Enhancing Global Surface Soil Moisture Estimation From ESA CCI and SMAP Product With a Conditional Variational Autoencoder

Changjiang Shi ⓘ, *Graduate Student Member, IEEE*, Zhijie Zhang ⓘ, Shengqing Xiong ⓘ, and Wanchang Zhang ⓘ, *Member, IEEE*

Abstract—High-quality soil moisture (SM) estimation is crucial for various applications, including drought monitoring, environmental assessment, and agricultural management. Advances in remote sensing technology have enabled the retrieval of near real-time Earth surface SM using both active and passive sensors. However, the ESA climate change initiative (CCI) SM product, which combines data from multiple sensors, sacrifices spatial-temporal resolution and coverage due to satellite orbit constraints and retrieval algorithms. To address this issue, an SM reconstruction approach based on a conditional variational autoencoder model was developed, leveraging the high spatial resolution of SMAP L4 data and the accuracy of CCI fused products across different land cover types. This method resulted in the creation of a global three-day SM product at 0.0625° spanning from 2015 to 2021. The reconstructed SM product underwent rigorous validation against global core SM sites and sparse observation networks. The evaluation employed multiple metrics, including the global unbiased root mean square error (ubRMSE) and correlation coefficient (CC). The validation yielded results, with ubRMSE values of approximately 0.029 and 0.071 m³/m³, and CC values of around 0.863 and 0.743 for core SM sites and sparse observation networks. This reconstructed product offers global coverage and enhanced accuracy compared to existing benchmarks.

Index Terms—Conditional variational autoencoder (CVAE), ESA climate change initiative (CCI), product reconstruction, SMAP L4, surface soil moisture (SM).

Changjiang Shi is with the Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China, and with the International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China, and also with the College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: shichangjiang20@mails.ucas.ac.cn).

Zhijie Zhang is with the Department of Environment and Society, Quinney College of Natural Resources, Utah State University, Logan, UT 84322 USA (e-mail: zhijiezhang@arizona.edu).

Shengqing Xiong is with the China Aero Geophysical Survey and Remote Sensing Center for Natural Resources, Beijing 100083, China (e-mail: xsqa-grs@126.com).

Wanchang Zhang is with the International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China, and also with the Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China (e-mail: zhangwc@radi.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2024.3393828

I. INTRODUCTION

SOIL moisture (SM) plays a crucial role in both the hydrological cycle and the energy balance at the Earth's surface, holding significant sway over the distribution of precipitation, incident radiation on Earth's land surface, and the regulation of various surface energy fluxes [1]. Its influence extends to crucial factors such as watershed runoff [2], radiation balance [3], carbon transport [4], and even the modulation of regional precipitation patterns [5]. Accurately depicting the spatial and temporal distributions of moisture within the topsoil layer is paramount for comprehending the Earth system, and it lays the foundation for meteorological, hydrological, and agricultural disaster forecasting and predictions. However, the inherently heterogeneous distribution of precipitation, soil textures, land cover types, and topography gives rise to substantial variability in SM on a spatial scale [6], [7]. This variation poses challenges for in situ SM observations, which often fall short in capturing watershed-scale estimations. In addition, the paucity of available in situ SM data and their limited spatial representativeness further hinder their applicability across various domains [8], [9].

To mitigate these constraints, land surface (hydrologic) models (LSMs) have been utilized to simulate SM across different soil layers, leveraging data from atmospheric general circulation models [10]. These models play a critical role in comprehending SM dynamics and have been extensively employed for operational drought monitoring. Nonetheless, the accuracy of LSM-generated SM is contingent upon external forcing and model parameterization, introducing uncertainties into the estimations [11], [12], [13].

Remote sensing emerges as a promising solution for monitoring regional and global SM, offering insights across different spatial and temporal scales. Microwave remote sensing, which is relatively unaffected by atmospheric conditions and cloud cover, holds a distinct advantage in assessing SM due to its penetrating capability [14], [15]. Satellite observations are utilized to extract SM information through both active and passive microwave remote sensing methods [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26]. SM retrieval methods based on microwave remote sensing can be grouped into five categories [27], [28]: physical, semiempirical, empirical, vegetation contribution-based, and change-detection-based models. Among them, physical models rely on the physics of microwave

interaction with soil, while semiempirical and empirical approaches use a mix of theory and observed data. Vegetation contribution models account for the influence of plants on the signal, and change-detection methods track variations in the signal over time to infer SM changes. However, these satellite-based observations are constrained by their sun-synchronous orbits, resulting in discontinuous spatial and temporal coverage. Furthermore, the presence of environmental factors, such as vegetation and snow, poses challenges to the accuracy and spatial resolution of the derived SM data [29], [30].

To overcome these challenges, the following three primary strategies have emerged:
1) fusing multisource satellite observations;
2) assimilating satellite data into LSMs;
3) incorporating SM-related land surface variables using machine learning techniques.

The first approach involves merging SM products from multimodal satellite sensors to enhance spatial consistency and resolution [31], [32], [33]. The climate change initiative (CCI) SM product, spanning decades and incorporating numerous microwave SM datasets, exemplifies this fusion strategy [34], [35]. The second strategy integrates satellite observations into LSMs through data assimilation, bolstering model accuracy, and forecasting capabilities [36], [37], [38], [39]. While this approach has demonstrated success, challenges persist in reconciling different assimilation operators, model calibrations, and temporal-spatial disparities [40], [41]. The third method utilizes machine learning techniques, including random forest, support vector machine, and artificial neural network, to enhance the accuracy of SM estimation by integrating relevant factors [42], [43], [44], [45], [46]. Owning to its potent capacity to fit a nonlinear relationship between independent and dependent variables [47], machine learning approaches, specifically random forest [42], [43], [48], [49], [50], [51], support vector machine [52], artificial neural networks [44], [45], [53], [54], [55], convolutional neural network [46], [54], [56], and long short-term memory neural network [57], [58] were adopted to improve and estimate SM in recent years. For instance, Long et al. [43] generated long-term daily seamless SM datasets based on the random forest model by combining high quality land surface temperature (LST), normalized difference vegetation index (NDVI), surface albedo, precipitation, soil texture, CCI, and in situ SM observations over the Hebei Province, China. Taking the features of the atmospheric and geophysical information derived from satellite and ground-based observations into account, Abbaszadeh et al. [59] proposed a random-forest-model-based scheme to reconstruct the level 3 daily SMAP radiometer SM product over the continental United States (CONUS). Yao et al. [60] transferred the merits of SMAP to AMSR-E/2 and developed a global daily SM product using artificial neural networks. Li et al. [49] presented a long-term SM dataset of 1-km spatial resolution derived through machine learning trained by the in situ SM observations from 1789 stations worldwide and 15 SM-related variables from ERA5-Land, USGS land cover, and MODIS leaf area index over China. SMAP-HydroBlocks (SMAP-HB) SM dataset leverages microwave satellite data, high-resolution (HR) land simulations, radiative transfer modeling, and machine learning, along with ground observations. By combining these approaches, SMAP-HB provides estimation of surface SM at 30-m resolution across the USA. However, most of these machine-learning-based methods strongly rely on the locations of the ground-based observations [61]; taking these observation data as training targets for machine learning is obviously lacking of regional representative but ignoring local distribution characteristics and spatial heterogeneity of SM [62], [63], which is difficult to extend to global scales for supervised machine-learning-based models due to the uneven distribution of observational sites and spatial heterogeneity of SM.

In order to overcome these constraints, we propose a novel approach for reconstructing a HR worldwide SM product. Leveraging the comprehensive ESA CCI SM dataset and SMAP assimilation products, we employ variational inference and a conditional variational autoencoder (CVAE) to seamlessly merge these datasets and enhance the precision and coverage of SM estimations. This approach overcomes challenges associated with missing values, limited use of multisource observations, and spatial distribution disparities. By incorporating the strengths of both datasets, we achieve a globally consistent and high-quality SM product.

The rest of this article is organized as follows. Section II introduces the ESA CCI, SMAP L4 SM products, and in situ dataset, outlining their roles in our proposed scheme. Section III details the main components of the proposed model. Sections IV and V present experimental settings, results, and validation processes, showcasing the effectiveness of our approach. Section VI delves into discussions regarding the comparative performance of our scheme against alternative methodologies. Finally, Section VII concludes this article.

II. DATA

A. ESA CCI and SMAP L4 SM Product

The SM CCI project, as a component of the ESA Program on Global Monitoring of Essential Climate Variables, was started in 2010 as an element of the CCI, which produced an updated SM product (ESA CCI SM) each year. The ESA CCI SM product integrates multiple SM data from both active and passive microwave sensors to create three products: ACTIVE, PASSIVE, and COMBINED product that utilizes both active and passive microwave data. We collected COMBINED microwave product of CCI as the basic data source for SM product reconstruction, which, respectively, rescaled whole level 2 observations into a common model-based climatology. This dataset spans 40 years with spatial resolution of 0.25°, available from November 1978 to December 31st, 2021.

On January 31st, 2015, NASA launched the SMAP satellite, which became the world's inaugural space-based system for monitoring SM at a depth of roughly 5 cm below the Earth's surface, employing a combination of both active and passive microwave observations at the L band frequency. Unfortunately, the active radar sensors onboard the satellite were damaged on July 7th, 2015, therefore, currently only the passive SM data are available. SMAP SM dataset include four levels of products. In this study, we collected SMAP Level 4 (L4) Global 9-km EASE-Grid Surface and Root Zone Soil Moisture Analysis
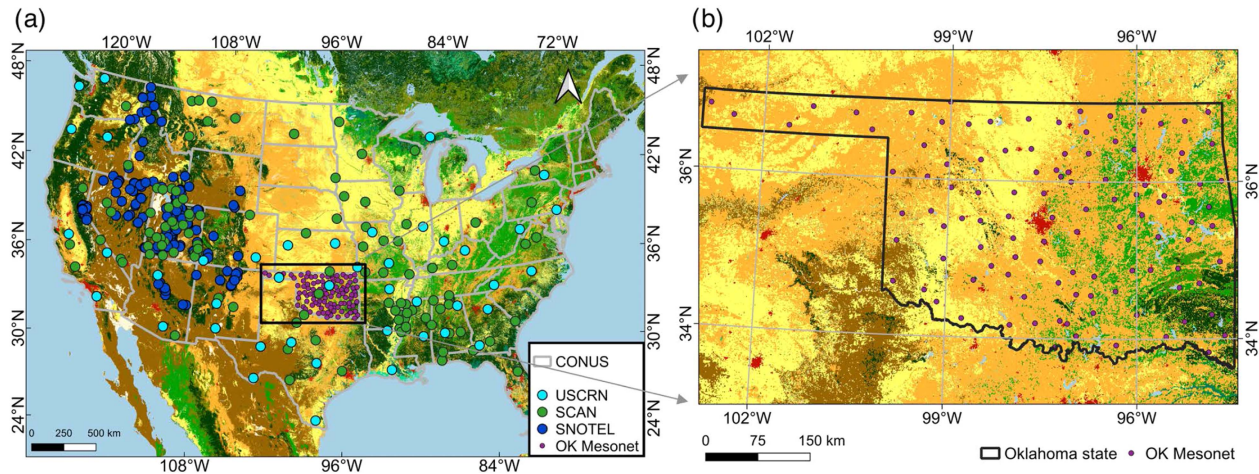
Fig. 1. (a) Geolocations of the sparsely distributed in situ SM observations for validations over the CONUS and (b) geolocations of the sparsely distributed in situ SM observations for validations over the Oklahoma state, USA.

