25-km SSM product (see Fig. 5). When comparing the differences of the density curves between 25-km SSM and the six downscaled SSM, SVATARK appears to have the closest match. Over the study area, the downscaled SSM by SVATARK is quite similar to the coarse SSM, both in spatial distribution and values. The following validations further demonstrate its improved performance.

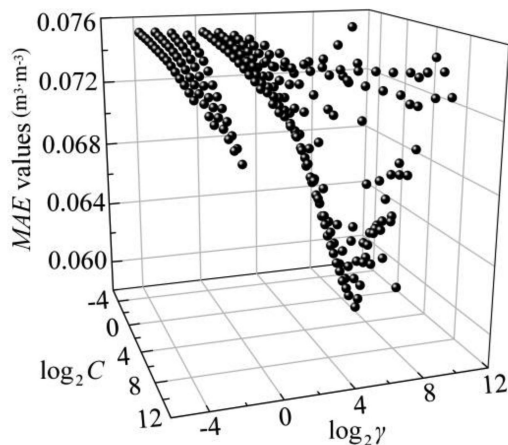


Fig. 6. MAE values of downscaled results on grasslands generated by SVATARK with varying C and γ .

Moreover, the mentioned penalty coefficient C and gamma γ of SVR model have great impacts on its performance, and would also affect the performance of SVATARK. Considering that most of the ground stations are located in grasslands, the SVATARK method was implemented repeatedly for grassland type with different C and γ to assess the influence of these two parameters on the downscaled results. Both $\log_2 C$ and $\log_2 \gamma$ were defined in a range from -4 to 10 with an interval of 1 . Fig. 6 shows the MAE values of downscaled results on grasslands generated by SVATARK with varying C and γ . It is observed that the MAE value becomes smaller as the two parameters increase in the beginning, then tends to be stable or to decrease. The model complexity and the learning ability would improve with increased C and γ , which might lead to an initial higher accuracy but subsequent lower accuracy with overfitting.

B. Validation With the In Situ SSM Measurements

The downscaled 1-km SSM of each algorithm were validated using the *in situ* observations from 57 ground stations over the Naqu region within the available days from 2010 to 2014. Fig. 7 shows the comparisons between ground observations and 1-km predictions of SVATARK, SVRK, SVRB, KNN, GWRK, and DISPATCH. It also displays the linear regression models between ground observations and downscaled predictions, which owns a different intercept for each downscaling case. The SVRK model produces more accurate predictions than those of the SVRB-based, GWRK-based, KNN-based, and DISPATCH-based SSM data with an RMSE value of $0.10 \text{ m}^3 \cdot \text{m}^{-3}$, a MAE value of $0.07 \text{ m}^3 \cdot \text{m}^{-3}$, and a SLOP value of 0.70 , but with slightly smaller R value of 0.64 . The KNN and GWRK models generate a better R value of 0.69 than the SVRK and SVRB methods. The DISPATCH model presents a better R value of 0.65 than the SVRK method. A smaller MAE seems to improve the RMSE, but not always corresponds to higher SLOP and R values. The comparison illustrates that the proposed SVATARK approach significantly outperforms the other four downscaling approaches with the smallest RMSE and MAE values of 0.08 and $0.06 \text{ m}^3 \cdot \text{m}^{-3}$, and the largest R and SLOP values of 0.72 and 0.71 . The scatterplot from the SVATARK approach visually

gathers along the 1:1 line and has the lowest dispersion. Because of the model prediction error and errors in input variables, there are some discrepancies between the 1-km downscaled results and *in situ* measurements. In addition, these discrepancies were also due to the representativeness errors among different resolutions, i.e., point values versus either $1 \times 1 \text{ km}$ regular grid values or $25 \times 25 \text{ km}$ regular grid values, the different time intervals in the temporal aggregation among various products, i.e., daily remote-sensed SSM and 30-min *in situ* SSM, and the impact of variables at different resolutions has been reduced by the aggregation of variables into the same resolution. Although the spatiotemporal prediction approach can help fill the missing values of remote-sensed data, the errors from this process can be propagated into the final results. In future research, more error analyses, especially before downscaling, should be performed to improve the downscaling accuracy. The improvements made by SVATARK are illustrated by increases in R (0.06 or 8.2% on average) and SLOP (0.08 or 12.8% on average), and decreases in RMSE ($0.03 \text{ m}^3 \cdot \text{m}^{-3}$ or 24.4% on average) and MAE ($0.02 \text{ m}^3 \cdot \text{m}^{-3}$ or 22.9% on average). A general improvement can be seen in Figs. 8–10.

Fig. 8 presents the comparisons between ground observations and six downscaled results for each year. The CDFs of SSM measured by *in situ* and downscaled results derived from six different algorithms are displayed in Fig. 9. The performance shown by the four statistical metrics appears inconsistent from year to year when comparing the results of SVRK, SVRB, KNN, GWRK, and DISPATCH methods. The other four results are frequently worse than the SVRK's during the five-year period. In comparison with the other five downscaling models, the SVATARK model produces the two highest statistical metrics (i.e., R and SLOP) and the two lowest statistical metrics (i.e., RMSE and MAE). The CDF comparison indicates that the downscaled results of SVATARK models show minimum deviations from the CDF calculated from the ground observations.

To explore the spatial distribution of the estimation errors, the MAE values from 2010 to 2014 of the 57 ground stations were calculated. For each ground station, Fig. 10 visualizes the MAE values of the downscaled results using SVATARK, SVRK, SVRB, KNN, GWRK, and DISPATCH with color bars, as well as the mean value of these MAEs with color circles. It displays the percentage of cloud-free days during available period in Fig. 10(b). As seen in Fig. 10, the MAE values tend to be higher in the upper left and middle part, likely due to higher topographic relief and the lack of the corresponding original remotely sensed observations, which could introduce errors from filling gaps. The SVATARK model has the smallest MAE values at each station, suggesting a better performance than the other five methods. Further SSM analyses at each station are shown in the next section. Although the above validations were all taken at stations of grasslands, which is the main vegetation type in the Naqu region, the SVATARK method could theoretically result in accurate downscaling predictions from other areas, given its ability to learn for small samples and the strong generalization of SVR, as well as the coherence of ATAK. The proposed method should be validated and applied to other land cover types in future work.

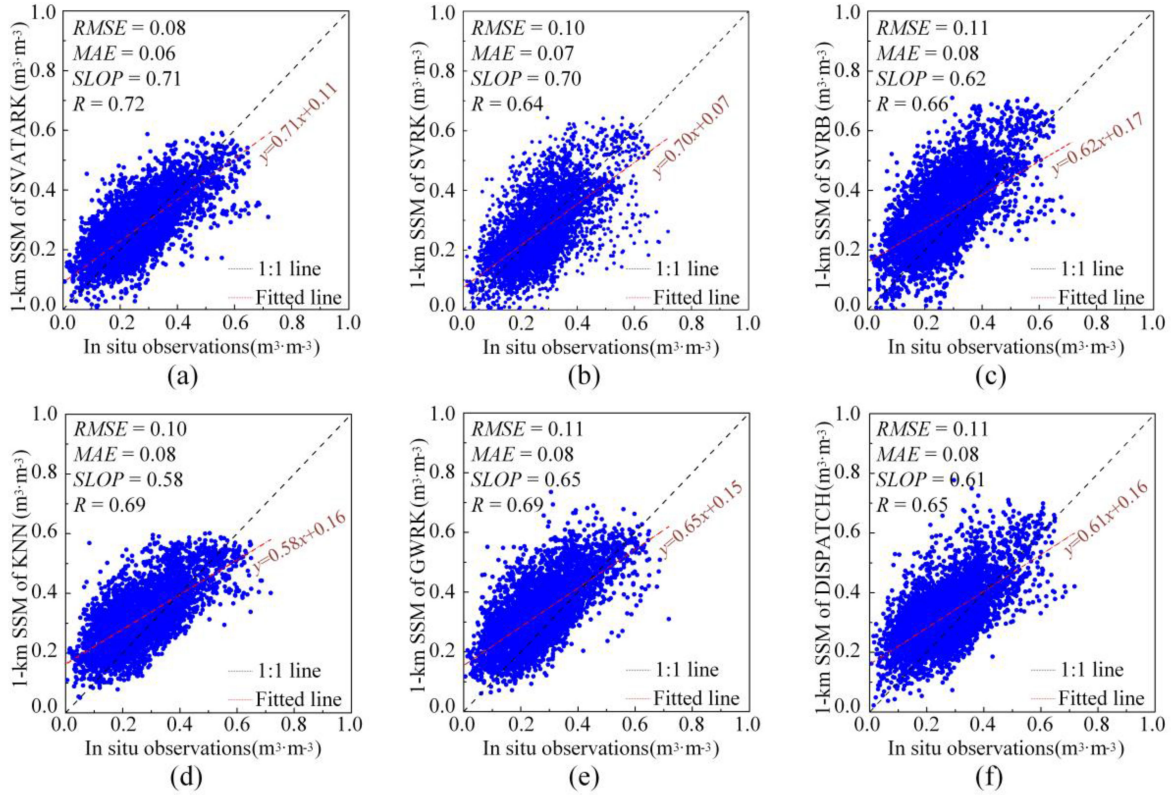


Fig. 7. Scatter plots between ground observations and downscaled 1-km SSM using (a) SVATARK, (b) SVRK, (c) SVRB, (d) KNN, (e) GWRK, and (f) DISPATCH during available period also displaying the linear regression models between ground observations and downscaled predictions.

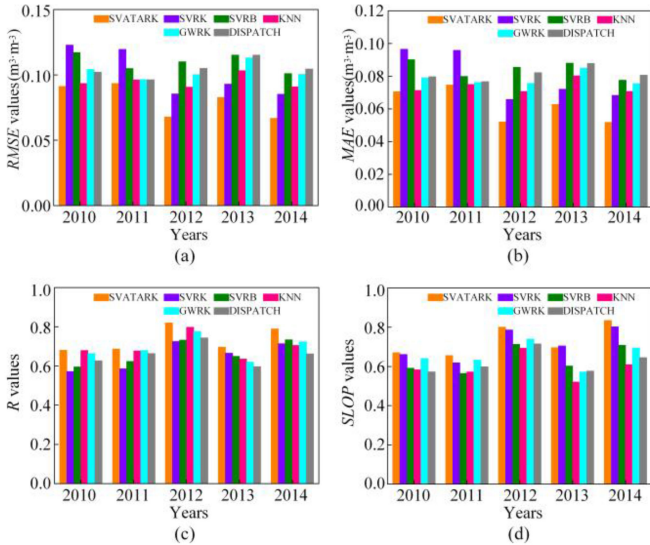


Fig. 8. Comparison metrics between the ground observations and six downscaled results from 2010 to 2014: (a) RMSE; (b) MAE; (c) R ; and (d) SLOP.