TABLE I
OVERVIEW OF THE SPARSELY DISTRIBUTED NETWORKS FOR SM IN-SITU OBSERVATIONS

Network name	Area	Number of sites	Sensor depth(cm)
SCAN	North America	120	5
USCRN	North America	51	5
SNOTEL	North America	270	5
OK Mesonet	Oklahoma	109	5
SMOSMANIA	Europe	16	5
TAHMO	Africa	14	10
LAB-net	South America	3	5
All networks		583	

dataset with spatial resolution of 9 km updated from National Snow and Ice Data Center (NSIDC).

The spatial coordinate system of SMAP L4 and ESA CCI SM products was unified as GCS WGS84, and the global coverage of ESA CCI SM product is about three days, therefore, the three-day mean composition was proceeded for ESA CCI SM products.

B. In Situ SM Dataset

SM validation primarily relies on comparisons with independent in situ measurements. These measurements can be categorized into two main types.

1) *Core validation sites:* These sites provide accurate SM estimates at specific spatial scales relevant to models and satellite retrievals, but are limited to a restricted set of climatic and land cover conditions.
2) *Sparse networks:* These networks offer SM estimates at individual point locations within grid cells, encompassing a wider range of environmental conditions. However, they may not capture the full spatial variability of SM.

To evaluate the accuracy of our simulated long-term SM dataset (2016–2021), we employed in situ observations from two sources: core validation sites and sparsely distributed representative sites worldwide (as described in [64]). Typically, sparse network SM observations were derived mostly by single

sensor positioned inside Table I, consisting of the soil climate analysis network (SCAN), U.S. climate reference network (USCRN), automated snow telemetry (SNOTEL), Oklahoma mesonet (OKM), trans African hydro-meteorological observatory network (TAHMO), Australian moisture monitoring network (OZN), and LAB-net. Fig. 1 exhibited the geolocations of sparse networks in the CONUS where many locally dense distributed sensor networks were set up for accurate SM at the grid cell scale of the SM product.

The SM datasets from core validation sites used in present study include those from the United States Department of Agriculture watershed sites (Walnut Gulch, Little Washita, Fort Cobb), Reynolds Creek and Little River sites from NSIDC, Tibetan Plateau sites (Maqu), Remedhus Network sites (REMEDHUS), Benin site in AMMA-catch network, Yanco sites, and the Shandian River Basin Network (Shandian River site), which were presented in Fig. 2 and summarized in Table II.

SM Datasets obtained from SCAN, USCRN, SNOTEL, REMEDHUS, TAHMO, and LAB-net sites were made available by the international soil moisture network (ISMN). Our proposed SM dataset reconstruction scheme belongs to an unsupervised model category, and all the in situ SM datasets were used only for validation of the model. When validations of the in situ SM dataset were performed, stations with integrity of time series of SM dataset greater than 80% were selected as the valid in situ dataset.

III. METHODOLOGY

A. Our Proposed SM Reconstruction Scheme

Based on the CVAE [65], a worldwide seamless long-term SM product reconstruction scheme was developed. The process of variational inference ensures the learned distribution aligns with a desired distribution. First of all, we assumed that the HR SM map of SMAP L4 product is $\mathbf{X_{HR}} \in \mathbb{R}^{m \times n}$, the high precision but low resolution (LR) SM map of ESA CCI product is $\mathbf{X_{LR}} \in \mathbb{R}^{\alpha m \times \alpha n}$, where (m,n) is the dimension of the map and α is the sampling factor. D is the degradation kernel that represents
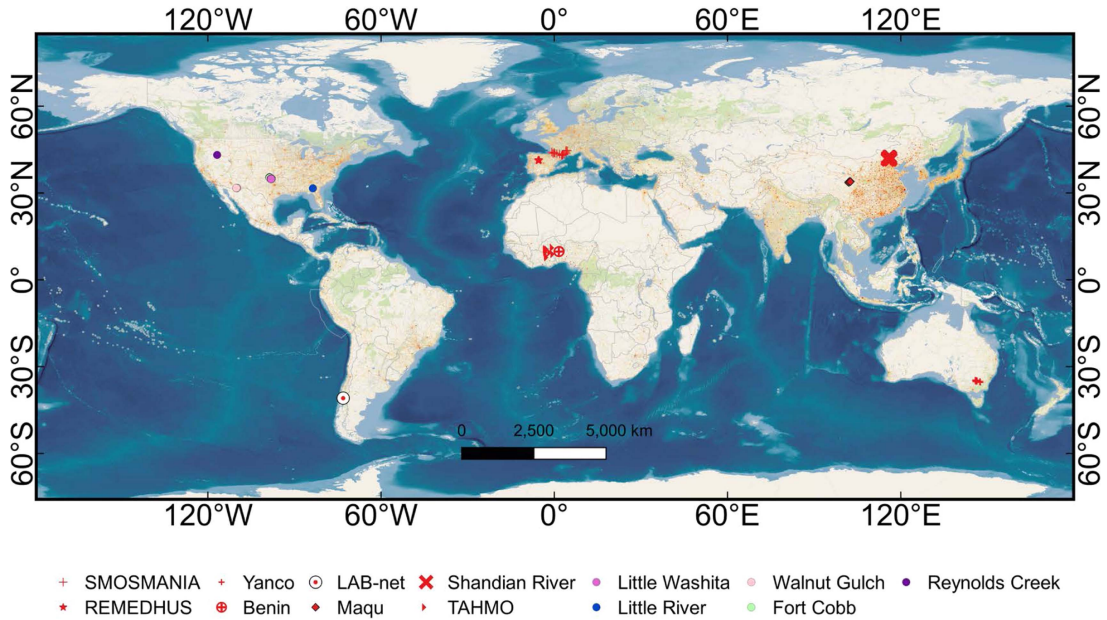
Fig. 2.   Spatial distribution of core validation sites all over the world.

TABLE II
OVERVIEW OF CORE VALIDATE SITES

Site name	Location	Climate regime	Number of sites	IGBP land cover	Sensor depths (cm)
Fort Cobb	North America	Temperate	12	Croplands	5
Walnut Gulch	North America	Arid	14	Shrub open rangeland	5
Little Washita	North America	Temperate	19	Grasslands/Cropland mosaic	5
Reynolds Creek	North America	Arid	8	Grasslands	5
Little River	North America	Temperate	8	Cropland/natural mosaic	5
REMEDHUS	Europe	Temperate	19	Croplands	5
Shandian River	Asia	Temperate	25	Grasslands	5
Maqu	Asia	Arid	8	Grasslands	5
Yanco	Australia	Semi-arid	9	Croplands/Grasslands	5
Benin	Africa	Arid	7	Savannas	5

various types of multiplicative noise and spatial sampling. The relationship between $\mathbf{X_{LR}}$ and $\mathbf{X_{HR}}$ can be expressed as

$$\mathbf{X_{LR}} = \mathrm{D} \otimes \mathbf{X_{HR}} + \mu \qquad (1)$$

where $\otimes$ stands for the convolution operator, μ is the additive noise, and both $\mathbf{X_{LR}}$ and $\mathbf{X_{HR}}$ are three-day composited SM maps. To reconstruct a high spatial resolution and high accuracy SM map, (1) can be solved as the maximum a posterior (MAP) problem as follows:

$$\widehat{\mathbf{X_{HR}}} = \underset{\mathbf{X_{HR}}}{\arg\max} \log P\left(\mathbf{X_{LR}} \mid \mathbf{X_{HR}}\right) + \log P\left(\mathbf{X_{HR}}\right) \quad (2)$$

where $\widehat{\mathbf{X_{HR}}}$ is the reconstructed SM map. $\log P(\mathbf{X_{LR}} \mid \mathbf{X_{HR}})$ is the log-likelihood of the observed SM given HR SM map and $\log P(\mathbf{X_{HR}})$ as the priori information of HR SM map used for model optimization. The temporal resolution of ESA CCI synthetic SM dataset is three days, while the original SMAP L4 dataset is 3 h. One scene of ESA CCI SM product corresponds to 24 scenes of SMAP within three days, in order to better learn the corresponding SM spatial variability and to utilize a larger training sample size, each CCI patch with its corresponding 24 SMAP patches was fed input into the model training separately for optimizing the solving process described in (2) to obtain

multiple reconstructed $\widehat{\mathbf{X_{HR}}}$

$$\widehat{\mathbf{X_{HR}}_i} = \underset{\mathbf{X_{HR}}}{\arg\max} \log P\left(\mathbf{X_{LR}}, \mathbf{R}_i \mid \mathbf{X_{HR}}\right) + \log P\left(\mathbf{X_{HR}}\right)$$

$$(3)$$

where $\mathbf{R_i} \in \mathbb{S}_{\mathrm{SMAP}}$ and $\widehat{\mathbf{X_{HR}}_i} \in \mathbb{S}_{\mathrm{SM}}$, and we defined the reconstructed SM space as $\mathbb{S}_{\mathrm{SM}}$ and the reference SMAP space as $\mathbb{S}_{\mathrm{SMAP}}$. Given reference SMAP time series $\mathbf{R_i}$, the intended reconstructed SM can be sampled from $\mathbb{S}_{\mathrm{SMAP}}$ space. To maintain the high spatial resolution of the reconstructed SM map, we used multiple scenes of the SMAP maps as a spatial distribution constraint variable while rebuilding high accuracy SM maps.

The entire architecture of the proposed scheme was presented in Fig. 3, which exhibited the steps for training and for the model inference process, the model's loss computation part was removed, with the same flow of inputs and computation graphs as training. The scheme can be subdivided into three parts with order of: SM encoder; CVAE; and SM decoder. In the training phase, the proposed scheme took the reference SMAP sequence ($\mathbf{R_i}$) and the composited CCI ($\mathbf{X_{LR}}$) as input. From composited CCI and reference SMAP series, feature maps F_X and F_{Ri} were extracted using the SM encoder. For conditional
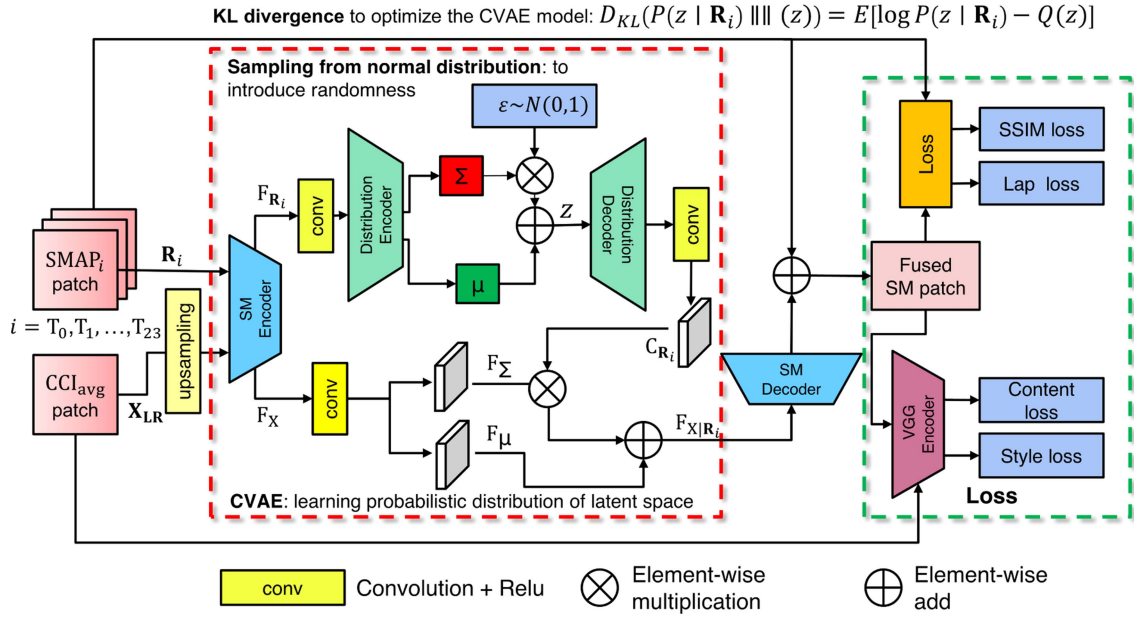
Fig. 3. Training procedure of the proposed mode. Feature encoder extracts latent variables z from reference features, and z represents the updated target distribution with new mean and variance. Feature decoder transforms z into a conditional feature map CR, and CR is used to guide the generation of new SM maps. Training target: Optimizes the model to match the learned distribution with a normal distribution, and KL divergence is employed to measure the difference between the learned and expected distributions.

feature sampling, reference SMAP features F_{Ri} were transferred to a hidden space $\mathbb{S}_{SMAP}$ using the CVAE. In other words, the CVAE produces synthetic features by mapping random feature vectors, thereby generating conditional features denoted as C_{Ri}. Concurrently, the CCI feature maps F_X were disassembled into separate mean and variance feature maps F_μ and F_Σ, respectively. These two individual components were subsequently integrated with the conditional features through a computational process, yielding the estimated features. The SM decoder then took the estimated features to rebuild the SM maps. In short, the CVAE uses latent variables from Encoding process to capture the essential information from reference SMAP data. Decoding process leverages the latent variables and conditional features to generate new SM data. The process of variational inference as aforementioned guides the learned distribution aligns with the desired CCI SM's distribution.

B. SM Encoder for Fusing SMAP and CCI SM Product

We took use of the pretrained VGG19 model [66] to build the encoder of the proposed framework, keeping only the convolutional layers while removing the fully connected layers. The choice of a smaller convolution kernel is the key of VGG; in addition to using less computation and introducing more nonlinearity to produce the same receptive field, the VGG's 3×3 convolution kernel can also improve the model's capability to fit data. As the result, we trained VGG19 to extract feature maps from SMAP ($F_{Ri} = G(R_i)$) and CCI ($F_X = G(X_{LR})$), where G denotes the feature extraction procedure of VGG. In this process, SM map of the ESA CCI SM product will be up sampled to four times of its original resolution, i.e., 0.0625°. And SM maps of the SMAP L4 SM product will be resized from 9 km

spatial resolution to 0.0625°. Using a pretrained VGG encoder has advantage for the following two reasons:

1) VGG was trained with many images with 2-D structure, resulting in generalized extracted feature maps for spatial grids with various contents;
2) With VGG, SMAP and CCI SM are projected to the same feature domain, which is convenient to fuse their features.

C. Conditional Variational AutoEncoder (CVAE)

The CVAE, as shown in Fig. 3, involves determining the mean and variance of the target distribution by encoding procedure with feature encoder, and then, reconstructing the outputs through decoding process with feature decoder. By measuring how closely the rebuilt distribution resembles the target distribution, the parameters involved are adjusted. In this study, employing the CVAE, the model can learn the hidden distribution by encoding the reference features from SMAP into a latent space as a latent variable z [67]. The reference features can be transferred by the feature decoder as conditional feature map CR. The feature encoder and decoder are two components of the variational inference process (specific components are shown in Fig. 4). The objective of variational inference is aimed to train the generative model for the reference SM map derived from SMAP data, which can be written as a normal distribution model as $P(z \mid R_i) = N \sim (z; \mu(R_i), \Sigma(R_i))$, where μ and $\Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$. To put it another way, we gather the spatial correlations and convert them into a probability model. Using a normal distribution $Q(z) = N \sim (0, I)$, encoded features are sampled from the reference model with $z = \mu + \epsilon \times \sigma$ to introduce randomness. KL divergence is used in the sampling
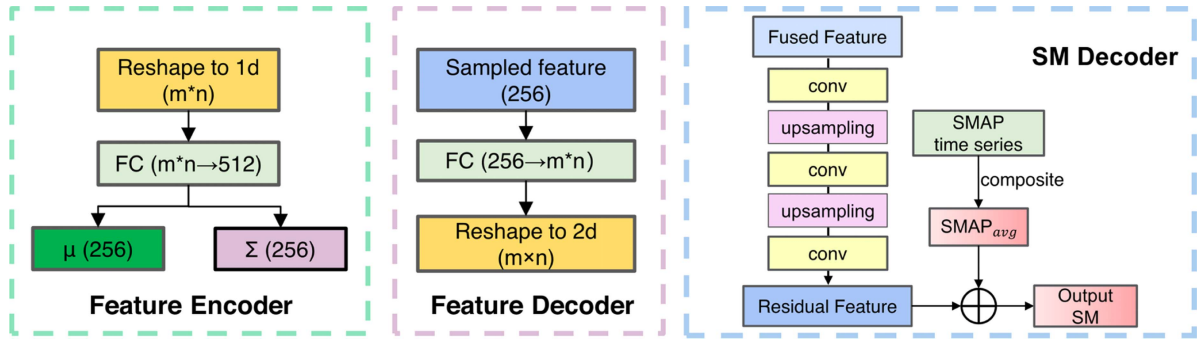
Fig. 4. Structure of the feature encoder, feature decoder, and SM decoder in the proposed CVAE.

process to optimize the model and ensure that the learned probability model is reasonably approximate to a normal distribution as follows:

$$D_{\text{KL}}(P(z \mid \mathbf{R}_i) \| \| (z)) = E[\log P(z \mid \mathbf{R}_i) - Q(z)]$$

$$= \frac{1}{2}\left(-\sum_i \left(\log \sigma_i^2 + 1\right) + \sum_i \sigma_i^2 + \sum_i \mu_i^2\right). \quad (4)$$

Like the conditions made previously for reconstruction in the training phase, we continue to employ the feature encoder during inference to derive a specific prior distribution, $P(z \mid \mathbf{R_i}) = N \sim (z; \mu(\mathbf{R_i}), \Sigma(\mathbf{R_i}))$. The conditional feature maps CR are then produced by returning the learned distribution to the spatial domain using one convolution block (we also interpolate and resample it to the same size as the input CCI SM maps). To propagate the conditional features CR onto the CCI SM feature map, we utilize a convolutional block to learn the spatial statistics of mean and variance for the CCI SM maps as F_μ and F_Σ. The fused features are then denoted by the formula $F_{\mathbf{X}|\mathbf{R}} = \text{CR} \cdot (1 + F_\Sigma) + F_\mu$.

D. SM Decoder For the HR SM Map

Finally, the SM Decoder, as exhibited in Fig. 4, learns to enhance the SM map from the reconstruct feature $F_{\mathbf{X}|\mathbf{R}}$. Like the VGG encoder, the SM decoder stacks three convolutional layers and two bilinear interpolation operators to restore the fused feature maps to the residual correction map. The residual map is summed with the synthesized SMAP SM map to obtain the final reconstructed SM map.

E. Training for Reconstruction

To train the proposed model to reconstruct the SM maps with high spatial resolution and high precision, a compound loss function is adopted in this study, and it is composed of six parts: KL loss, content loss, style loss, SSIM loss [68], Lap loss, and total variation loss. The total loss for SM reconstruction can be formulated as

$$L_{\text{total}} = \lambda_{\text{KL}} L_{\text{KL}} + \lambda_{\text{content}} L_{\text{content}} + \lambda_{\text{style}} L_{\text{style}} +$$

$$\lambda_{\text{SSIM}} L_{\text{SSIM}} + \lambda_{\text{Lap}} L_{\text{Lap}} + \lambda_{\text{TV}} L_{\text{TV}}. \quad (5)$$

1) KL loss is employed for the purpose of optimizing the variational lower bound of the logarithmic likelihood function, as formulated in (4).
   Utilizing style and content losses for style transfer is an effective method to transmit the reference distribution to the target while keeping the spatial information unchanged. It is appropriate for SM reconstruction using references as well. Our objective is to reconstruct the SM product with data precision close to the CCI and with spatial resolution equivalent to the reference SMAP. To put it differently, we attempt to guarantee that the reference features are transferred to the anticipated SM.
2) With regard to content loss, we use VGG-19 to reconstruct the SM and extract features for the CCI (the feature map from the last layer is taken from VGG-19). These features are referred to as W_{CCI} and $W_{(\text{SM})}$, respectively. We also have the loss between reconstructed SM and CCI SM for the pixel-wised difference. Total content loss is as follows:

$$L_{\text{content}} = \left\| W_{\text{CCI}} - W_{\widehat{\text{SM}}} \right\|_1 + \left\| \text{CCI} - \widehat{\text{SM}} \right\|_1. \quad (6)$$

3) We calculate the mean and variance of the feature maps V_{CCI} and $V_{(\text{SM})}$ of the CCI and the rebuilt map RSM from different layers of VGG-19 to compare how well the reconstructed features match with the extracted features as follows:

$$L_{\text{style}} = \sum_i \| \text{mean} \left(V_{\text{CCI}}^i - V_{\text{SM}}^i \right) \|_1$$

$$+ \| \text{var} \left(V_{\text{CCI}}^i - V_{\text{SM}}^i \right) \|_1 \quad (7)$$

where i is the index of the VGG-19 layers, and mean and var are procedures for calculating the mean and variance of the feature maps.
4) A perception-based model called structural similarity index measure (SSIM) is used to compare two images by taking image deterioration into account as a perceived change in structural information, which leads to the definition of the SSIM loss function as

$$L_{SSIM} = \frac{\left(2\mu_{\text{SMAP}}\mu_{\widehat{\text{SM}}} + C_1\right)\left(2\sigma_{\text{SMAP\_}\widehat{\text{SM}}} + C_2\right)}{\left(\mu_{\text{SMAP}}^2 + \mu_{\widehat{\text{SM}}}^2 + C_1\right)\left(\sigma_{\text{SMAP}}^2 + \sigma_{\widehat{\text{SM}}}^2 + C_2\right)} \quad (8)$$

$C_1 = (0.01 * l)^2, C_2 = (0.03 * l)^2, C_3 = C_2/2$, where 1 is the specified range value of SM, and μ_{SMAP}, $\mu_{\widehat{\text{SM}}}$ are the mean value of SMAP and the reconstructed SM map, respectively. σ^2_{SMAP}, $\sigma_{\widehat{\text{SM}}}$, and $\sigma_{\text{SMAP\_}\widehat{\text{SM}}}$ are the variance of SMAP, variance of $\widehat{\text{SM}}$ and covariance between SMAP and $\widehat{\text{SM}}$ respectively.

5) The Laplacian loss quantifies the difference between the SMAP and the rebuilt SM map's Laplacian matrix and, consequently, expresses the difference in detail structure between two maps. The loss can be written as

$$L_{\text{Lap}} = \sum_j 2^{2j} \left\| L^j(\text{SMAP}) - L^j(\widehat{\text{SM}}) \right\|_1 \quad (9)$$

where $L^j(\text{SMAP})$ and $L^j(\widehat{\text{SM}})$ are the jth level of the Laplacian pyramid representation of SMAP and the reconstructed SM map, respectively.