In addition to the spatial representativeness errors, there is a bias between the coarse SSM product and the *in situ* observations. The ground-based measurements and original coarse SSM product have different observation intervals, with 30 min and daily, respectively. The eight-day *in situ* and coarse SSM could be derived by averaging all available corresponding observations within eight days. The different time intervals in

the averaging process would also cause the bias between the eight-day averaged remote-sensed and *in situ* SSM, which would also be propagated into the 1-km downscaled results. Compared with the eight-day averages of ground observations, the eight-day coarse SSM product owns an RMSE value of $0.13 \text{ m}^3 \cdot \text{m}^{-3}$, a MAE value of $0.10 \text{ m}^3 \cdot \text{m}^{-3}$, and with very small SLOP and R values of -0.014 and -0.03 . To reduce the negative effect of the bias in the downscaled prediction accuracy, in this experiment, the mean difference between the eight-day averaged 25-km SSM values and the point supports was removed before downscaling. It narrowed the discrepancy between the input coarse SSM and ground observations, with smaller RMSE and MAE values of 0.12 and $0.09 \text{ m}^3 \cdot \text{m}^{-3}$. Because of the same input variables in each downscaling method, it would not generate significant impact on the relative performances among various downscaled results. From a validation perspective, the 1-km downscaled SSM results are at an obvious advantage compared to the 25-km SSM with the improvements of all statistical metrics, especially for the 1-km SSM derived by SVATARK, which has a decrease of 33.3% in both RMSE and MAE. Considering that all the ground stations were installed on grasslands, this bias step was used without consideration of the differences resulting from topography and LC types. However, the remotely sensed product might perform differently over various surfaces and therefore, incorporation of the impact of LC types may be beneficial for improving the bias correction accuracy. In addition, how to determine whether bias needs to be applied to all coarse grids is still a problem, if the SSM can be effectively observed at some

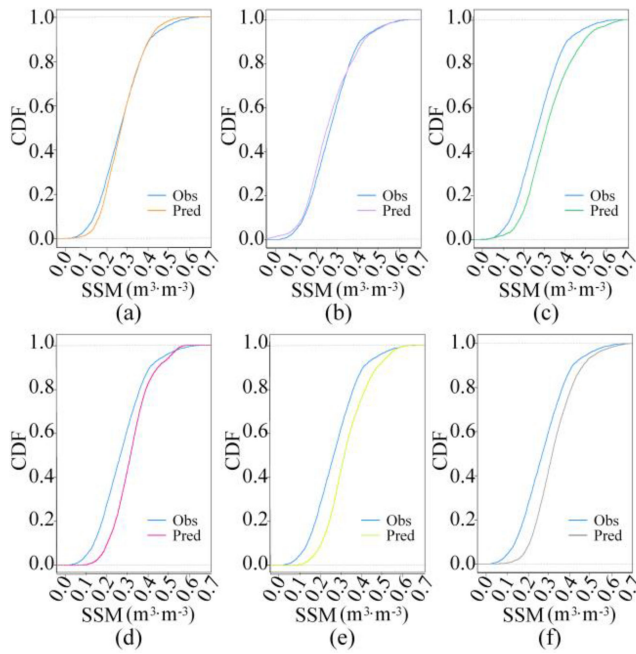


Fig. 9. CDFs of the ground observations and downscaled 1-km SSM using (a) SVATARK, (b) SVRK, (c) SVRB, (d) KNN, (e) GWRK, and (f) DISPATCH during available period. (The ground observations and downscaled results are marked as Obs and Pred, respectively.)

grids. Moreover, the employed average values of SSM within eight days in this article could reflect the average state of SSM; however, it would ignore the rapid changes of instantaneous SSM.

C. Dynamic Analysis of Downscaled SSM

The downscaled maps and validation analyses described in this study illustrate that the downscaled SSM results generally show a good performance compared with ground-based measurements, and their spatial pattern follows those of the coarse SSM. In this section, we investigate whether the fine-resolution SSM predictions from the six downscaled methods also capture the temporal dynamics of ground-based SSM observations during the study period. Fig. 11 shows the temporal variations of 1-km downscaled SSM derived from all six downscaling methods and *in situ* observations at five ground stations (in Figs. 2 and 10) and at network scale (i.e., network area in Fig. 2). There is a significant seasonal variation in the time series, generally reaching its highest value in August. By using average aggregation within the network domain, the aggregated values of ground measurements and 1-km SSM were obtained. Two statistical metrics, MAE and RMSE, were used to evaluate their performance. The MAE and RMSE values of these time-series SSM seem better than the comparisons results of the scatter plot in Fig. 7, which might be caused by the varying performance of downscaled results at different stations.