6) The first-order horizontal and vertical pixel gradients are calculated using the total variation loss to promote smoothness in the reconstructed SM maps.

$$L_{\text{TV}} = \sum_{m,n} ((\widehat{\text{SM}_{m,n-1}} - \widehat{\text{SM}_{m,n}})^2$$
$$+ (\widehat{\text{SM}_{m-1,n}} - \widehat{\text{SM}_{m,n}})^2)^{1/2} \quad (10)$$

where m and n are the spatial indices of the SM map, respectively.

IV. EXPERIMENTS SETTINGS

A. Dataset and Settings

Since the ESA CCI SM product cannot cover the entire world seamlessly in a single day, SM dataset from the last and the next days are also utilized, i.e., the data were composited by moving-average windows over three days. Although the data were temporally composited, ice, snow, frozen soil, water bodies, and so on are considered incapable of retrieval of SM by satellite observations. As a result, if grid points located in these regions were invalid values, they are removed from the training SM map. Since such land cover types take account of relatively minor part of the Earth surface, the regional statistical distribution properties remain unaffected. The training samples were prepared for model training after the preprocessing as described previously. However, when performing model validation and comparisons, site measurements, SMAP, and CCI were still used as the validation set when valid observations existed in the snow regions. Note that the proposed deep-learning-based model is an unsupervised model without in situ observation data as labels involved, and the CVAE reconstruction model adjusts the model parameters based on the similarity between the statistical distribution of the reconstructed data and the statistical distribution of the input/reference data. Therefore, the whole datasets of SMAP and CCI SM products was input to the proposed model for "map to map" training rather than pixel-wise training.

Since the objective of our study is aimed to reconstruct a globally seamless SM product with high data precision and
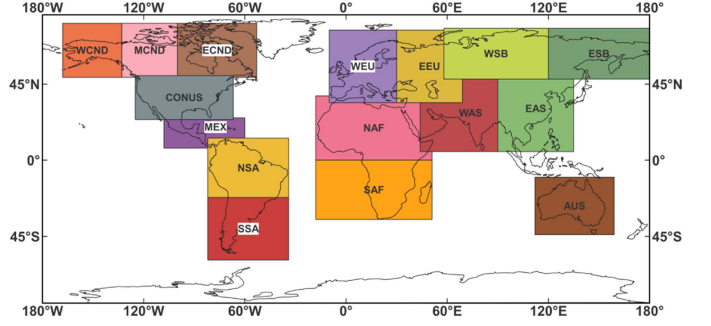

Fig. 5. Colorful boxes correspond to divided regions with shortened letters denoting the region names.

spatial resolutions compared to the currently available SM product, for the sake of fully taking regional heterogeneity and the reduction of high demand of computer hardware into account, we subdivided the global land surface into 16 parts (regions such as Japan, Indonesia, New Zealand and Greenland whose spatial distribution is relatively fragmented, were not considered for SM reconstructions; these regions can be processed individually in small-scale chunks, whose processing can be carried out in the subsequent data updating), as shown in Fig. 5. Regarding in situ SM sites for validations, we removed sites with no more than 80% time series completeness and retained sites with more valid data values.

With a windows server equipped with an Intel i9-10900 K processor and a GeForce RTX 2080Ti GPU, all the trials for this study were conducted. The model was implemented by the PyTorch framework. In the training phase, the learning rate was set to 0.0001 and Adam was employed as an optimizer. We trained the model for 5×10^4 iterations totally. The proposed model comprises a total of 122 713 026 parameters are trainable. It takes 18 minutes on the server to complete the training of one epoch of the soil moisture map covering the entire region of the continental United States with a timeseries length of 728.

B. Validation and Assessment

This study employed the correlation coefficient (CC) and unbiased root mean square error (ubRMSE) as the criterions to evaluate the performance of the model. The degree of linear correlation between variables can be effectively determined using the CC.

$$\text{CC}_{x,y} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \quad (11)$$

where x_i and y_i are the predicted and the measured SM on the ith day, respectively; $\bar{x}$ and $\bar{y}$ are arithmetic average of the predicted and the measured SM, respectively; and n denotes the number of days during the study period.

$$\text{bias} = \bar{x} - \bar{y}. \quad (12)$$

The degree of overestimation and underestimation of observations can be measured as bias. In contrast, the root mean square error (RMSE) could numerically show how dispersed

the predictions are

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}}. \qquad (13)$$

SM deviations may vary with the seasons. Considering this slowly changing deviation, ubRMSE reflects the error in the measured and the predicted SM values by removing the past average seasonal cycle variation information

$$\text{ubRMSE} = \sqrt{\frac{\sum_{i=1}^{n}\{(x_i - \bar{x}) - (y_i - \bar{y})\}^2}{n}}. \qquad (14)$$

Therefore, these metrics enable us to investigate the model's competence in SM reconstruction over the case study regions systematically and comprehensively.

For sparse networks, a single measuring station within a raster grid provides predicted SM for validation. In contrast, core validation networks contain multiple in situ sites per grid, using the average SM. Diverse SM products have varying precision, temporal, and spatial resolutions, often with HR grids for each LR grid. In this study, to compare different scale SM products accurately [69], we adopt the evaluation strategy from [64]. When comparing SM estimates at 0.25° resolution against corresponding in situ measurements, bilinear interpolation is employed to extract SM values with the reference site locations. For reference pixels defined at 0.0625° resolution, HR SM estimates are aggregated using an area-weighted averaging approach to match the coarser spatial scale at 0.25°resolution. Consequently, the SM values of different spatial resolution products corresponding to each site can thus be calculated.

To verify the superiority of our proposed method, we compared it with the cumulative distribution function (CDF) matching algorithm [70]. This algorithm utilizes the cumulative distribution frequencies of two datasets to establish a mapping relationship between them. We applied it to both SMAP L4 and CCI data by dividing the cumulative distribution curves into six parts at seven frequency points (0, 10%, 30%, 50%, 70%, 90%, and 100%). The CDF matching adjusted the target SMAP L4 data to generate CDF SM data (CDFsm) that approximates the distribution of the CCI data, which is further discussed in Section IV-A.

V. Experimental Results

Through the trained CVAE, the seamless global three-day composited SM products from April 1, 2015 to December 31, 2021 were eventually generated. The reconstructed SM product saved in NetCDF format can be downloaded online[1] for free. Codes were made available online.[2]

To investigate the variability of the proposed scheme, we quantified the accuracy of the reconstructed SM product (CVAEsm) in terms of both its spatial and temporal accuracy using in situ observations both from sparse networks and core validation sites over the world (as shown in Figs. 1 and 2). Comparisons were made between the reconstructed SM product with

[1][Online]. Available: https://doi.org/10.5281/zenodo.800060
[2][Online]. Available: https://github.com/YangTze7/CVAE-SM

TABLE III
METRICS FOR THE COMPARISONS BETWEEN CVAEsm, CCIsm, AND SMAPsm OVER SPARSE NETWORK IN THE CONUS

Sparse network	CVAE		CCI		SMAP L4	
	CC	ubRMSE	CC	ubRMSE	CC	ubRMSE
SCAN	0.781	0.087	0.764	0.090	0.754	0.089
USCRN	0.678	0.099	0.704	0.097	0.584	0.118
SNOTEL	0.516	0.103	0.479	0.105	0.498	0.103
OK mesonet	0.637	0.071	0.600	0.074	0.617	0.077
Average	**0.653**	**0.090**	0.636	0.092	0.613	0.097

the original CCI SM (CCIsm) and SMAP L4 SM (SMAPsm) products (being the baseline products). Evaluation on the quality of the reconstructed SM product was done by use of the average daily in situ observations within the grid cells where all the SM products are available at that time simultaneously.

A. Sparse Network Validation

In situ SM observations from 2016 to 2020 were used to validate the reconstructed SM product. The in situ observation sites are located in evenly distributed areas across the CONUS with diverse topography, land use, and soil types.

1) Sparse Networks in the CONUS:

a) Temporal statistics: Metrics for the comparisons between CVAEsm, CCIsm, and SMAPsm over sparse network in the CONUS were examined, the results were listed in Table III. It was noted that the reconstructed SM product CVAEsm performed better according to various criterions.

Similarly, the performance of the CVAEsm over the sparse networks in the CONUS was shown in Fig. 6. It is observed that the averaged CC of CCIsm and SMAPsm are 0.636 and 0.613, respectively, while the value for CVAEsm is 0.653. Especially, CVAEsm performed better than CCIsm and SMAPsm regarding temporal correlation in regions with mountains and a prevalence of snow (e.g., at SNOTEL sites), with CC of 0.516, 0.479, and 0.498, correspondingly. To demonstrate the SM dynamics of these datasets, the three-day composited time series of SM derived from CVAEsm, CCIsm, SMAPsm, and CDFsm in the sparse in situ observation sites was shown in Fig. 7.

The results indicated that the CVAEsm presented the highest median and mean temporal correlation of 0.733 and 0.700 among the sparse sites in CONUS, respectively. The reconstructed SM product CVAEsm suggested mean improvement of 0.012 and 0.055 in temporal correlation concerning the SMAPsm and CCIsm, with ubRMSE reduction of 0.001 and 0.003 relative to the SMAPsm and the CCIsm, respectively.

In the CONUS, ubRMSE values and correlation of the CVAEsm are regionally homogeneous, as can be seen in Fig. 8(a)–(c), which the reconstructed SM product was improved over most observational sites in comparison to the standard products. The differences in the assessment metrics between the CVAEsm and the baseline products were exhibited in Fig. 8(d)–(l). Regarding the relative difference of CCs, the dots in CCI and SMAP are dominated by blue tones, indicating that CCs of most regions are improved by the CVAE. Regarding the relative difference of ubRMSE, the points in CCI and SMAP
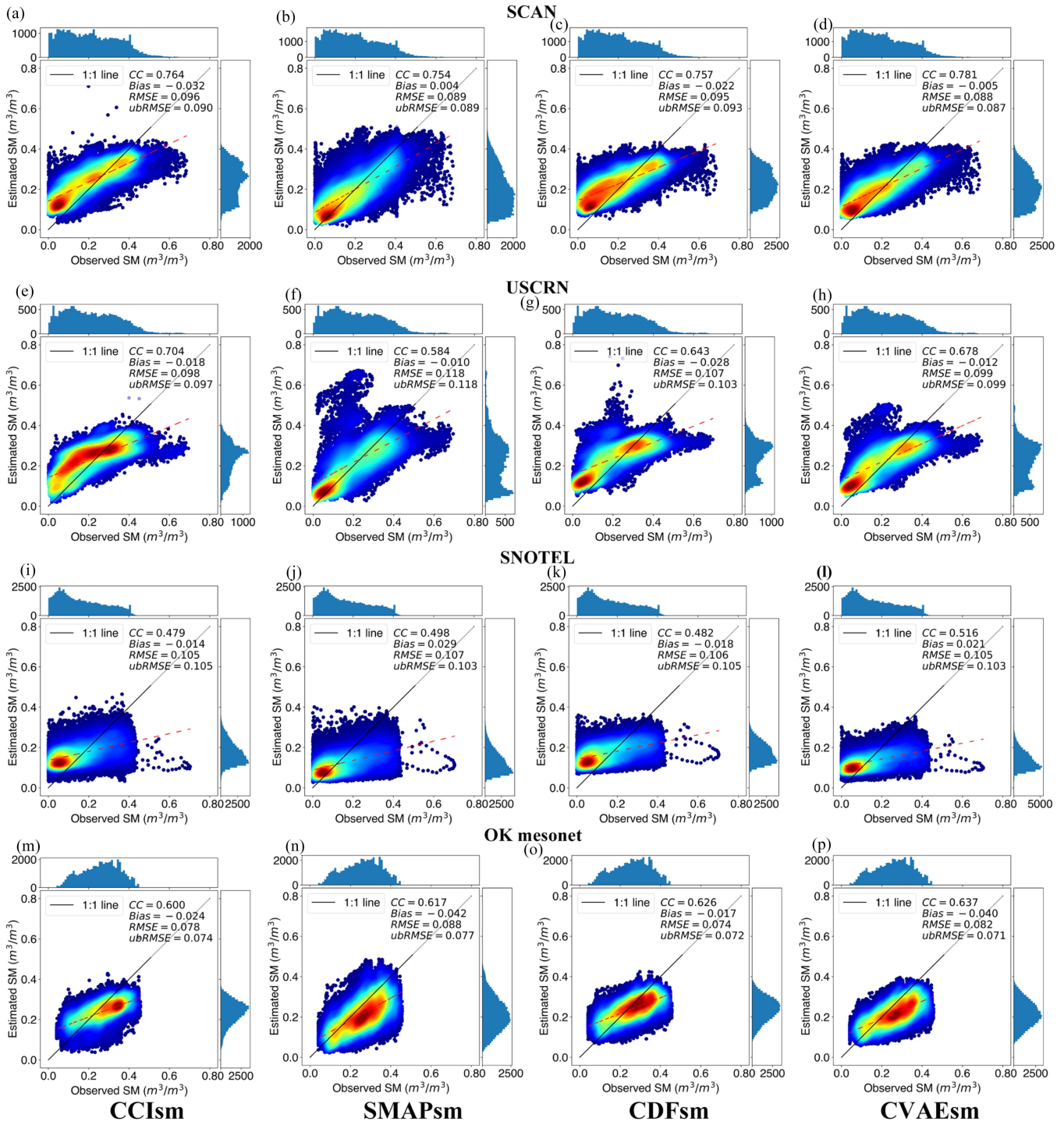
Fig. 6.   Scatterplot of CCIsm, SMAPsm, CDFsm, and CVAEsm versus sparse in situ SM over the CONUS.