In the proposed SVATARK downscaling method, the values of MAE and RMSE at all five stations range from 0.033 to 0.065 $\text{m}^3 \cdot \text{m}^{-3}$ and from 0.041 to 0.076 $\text{m}^3 \cdot \text{m}^{-3}$, respectively, where SVATARK is found to be more accurate than the other five approaches. From the time series comparisons at different

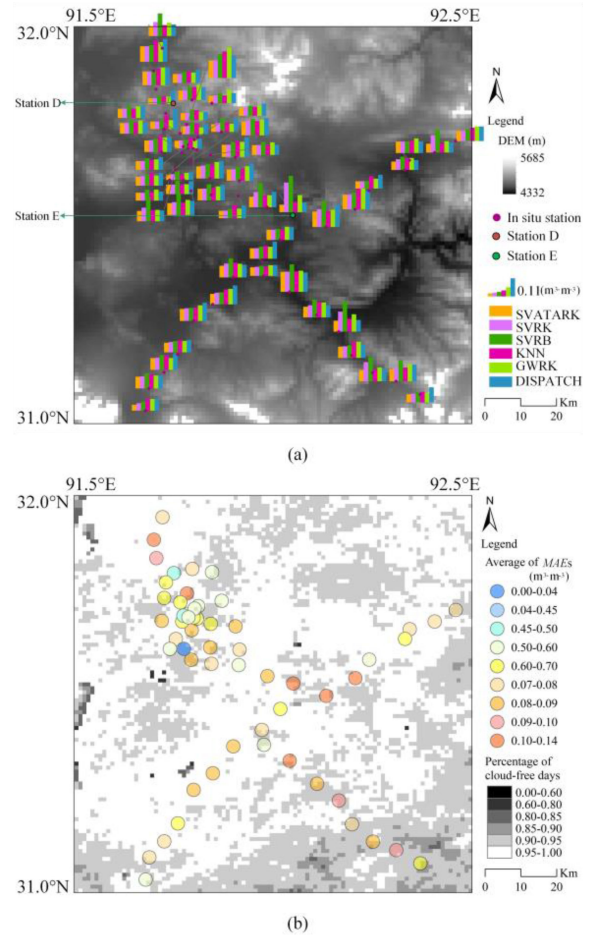


Fig. 10. Spatial distribution of (a) MAE between ground observations and six downscaled results and (b) mean value of different MAE derived by six methods for each station and the percentage of cloud-free days during the available period. (The lowest and highest average MAE for Station D and Station E, respectively, were employed to analyze temporal characteristics.)

ground stations [see Fig. 11(a)–(e)], the downscaled SSM predictions, especially of the SVATARK method, show temporal consistency with the *in situ* observations, although there is a significant bias between them. This indicates that the downscaled SSM of SVATARK can describe the temporal changes of the *in situ* SSM. The discrepancies are mainly because of the large-scale differences between the 1-km predictions and point observations. For five statistical downscaling methods, the best performance was obtained at Station D with the lowest MAE of all stations, which did not happen in DISPATCH. It might be due to the different ancillary variables employed in statistical and physical-based downscaling methods, respectively. The variation in performance of various stations might be the result of the station's location, which would affect the soil type and have different accuracy of the input variables. The ranges of values for the downscaled SSM for almost all are less than the ground measurements' range. This matches well with the fact that the range of SSM decreases dynamically from fine to coarse scales [34]. Areal-averaged downscaled SSM agrees well with the ground-based SSM in Fig. 11(f). However, the discrepancies between the six downscaled SSM and ground observations in the

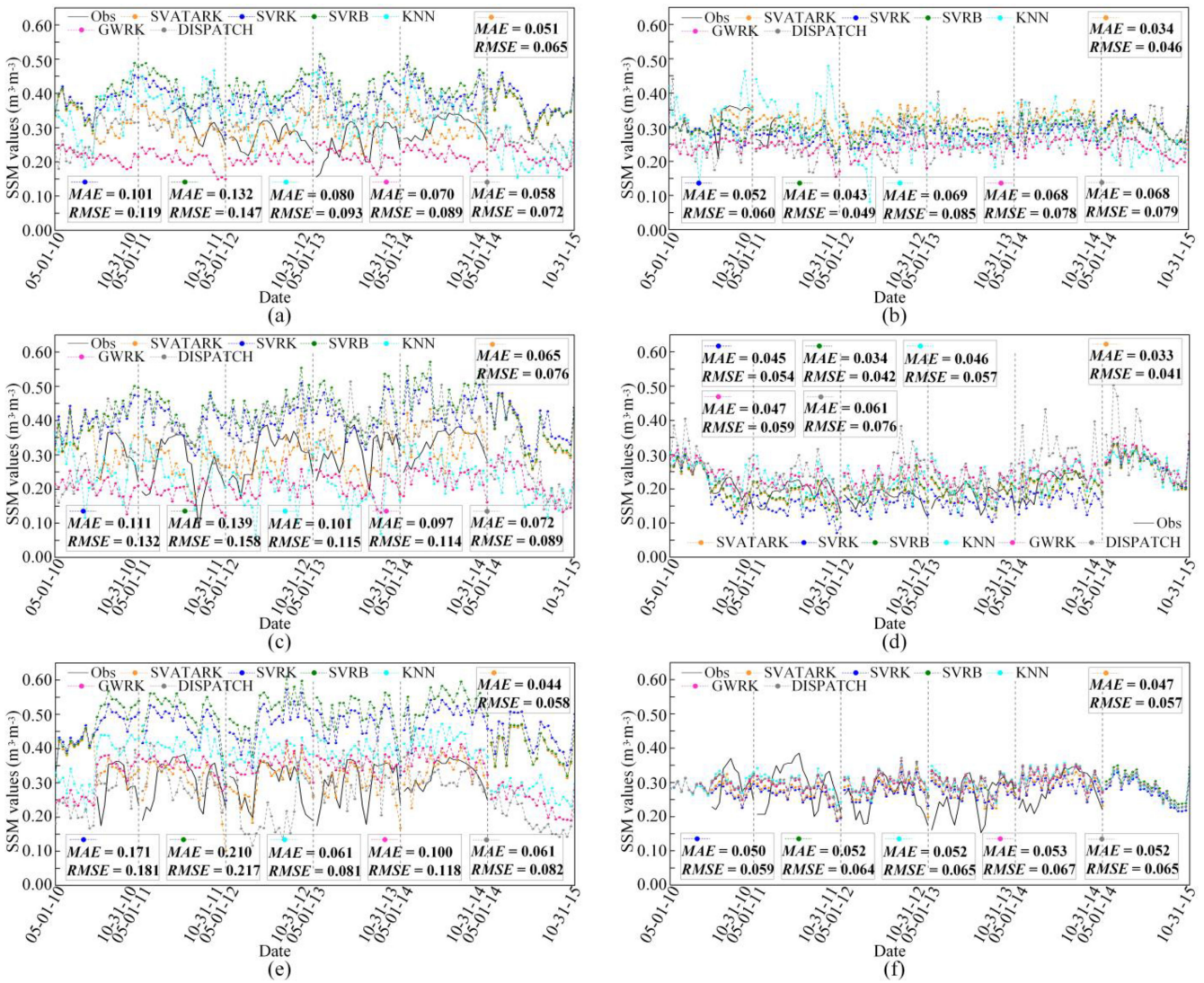


Fig. 11. Time series of the 1-km SSM and *in situ* observations at the five selected ground stations and network areas: (a) Station A; (b) Station B; (c) Station C; (d) Station D; (e) Station E; and (f) Network area.

network area seem smaller than those of the stations, particularly for the SVRK, SVRB, KNN, GWRK, and DISPATCH methods, perhaps due to the comparisons at the same scale for avoiding scaling differences in the validation. The errors associated with upscaling the SSM from the 1-km and point scale to the network scale require further research and exploration. Note that a better performance of the downscaled SSM is also obtained by using SVATARK.

Soil moisture is a direct indicator of agricultural drought. *In situ* observations of SSM may not be able to assess drought conditions in a region, whereas the 1-km downscaled predictions could provide powerful data support. A simple relative drought analysis was attempted using the downscaled SSM of SVATARK. The pixels with anomalously low values in the downscaled images were counted by comparing the pixel values on the same date every year. The mean and standard deviation were calculated for each pixel. Pixels with a larger absolute value than the standard deviation were considered to be in a drought condition. The main idea behind this assumption is to find the

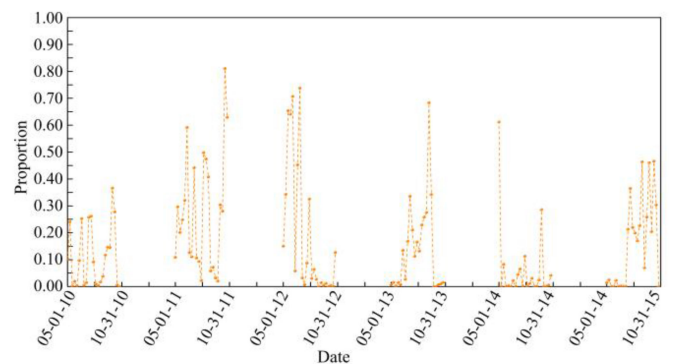


Fig. 12. Proportion of pixels with relatively smaller SSM in downscaled images of SVATARK during study period.

pixel which has a low value and relatively large variation in SSM over the same period. Fig. 12 shows the proportion of pixels with relatively smaller SSM in the downscaled images using SVATARK during the study period. Several proportions are

larger than 0.30, meaning that 30% of the 1×1 km pixels in the corresponding date have abnormally low values. The proportion values suddenly increase in mid-July 2015, indicating relative drought conditions. These results are consistent with Zhu *et al.* [52]. Although the SSM at 0–5 cm depth might have a limited ability to reflect soil drought without deep soil moisture, this preliminary attempt demonstrates that the proposed downscaling method could be used in drought remote sensing monitoring applications for a large area. The downscaled SSM could also help understand how often and where these droughts occur. In this study, only four types of ancillary variables were employed, but rainfall (including its infiltration and runoff) also affects SSM variations, and should be explored as an ancillary variable in the downscaling process in future work.

V. CONCLUSION

In order to transform coarse-resolution remote sensed data to a fine resolution, we proposed a machine learning-based geostatistical downscaling method in this study. The proposed SVATARK relies on SVR that expresses the nonlinear relationship between target (i.e., SSM) and ancillary variables (i.e., LC, LST, NDVI, BSA, and terrain factors), and utilizes ATAK to achieve the predictions on changed supports. SVATARK was compared to the benchmark methods SVRK, SVRB, KNN, GWRK, and DISPATCH to obtain 1-km predictions from a 25-km SSM product over the Naqu region during a 36-month period from 2010–2015. The downscaled predictions were validated using ground stations. In general, the adopted statistical metrics indicate that the SVATARK downscaling approach obtained the greatest accuracy, and the dynamic analysis of 1-km SSM at five different stations reached the same conclusion. The comparison among the aggregated SSM values for network area also shows the high potential of our proposed method for producing fine-resolution SSM images. Moreover, the SVATARK-based downscaled predictions were extended to monitor the relative drought by capturing the abnormally low SSM, which might be further generalized with systematic analyses. The SVATARK method is entirely general, and it can be employed to downscale or even upscale other continuous variables owing to the changes in the supports in ATAK. In the future work, other machine learning or deep learning methods, such as random forest or neural network algorithms, could be applied in trend predictions and could be integrated with ATAK for spatial scaling.