are dominated by red tones, indicating that ubRMSE of most regions is lowered through the CVAE.

2) Spatial Statistics: The spatial correlation for SM products every three days was determined to evaluate to which degree the SM products are of spatial representative by comparing the daily SM products at grid-cell level with daily in situ observations across CONUS. Fig. 9 presented variations of the spatial correlations between each of the SM products of SMAPsm, CCIsm, and CDFsm with the in situ observations for each time step over

the CONUS. It can be observed that the CVAEsm presented the higher spatial correlation (0.584) across the CONUS than CCIsm and SMAPsm, and the CVAEsm showed mean improvement of 0.036 and 0.013 in spatial correlation with respect to the SMAPsm and CCIsm, as shown in Fig. 9, over the time span of time series of SM products from 2016 to 2019. After three days of average composition the time series of SM dataset is 485 time steps. The CVAEsm exhibits greater spatial correlation compared to SMAPsm and CDFsm across the majority of time
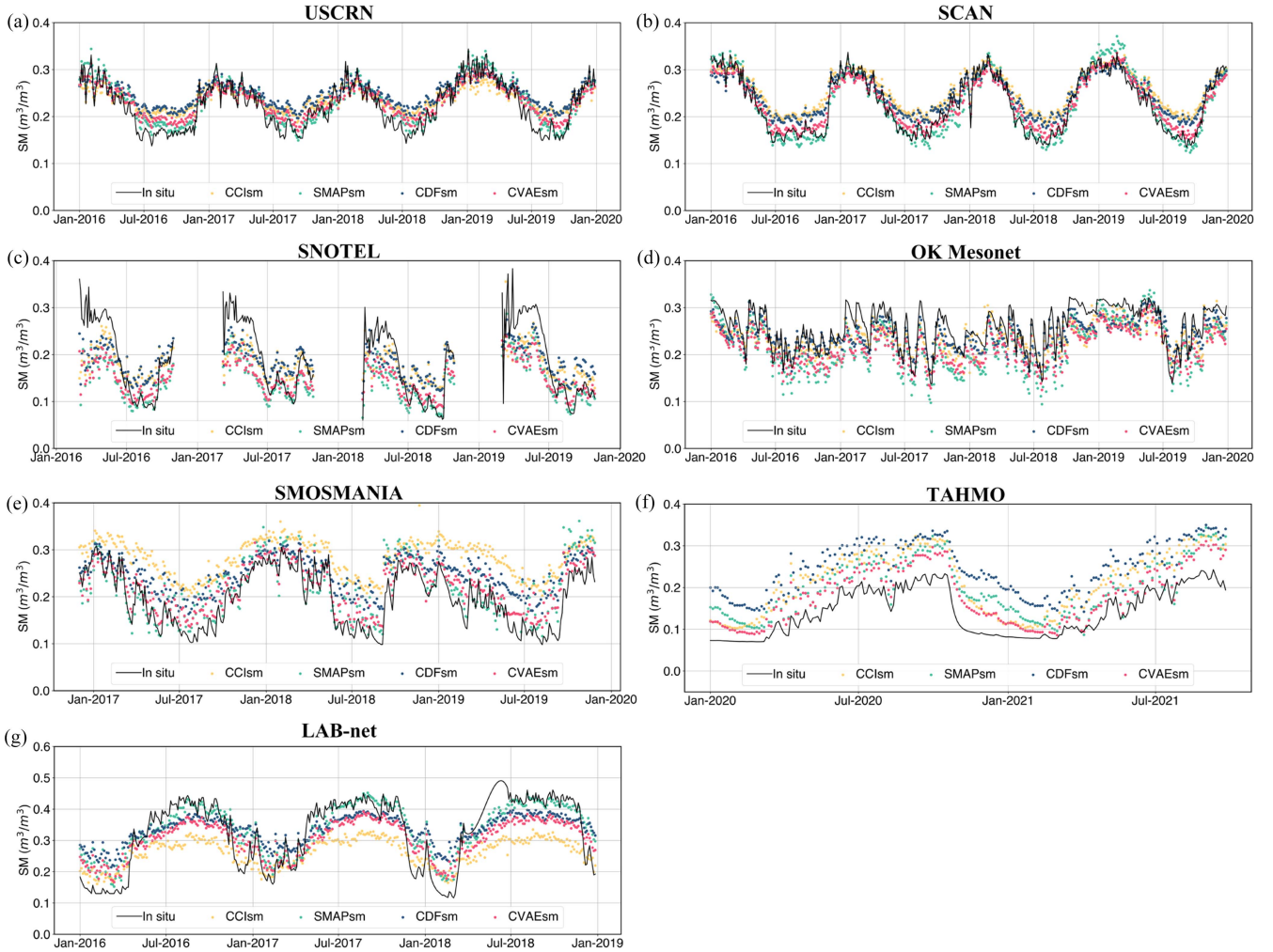
Fig. 7.   Three-day composited time series of SM derived from CVAEsm, CCIsm, SMAPsm, and CDFsm in the sparse in situ observation sites all over the world.

TABLE IV
METRICS FOR THE COMPARISONS BETWEEN CVAEsm, CCIsm, AND SMAPsm
OVER SPARSE NETWORK OUTSIDE THE CONUS

Sparse network	CVAE		CCI		SMAP L4	
	CC	ubRMSE	CC	ubRMSE	CC	ubRMSE
SMOSMANIA	0.937	0.024	0.913	0.029	0.939	0.024
TAHMO	0.726	0.053	0.698	0.058	0.646	0.061
LAB- net	0.925	0.062	0.917	0.074	0.911	0.052
Average	**0.863**	**0.046**	0.843	0.054	0.832	0.046

steps, and the spatial correlation of CVAEsm is higher than that of CCI at 292 time steps and lower than that of CCI at 193 time steps

3) Sparse Networks Outside the CONUS: The performance of the CVAEsm over the sparse sites outside the CONUS (including LAB-net, SMOSMANIA and TAHMO networks) were analyzed with comparisons between the SM products and the in situ SM observations. The results were illustrated in Fig. 10 and a statistical matrix summarized for the analyses was listed in Table IV.

As portrayed in Fig. 10, the scatters showing the relationship between the estimated SM products and the in situ observed SM from the sparse sites over the LAB-net, SMOSMANIA, and TAHMO demonstrated the better accuracy of the CVAEsm compared with other products. CVAEsm performed better over the LAB-net and TAHMO sites, and the CCs of CVAEsm were 0.925 and 0.726 with the ubRMSE about 0.062 and 0.054, respectively, in the two sparse networks. Statistics of the CVAEsm were close to SMAP L4 and better than CCI products over the SMOSMANIA networks, with a CC of 0.937 and ubRMSE values of 0.024 m^3/m^3. Overall, the averaged CC and ubRMSE of the CVAEsm, CCIsm, and SMAPsm products over sparse network outside the CONUS were 0.863 (0.843, 0.832) and 0.046 (0.054, 0.046), respectively. All aforementioned explored issues indicated that the CVAEsm performed better over the sparse sites outside the CONUS.

B. Core Site Validation

The core validation sites are underrepresented in spatial coverage compared to the sparse networks. Yet, they have the denser sensor networks to provide accurate SM measurements within
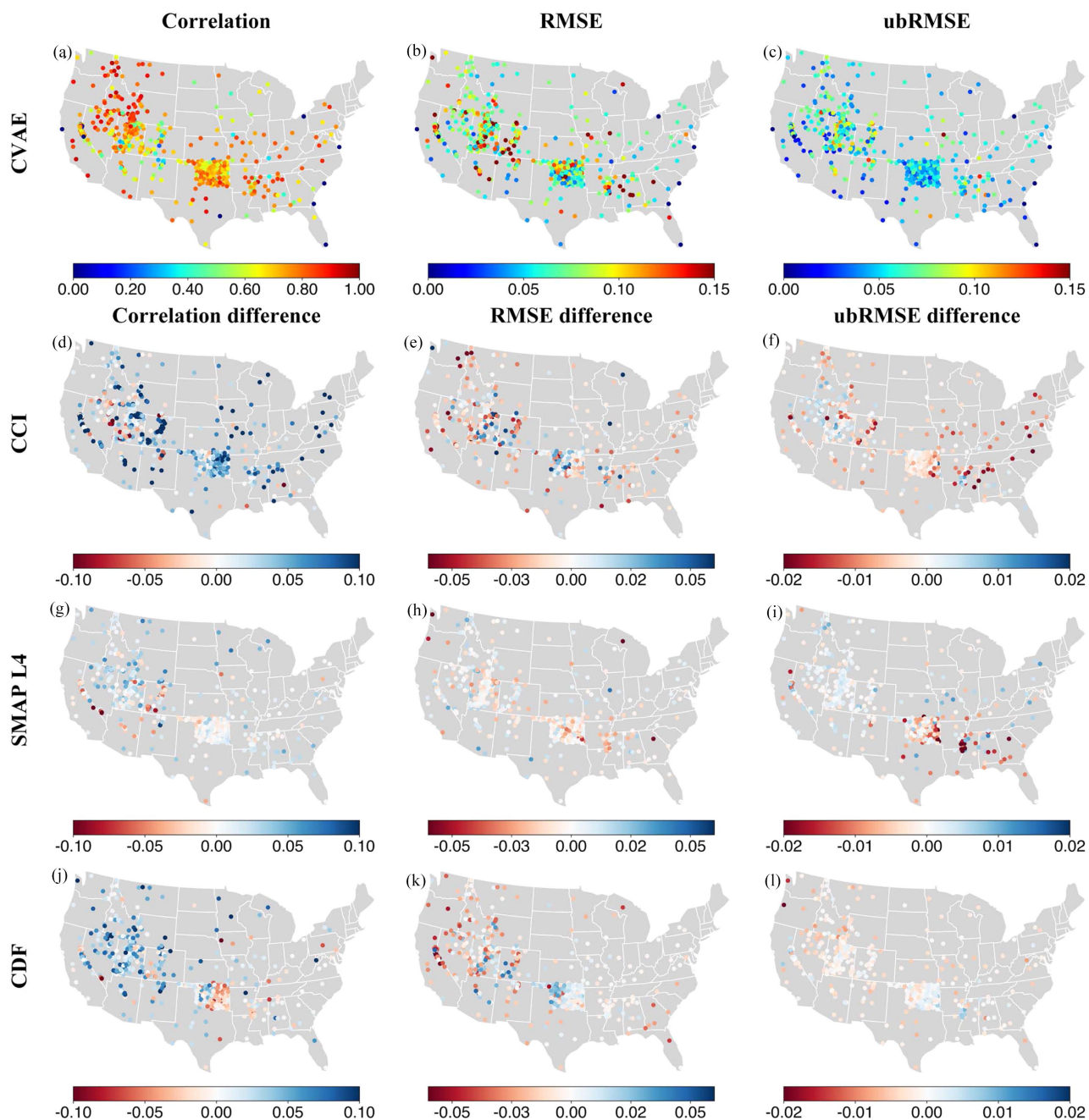
Fig. 8. Spatial distribution of the temporal statistics of the generated product and how they differ from those of the base products. The correlation, RMSE, and ubRMSE for CVAEsm for all sites are displayed in the first row. The subsequent rows display variances in the assessment metrics comparing CVAEsm with the baseline products.