ACKNOWLEDGMENT

The authors would like to thank all the data producers.

REFERENCES

- [1] S. I. Seneviratne *et al.*, “Investigating soil moisture–climate interactions in a changing climate: A review,” *Earth Sci. Rev.*, vol. 99, no. 3/4, pp. 125–161, 2010.
- [2] S. M. Bateni and D. Entekhabi, “Relative efficiency of land surface energy balance components,” *Water Resour. Res.*, vol. 48, no. 4, 2012, Art. no. W04510.
- [3] P. Krishnan *et al.*, “Impact of changing soil moisture distribution on net ecosystem productivity of a boreal aspen forest during and following drought,” *Agricultural Forest Meteorol.*, vol. 139, no. 3/4, pp. 208–223, 2006.
- [4] H. Wang *et al.*, “Monitoring winter wheat drought threat in northern China using multiple climate-based drought indices and soil moisture during 2000–2013,” *Agricultural Forest Meteorol.*, vol. 228/229, pp. 1–12, 2016.
- [5] W. Dorigo *et al.*, “ESA CCI soil moisture for improved earth system understanding: State-of-the art and future directions,” *Remote Sens. Environ.*, vol. 203, pp. 185–213, 2017.
- [6] P. Dobriyal, A. Qureshi, R. Badola, and S. A. Hussain, “A review of the methods available for estimating soil moisture and its implications for water resource management,” *J. Hydrol.*, vol. 458/459, pp. 110–117, 2012.
- [7] B. Kerkez, S. D. Glaser, R. C. Bales, and M. W. Meadows, “Design and performance of a wireless sensor network for catchment-scale snow and soil moisture measurements,” *Water Resour. Res.*, vol. 48, 2012, Art. no. W09515.
- [8] M. Zreda *et al.*, “COSMOS: The cosmic-ray soil moisture observing system,” *Hydrol. Earth Syst. Sci.*, vol. 16, pp. 1–21, 2012.
- [9] G. P. Petropoulos, G. Ireland, and B. Barrett, “Surface soil moisture retrievals from remote sensing: Current status, products & future trends,” *Phys. Chem. Earth A/B/C*, vol. 83/84, pp. 36–56, 2015.
- [10] E. G. Njoku, T. J. Jackson, V. Lakshmi, T. K. Chan, and S. V. Nghiem, “Soil moisture retrieval from AMSR-E,” *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 2, pp. 215–229, Feb. 2003.
- [11] R. M. Parinussa *et al.*, “Global surface soil moisture from the microwave radiation imager onboard the Fengyun-3B satellite,” *Int. J. Remote Sens.*, vol. 35, no. 19, pp. 7007–7029, 2014.
- [12] T. Meissner, F. J. Wentz, and D. M. L. Vine, “The salinity retrieval algorithms for the NASA Aquarius version 5 and SMAP version 3 releases,” *Remote Sens.*, vol. 10, 2018, Art. no. 1121.
- [13] B. Molero *et al.*, “SMOS disaggregated soil moisture product at 1 km resolution: Processor overview and first validation results,” *Remote Sens. Environ.*, vol. 180, pp. 361–376, 2016.
- [14] J. Peng *et al.*, “A review of spatial downscaling of satellite remotely sensed soil moisture,” *Rev. Geophys.*, vol. 55, no. 2, pp. 341–366, 2017.
- [15] Y. Ge *et al.*, “Principles and methods of scaling geospatial earth science data,” *Earth Sci. Rev.*, vol. 197, 2019, Art. no. 102897.
- [16] M. Piles, D. Entekhabi, and A. Camps, “A change detection algorithm for retrieving high-resolution soil moisture from SMAP radar and radiometer observations,” *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 12, pp. 4125–4131, Dec. 2009.
- [17] R. van der Velde, M. S. Salama, O. A. Eweys, J. Wen, and Q. Wang, “Soil moisture mapping using combined active/passive microwave observations over the east of The Netherlands,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 9, pp. 4355–4372, Sep. 2015.
- [18] X. Zhan, P. R. Houser, J. P. Walker, and W. T. Crow, “A method for retrieving high-resolution surface soil moisture from Hydros L-band radiometer and radar observations,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1534–1544, Jun. 2006.
- [19] X. Wu, J. P. Walker, C. Rüdiger, R. Panciera, and Y. Gao, “Medium-resolution soil moisture retrieval using the Bayesian merging method,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6482–6493, Nov. 2017.
- [20] O. Merlin *et al.*, “Performance metrics for soil moisture downscaling methods: Application to DISPATCH data in central Morocco,” *Remote Sens.*, vol. 7, no. 4, pp. 3783–3807, 2015.
- [21] Y. Malbêteau *et al.*, “DisPATCH as a tool to evaluate coarse-scale remotely sensed soil moisture using localized in situ measurements: Application to SMOS and AMSR-E data in southeastern Australia,” *Int. J. Appl. Earth Observ. Geoinf.*, vol. 45, pp. 221–234, 2016.
- [22] Y. Yang *et al.*, “Estimation of surface soil moisture from thermal infrared remote sensing using an improved trapezoid method,” *Remote Sens.*, vol. 7, no. 7, pp. 8250–8270, 2015.
- [23] E. Babelaian *et al.*, “Mapping soil moisture with the OPTical TRapezoid model (OPTRAM) based on long-term MODIS observations,” *Remote Sens. Environ.*, vol. 211, pp. 425–440, 2018.
- [24] S. B. Duan and Z. L. Li, “Spatial downscaling of MODIS land surface temperatures using geographically weighted regression: Case study in northern China,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6458–6469, Nov. 2016.
- [25] Y. Liu, Y. Yang, W. Jing, and X. Yue, “Comparison of different machine learning approaches for monthly satellite-based soil moisture downscaling over northeast China,” *Remote Sens.*, vol. 10, no. 1, 2018, Art. no. 31.
- [26] S. Mukherjee, P. K. Joshi, and R. D. Garg, “Regression-Kriging technique to downscale satellite-derived land surface temperature in heterogeneous agricultural landscape,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 3, pp. 1245–1250, Mar. 2015.