the grid cells of the SM map. The estimated SM was validated against core validation sites for the period from 2016 to 2020 all over the world. Representative core validation sites, including Walnut Gulch watershed sites, Little Washita sites, Fort Cobb sites, Little River sites, Reynolds Creek sites, REMED-HUS sites, Tibetan Plateau Maqu site, Shandian River sites, Yanco sites, and Benin sites, were selected for validations. With their extensive coverage of global geography, land cover types, and soil textures, these ground-based sites serve as the important validation points for satellite SM products. The CVAEsm

over the core validation sites was validated by comparing the estimated SM product with the in situ observations from those sites, and the results were exhibited in the scatter plots in Figs. 11 and 12, a statistical matrix summarized these results was listed in Table V.

As portrayed in Figs. 11 and 12 and listed in Table V, the scatter plots and score metrics derived from ten areas demonstrate the higher accuracy of the CVAEsm product compared to the other SM products. In general, the reconstructed SM by the CVAE showed better accuracy over ten core validation sites, with an
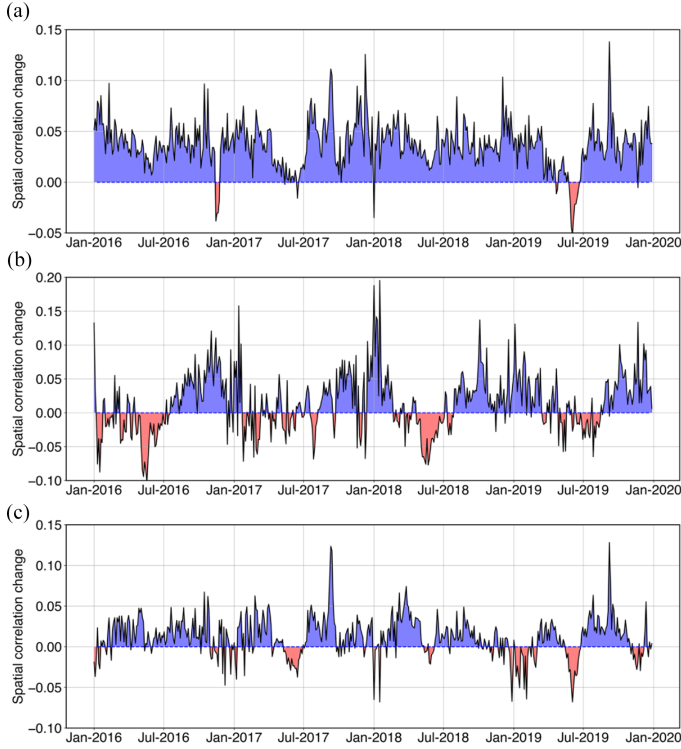
(a)

(b)

(c)

Fig. 9. Increased value of CVAEsm's spatial correlation with the in situ observations for each time step over the CONUS than the SM products (SMAPsm, CCIsm, and CDFsm).

TABLE V
METRICS FOR THE COMPARISONS AMONG THE CVAEsm, CCIsm, AND SMAPsm OVER THE CORE VALIDATION SITES

Core validation sites	CVAE		CCI		SMAP L4	
	CC	ubRMSE	CC	ubRMSE	CC	ubRMSE
Fort Cobb	0.882	0.025	0.868	0.028	0.876	0.028
Walnut Gulch	0.856	0.025	0.815	0.029	0.837	0.027
Little Washita	0.900	0.023	0.882	0.025	0.889	0.024
Reynolds Creek	0.897	0.031	0.883	0.031	0.869	0.033
Little River	0.898	0.019	0.821	0.025	0.885	0.028
REMEDHUS	0.860	0.026	0.869	0.029	0.842	0.033
Shandian River	0.661	0.029	0.566	0.031	0.623	0.033
Maqu	0.755	0.034	0.640	0.039	0.756	0.042
Yanco	0.964	0.031	0.945	0.046	0.968	0.024
Benin	0.958	0.045	0.909	0.057	0.929	0.049
Average	**0.863**	**0.029**	0.820	0.034	0.841	0.032

average CC of 0.863 and an average ubRMSE of 0.029 m³/m³. The CCs of CVAEsm with the in situ observed SM from the most core validation sites ranked the highest. For the REMEDHUS and Maqu core validation sites, although the CCs of CVAEsm did not reach the highest, they were only slightly lower than the best products by 0.009 and 0.001, respectively. And comparisons of three-day composited time series of SM from the CVAEsm, CCI, SMAP L4, CDFsm, and the in situ observed SM from the core validation sites over the world were shown in Fig. 13.

VI. DISCUSSIONS

A. Comparison With CDF Matching-Based SM

The CVAE-based reconstruction of SM against the CDF matching-based ones were validated, both of which use CCIsm

TABLE VI
METRICS FOR THE COMPARISONS BETWEEN CVAEsm AND CDFsm OVER CORE VALIDATION SITES

Core validation sites	CVAE		CDF	
	CC	ubRMSE	CC	ubRMSE
Fort Cobb	0.882	0.025	0.865	0.027
Walnut Gulch	0.856	0.025	0.814	0.029
Little Washita	0.900	0.023	0.884	0.024
Reynolds Creek	0.897	0.031	0.824	0.031
Little River	0.898	0.019	0.876	0.021
REMEDHUS	0.860	0.026	0.844	0.027
Shandian River	0.661	0.029	0.687	0.029
Maqu	0.755	0.034	0.75	0.035
Yanco	0.964	0.031	0.929	0.045
Benin	0.958	0.045	0.913	0.036
Average	**0.863**	**0.029**	0.838	0.031

TABLE VII
METRICS FOR THE COMPARISONS BETWEEN CVAEsm AND CDFsm OVER SPARSE NETWORK

Sparse network	CVAE		CDF	
	CC	ubRMSE	CC	ubRMSE
SCAN	0.781	0.087	0.757	0.093
USCRN	0.678	0.099	0.643	0.103
SNOTEL	0.516	0.103	0.482	0.105
OK Mesonet	0.637	0.071	0.626	0.072
SMOSMANIA	0.937	0.024	0.924	0.024
TAHMO	0.726	0.053	0.634	0.060
LAB-net	0.925	0.062	0.902	0.074
Average	**0.743**	**0.071**	0.710	0.076

and SMAPsm as inputs and do not require additional variables, as well as ground-based measurements as inputs. Comparisons were made between the CVAEsm and the CDFsm products through the core validation sites and sparse networks, and the results were listed in Tables VI and VII. The averaged CC and ubRMSE of the CVAEsm and CDFsm products over the core validation sites are 0.863 (0.838) and 0.029 (0.031), respectively. And the averaged CC and ubRMSE of the CVAEsm and CDFsm products over the sparse networks in the CONUS are 0.743 (0.710) and 0.071 (0.076), respectively.

B. Comparisons With Other SM Products

In addition, we compared the accuracy of CVAEsm with those of four SM products proposed by recent studies, including the following:
1) ANN-based SM data product (NNsm) that transfers the merits of SMAP to AMSR-E/2 with in situ SM observations [60];
2) gap-filled AMSR2 SM data product (SGDsm) based on partial convolutional neural network with in situ SM observations [56];
3) ERA5 SM product (ERA5sm), one of the most commonly used sources of SM in geoscience studies [71];
4) SMAP-HydroBlocks as listed in Tables VIII–XI and Fig. 17.

The comparisons were made for the period from 2016 to 2019. CVAEsm outperforms NNsm, SGDsm, and ERA5sm regarding
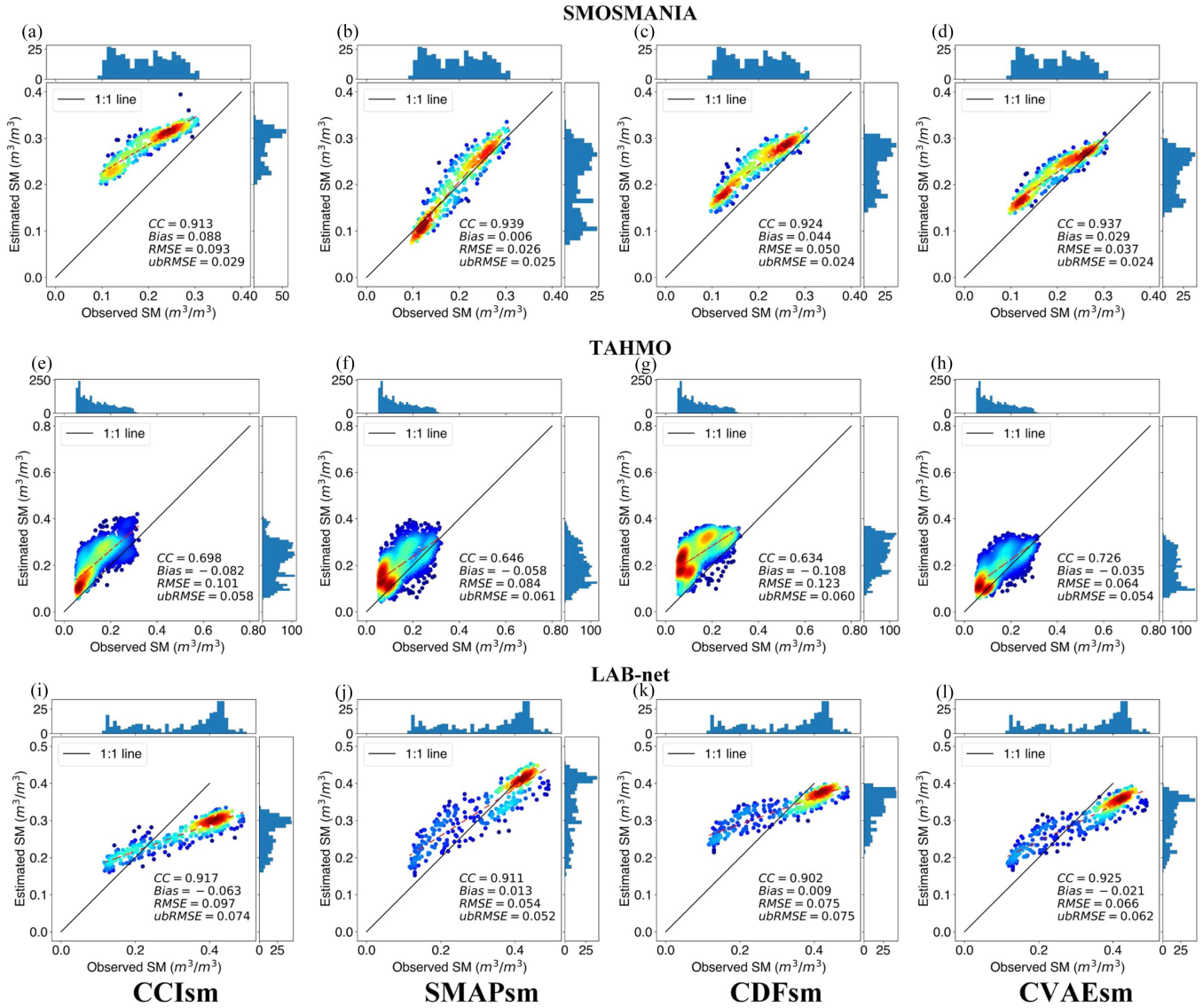
Fig. 10. Scatterplot of CCIsm, SMAPsm, CDFsm, and CVAEsm products versus sparse in situ SM observations outside CONUS.

TABLE VIII
METRICS FOR THE COMPARISONS BETWEEN CVAEsm, NNsm, AND SGDsm OVER THE CORE VALIDATION SITES

Core validation sites	CVAE		NNsm		SGDsm	
	CC	ubRMSE	CC	ubRMSE	CC	ubRMSE
Fort Cobb	0.882	0.025	0.843	0.028	0.766	0.047
Walnut Gulch	0.856	0.025	0.714	0.034	0.686	0.035
Little Washita	0.900	0.023	0.801	0.032	0.688	0.061
REMEDHUS	0.860	0.026	0.851	0.032	0.833	0.083
Shandian River	0.661	0.029	0.436	0.038	0.658	0.041
Maqu	0.755	0.034	0.678	0.053	0.306	0.068
Yanco	0.964	0.031	0.874	0.054	0.815	0.051
Benin	0.958	0.045	0.809	0.072	0.754	0.063
Reynolds Creek	0.897	0.031	0.507	0.053	0.745	0.083
Little River	0.898	0.019	0.805	0.033	0.520	0.074
Average	**0.863**	**0.029**	0.732	0.043	0.677	0.061

TABLE IX
METRICS FOR THE COMPARISONS BETWEEN CVAEsm, NNsm, AND SGDsm OVER THE SPARSE NETWORKS

Sparse network	CVAE		NNsm		SGDsm	
	CC	ubRMSE	CC	ubRMSE	CC	ubRMSE
SCAN	0.781	0.087	0.645	0.109	0.497	0.169
USCRN	0.678	0.099	0.540	0.126	0.434	0.221
SNOTEL	0.516	0.103	0.099	0.143	0.155	0.215
OK Mesonet	0.637	0.071	0.491	0.087	0.391	0.145
SMOSMANIA	0.937	0.024	0.878	0.032	0.807	0.049
TAHMO	0.726	0.053	N/A	N/A	N/A	N/A
LAB-net	0.925	0.062	0.217	0.110	0.241	0.122
Average	**0.743**	**0.071**	0.478	0.101	0.421	0.154

the CC and ubRMSE for both core validation sites and sparse networks. The averaged CC and ubRMSE of the CVAEsm,

NNsm, SGDsm, and ERA5sm products over core validation sites are 0.863 (0.732, 0.677, 0.813) and 0.029 (0.043, 0.061, 0.048), respectively. And the averaged CC and ubRMSE of the CVAEsm, NNsm SGDsm, and ERA5sm products over the
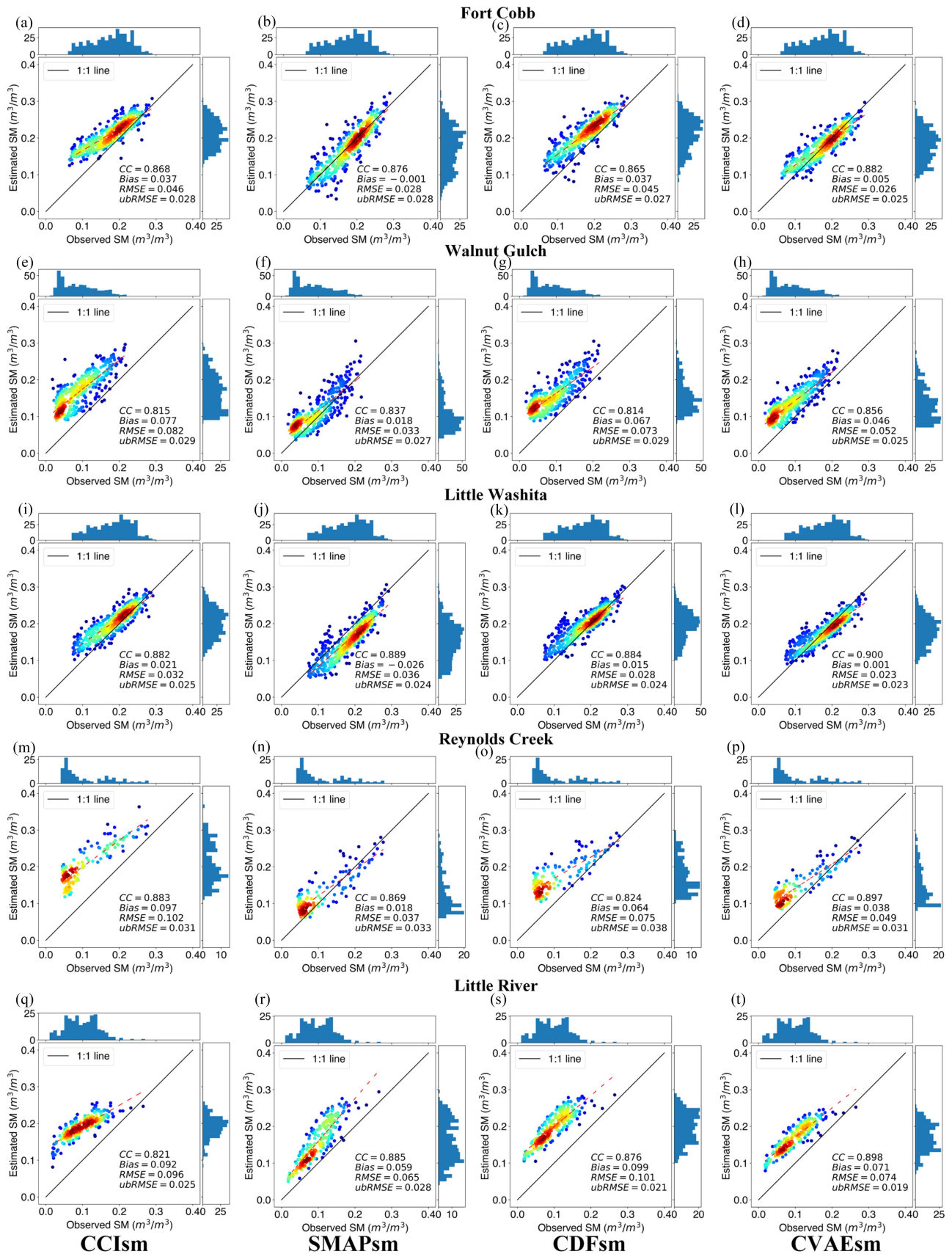
Fig. 11.　Scatterplots of the CCIsm, SMAPsm, CDFsm, and CVAEsm versus in situ SM observations from the core validation sites over the CONUS.
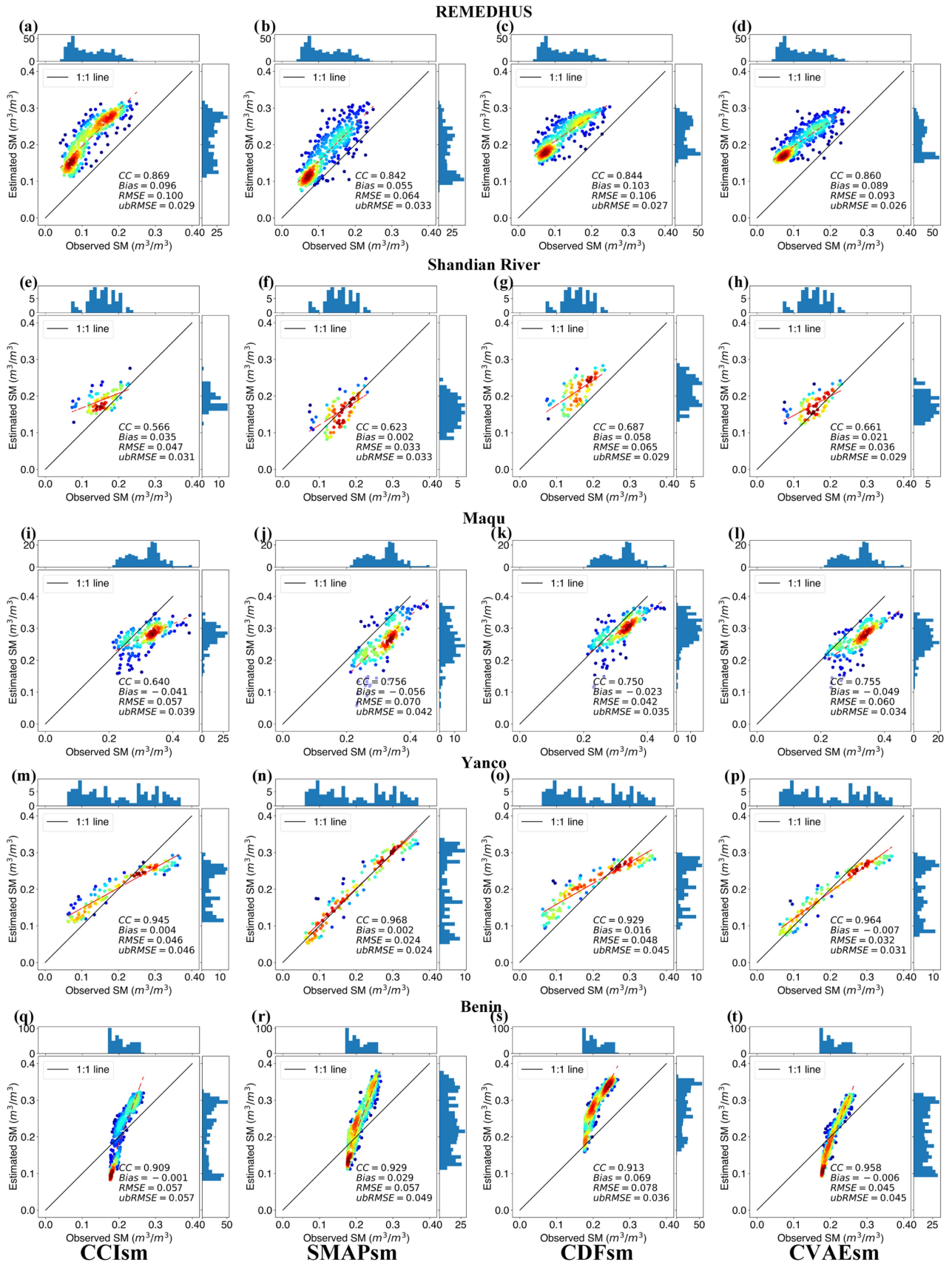
Fig. 12. Scatterplot of the CCIsm, SMAPsm, CDFsm, and CVAEsm versus in situ SM observations from the core validation sites outside CONUS, USA.
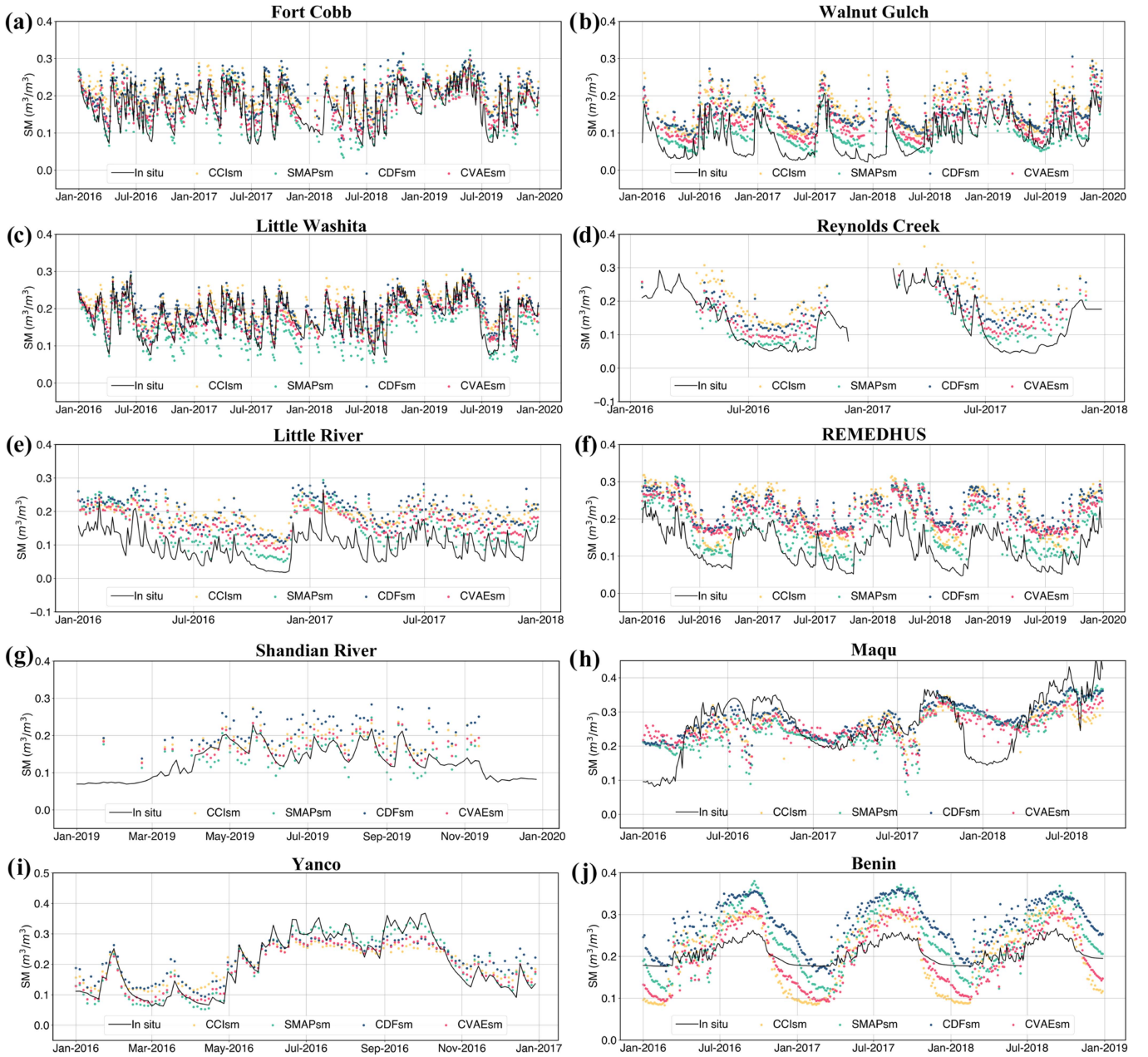
Fig. 13.   Comparisons of three-day composited time series of SM from the CVAEsm, CCI, SMAP L4, CDFsm, and the in situ observed SM from the core validation sites over the world.