- [27] Y. Jin, Y. Ge, J. H. Wang, Y. Chen, and G. B. M. Heuvelink, "Downscaling AMSR-2 soil moisture data with geographically weighted area-to-area regression kriging," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2362–2376, Apr. 2018.
- [28] N. S. Chauhan, S. Miller, and P. Ardanuy, "Spaceborne soil moisture estimation at high resolution: A microwave-optical/IR synergistic approach," *Int. J. Remote Sens.*, vol. 24, no. 22, pp. 4599–4622, 2003.
- [29] M. Piles *et al.*, "A downscaling approach for SMOS land observations: Evaluation of high-resolution soil moisture maps over the Iberian peninsula," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 9, pp. 3845–3857, Sep. 2014.
- [30] K. R. Knipper, T. S. Hogue, K. J. Franz, and R. L. Scott, "Downscaling SMAP and SMOS soil moisture with moderate-resolution imaging spectroradiometer visible and infrared products over southern Arizona," *Int. J. Remote Sens.*, vol. 11, no. 2, 2017, Art. no. 026021.
- [31] P. Song, J. Huang, and L. R. Mansaray, "An improved surface soil moisture downscaling approach over cloudy areas based on geographically weighted regression," *Agricultural Forest Meteorol.*, vol. 275, pp. 146–158, 2019.
- [32] Y. H. Zhang, Y. Du, F. Ling, S. Fang, and X. Li, "Example-based super-resolution land cover mapping using support vector regression," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1271–1283, Apr. 2014.
- [33] W. Zhao, N. Sánchez, H. Lu, and A. Li, "A spatial downscaling approach for the SMAP passive surface soil moisture product using random forest regression," *J. Hydrol.*, vol. 563, pp. 1009–1024, 2018.
- [34] P. Abbaszadeh, H. Moradkhani, and X. Zhan, "Downscaling SMAP radiometer soil moisture over the CONUS using an ensemble learning method," *Water Resour. Res.*, vol. 55, no. 1, pp. 324–344, 2019.
- [35] J. Im *et al.*, "Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches," *Environ. Earth Sci.*, vol. 75, no. 15, 2016, Art. no. 1120.
- [36] Z. Wei *et al.*, "Downscaling SMAP soil moisture estimation with gradient boosting decision tree regression over the Tibetan plateau," *Remote Sens. Environ.*, vol. 225, pp. 30–44, 2019.
- [37] C. Song and L. Jia, "A method for downscaling FengYun-3B soil moisture based on apparent thermal inertia," *Remote Sens.*, vol. 8, no. 9, 2016, Art. no. 703.
- [38] S. Chen *et al.*, "Spatial downscaling methods of soil moisture based on multisource remote sensing data and its application," *Water*, vol. 11, no. 7, 2019, Art. no. 1401.
- [39] Y. H. Kaheil, M. K. Gill, M. McKee, L. A. Bastidas, and E. Rosero, "Downscaling and assimilation of surface soil moisture using ground truth measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1375–1384, May 2008.
- [40] N. Djamaï, "A combination of DISPATCH downscaling algorithm with CLASS land surface scheme for soil moisture estimation at fine scale during cloudy days," *Remote Sens. Environ.*, vol. 184, pp. 1–14, 2016.
- [41] P. K. Srivastava, D. Han, M. R. Ramirez, and T. Islam, "Machine learning techniques for downscaling SMOS satellite soil moisture using MODIS land surface temperature for hydrological application," *Water Resour. Manage.*, vol. 27, no. 8, pp. 3127–3144, 2013.
- [42] R. N. Sujay and P. C. Deka, "Support vector machine applications in the field of hydrology: A review," *Appl. Soft Comput.*, vol. 19, pp. 372–386, 2014.
- [43] P. C. Kyriakidis, "A geostatistical framework for area-to-point spatial interpolation," *Geographical Anal.*, vol. 36, no. 3, pp. 259–289, 2004.
- [44] V. Vapnik, S. E. Golowich, and A. J. Smola, "Support vector method for function approximation, regression estimation and signal processing," in *Proc. Neural Inf. Process. Syst. Conf.*, 1997, pp. 281–287.
- [45] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statist. Comput.*, vol. 14, no. 3, pp. 199–222, 2004.
- [46] D. Meyer *et al.*, "E1071: Misc functions of the department of statistics, probability theory group (Formerly: E1071), TU Wien," *R package Version 1.7-3*, vol. 1, pp. 7–3, 2015.
- [47] P. Goovaerts, "Kriging and semivariogram deconvolution in the presence of irregular geographical units," *Math. Geosci.*, vol. 40, no. 1, pp. 101–128, 2008.
- [48] T. Jing, X. Deng, and H. Su, "Intercomparison of two trapezoid-based soil moisture downscaling methods using three scaling factors," *Int. J. Digit. Earth*, vol. 12, no. 4, pp. 485–499, 2019.
- [49] F. Gerber, R. de Jong, M. E. Schaepman, G. Schaepman-Strub, and R. Furrer, "Predicting missing values in spatio-temporal remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2841–2853, May 2018.
- [50] K. Yang *et al.*, "A multiscale soil moisture and freeze–thaw monitoring network on the third pole," *Bull. Amer. Meteorol. Soc.*, vol. 94, no. 12, pp. 1907–1916, 2013.
- [51] W. Wagner *et al.*, "Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 7, pp. 315–321, 2012.
- [52] X. C. Zhu *et al.*, "Application of cosmic-ray neutron sensing to monitor soil water content in an alpine meadow ecosystem on the northern Tibetan plateau," *J. Hydrol.*, vol. 536, pp. 247–254, 2016.



Yan Jin received the B.S. degree in information and computation science and the M.S. degree in applied mathematics from Chang'an University, Xi'an, China, in 2011 and 2014, respectively, and the Ph.D. degree in cartography and geographical information system from the State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, University of Chinese Academy of Sciences, Beijing, China, in 2018.

She is currently a Lecturer with the School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing, China. Her current research interests include scale transformation, data fusion, geostatistics, and remote sensing applications.



Yong Ge (Member, IEEE) received the Ph.D. degree in cartography and geographical information system from the Chinese Academy of Sciences (CAS), Beijing, China, in 2001.

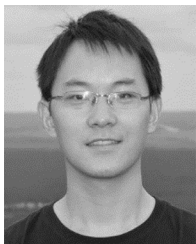
She is a Professor with the State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, CAS. She has directed research in more than ten national projects. She is the author or coauthor of more than 100 scientific papers in refereed journals, a book, and six book chapters, the editor of a book, and holds eight granted patents in improving the accuracy of information extraction from remotely sensed imagery. Her research interests include spatial data analysis and data quality assessment.

Dr. Ge has been involved in the organization of several international conferences and workshops. She is a Steering Committee Member of the International Spatial Accuracy Research Association (ISARA) and the Editorial Board Member of *Spatial Statistics* (Elsevier).



Yaojie Liu received the B.S. degree in geoinformation science and technology from Tongji University, Shanghai, China, in 2011, the M.S. degree in cartography and geographical information system from Capital Normal University, Beijing, China, in 2014, and the Ph.D. degree in ecology from the Chinese Academy of Sciences (CAS), Beijing, China, in 2018.

He is currently a Postdoctoral Researcher with the International Institute for Earth System Science, Nanjing University, Nanjing, China. His current research interests include global change ecology, remote sensing technology in ecological and environmental applications, and carbon cycle.



Yuehong Chen received the B.S. degree from Hohai University, Nanjing, China, in 2010, and the M.Sc. and Ph.D. degrees from the State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, University of Chinese Academy of Sciences, China, in 2013 and 2016, respectively.

He is currently an Associate Professor with the School of Earth Sciences and Engineering, Hohai University. His current research interests include geospatial data analysis, geostatistics, and scale trans-

formation.



Haitao Zhang received the Ph.D. degree in cartography and geographical information system from Nanjing Normal University, Nanjing, China, in 2006.

He is currently a Professor with the School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing, China. His current research interests include data mining, spatiotemporal reasoning, and LBS privacy protection.



Gerard B. M. Heuvelink received the M.Sc. degree in applied mathematics from Twente Technical University, Enschede, The Netherlands, in 1987, and the Ph.D. degree in environmental sciences from Utrecht University, Utrecht, The Netherlands, in 1993.

He was an Assistant Professor with the University of Amsterdam, Amsterdam, The Netherlands, until 2003, when he accepted a dual appointment as Senior Researcher with Alterra and an Associate Professor with Wageningen University, Wageningen, The Netherlands. In 2011, he changed his appointment

with Alterra for as a Senior Researcher with ISRIC - World Soil Information, Wageningen, The Netherlands. Since 2011, he has been a Visiting Professor with the University of the Chinese Academy of Sciences, Beijing, China. In 2017, he was appointed Special Professor in Pedometrics and Digital Soil Mapping with Wageningen University. He is the author or coauthor of more than 250 scientific publications on geostatistics, spatial uncertainty analysis, and pedometrics, about 110 of which appeared in peer-reviewed international journals.

Dr. Heuvelink is an Associate Editor for the *European Journal of Soil Science* and of *Spatial Statistics*. He was the recipient of the Richard Webster Medal from the Pedometrics Commission of the International Union of Soil Science, in 2014.