sparse network are 0.743 (0.447, 0.380, 0.713) and 0.071 (0.101, 0.154, 0.080), respectively. CVAEsm is statistically superior to the other SM products in both core validation sites and sparse networks and is of more obvious advantages over NNsm and SGDsm. Since SMAP-HB introduces several HR auxiliary variables to participate in the modeling, its resolution reaches about 30 m, which is a large scale difference compared with the reconstructed data by our proposed method. To reduce the influence of spatial scale effects on the validation results, we aggregated the SMAP-HB data to the same spatial scale as CVAEsm and selected the OK mesonet in situ measured data from 2016 to 2019 for accuracy assessment. As shown in Fig. 17, while SMAP-HydroBlocks achieves the highest agreement

(CC) with the reference SM, it also has the highest bias and RMSE. This suggests that SMAP-HydroBlocks might capture the overall distribution of SM well but may have a tendency to underestimate the values in the area. SM enhanced by CVAE has a good balance between agreement, bias, and RMSE. It has a very close CC to SMAP-HydroBlocks, the relatively lower bias and RMSE.

C. Validations at Different Spatial Scales

Accuracy assessment at different spatial scales was conducted using in situ SM observation sites all over the world, CONUS, and Oklahoma state. Due to the uneven spatial distribution of
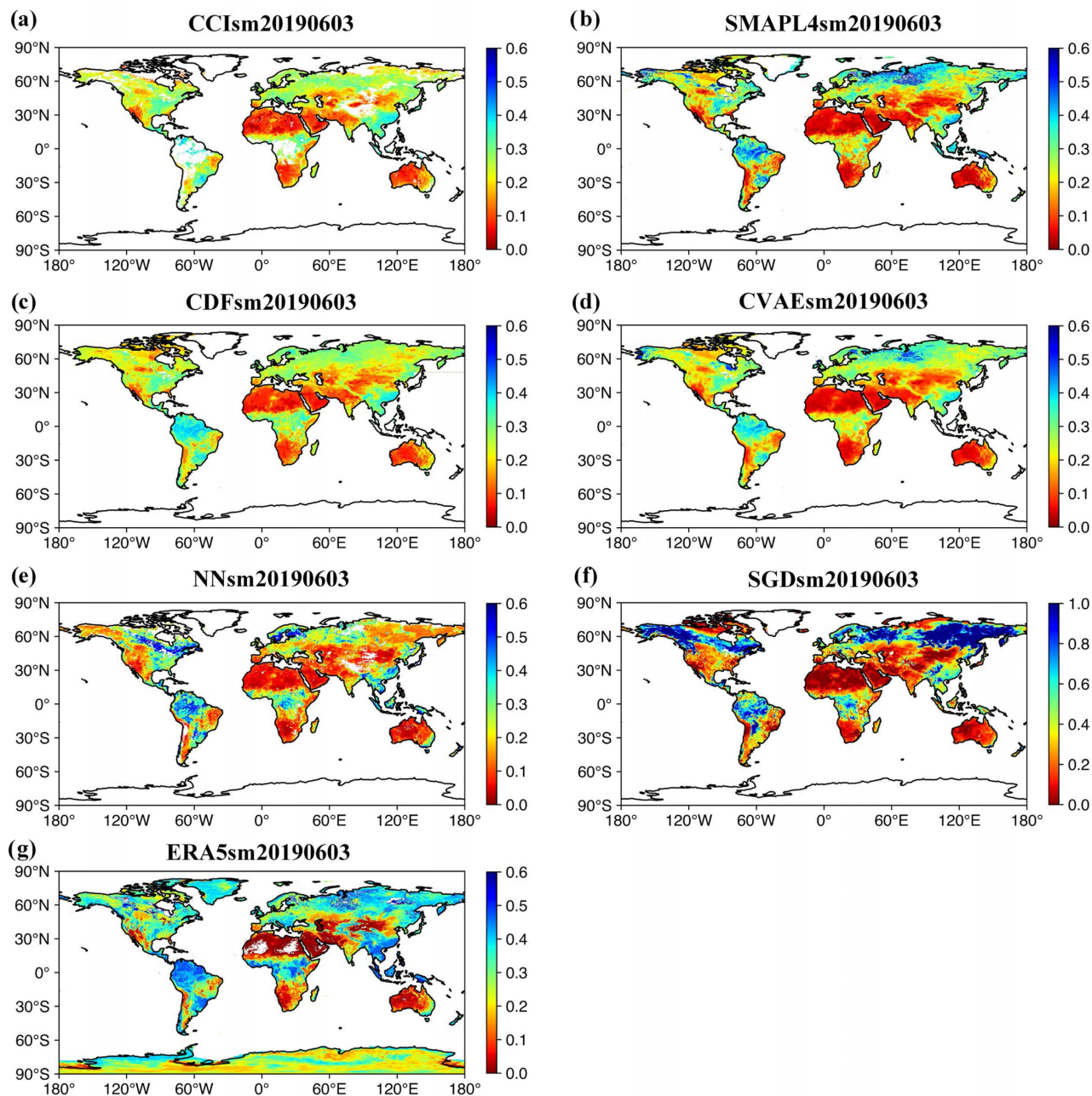
Fig. 14. Global map of CCIsm, SMAPsm, CDFsm, CVAEsm, NNsm, SGDsm, and ERA5sm on June 3, 2019. Since many areas in SGDsm's global map have values greater than 0.6, the color bar range is adjusted to 0–1.

sparse networks around the world, the average value of verification accuracy by core validation sites was used to evaluate the accuracy of SM products at the global scale, while the validations at the CONUS and Oklahoma state were performed over the core validation sites. The assessments listed in Table XII indicated that CVAEsm achieved better accuracy on all three spatial scales compared with CCIsm, SMAPsm, and CDFsm. Overall, the average accuracy of CCIsm in the global and Oklahoma state is lower than SMAPsm and higher than SMAPsm in the continental scale of the United States. And the CDFsm obtained through CDF matching performs worse than CVAEsm, and its precision

at the global scale is lower than SMAPsm, while it is better than CCIsm and SMAPsm in the CONUS and Oklahoma state.

Visual comparisons on the global, CONUS, and OKM SM maps of these products on June 3, 2019 were shown in Figs. 14–16. CVAEsm, CDFsm SMAPsm, and SMAP-HB have richer spatial texture and more continuous variation at spatial scales. In contrast, CCIsm, NNsm, and SGDsm are poor in characterizing regional heterogeneity and have less spatial continuity due to their lower spatial resolution. Although ERA5sm has a close spatial resolution of 0.1° to SMAP L4, ERA5sm is less representative on the spatial heterogeneity of SM.
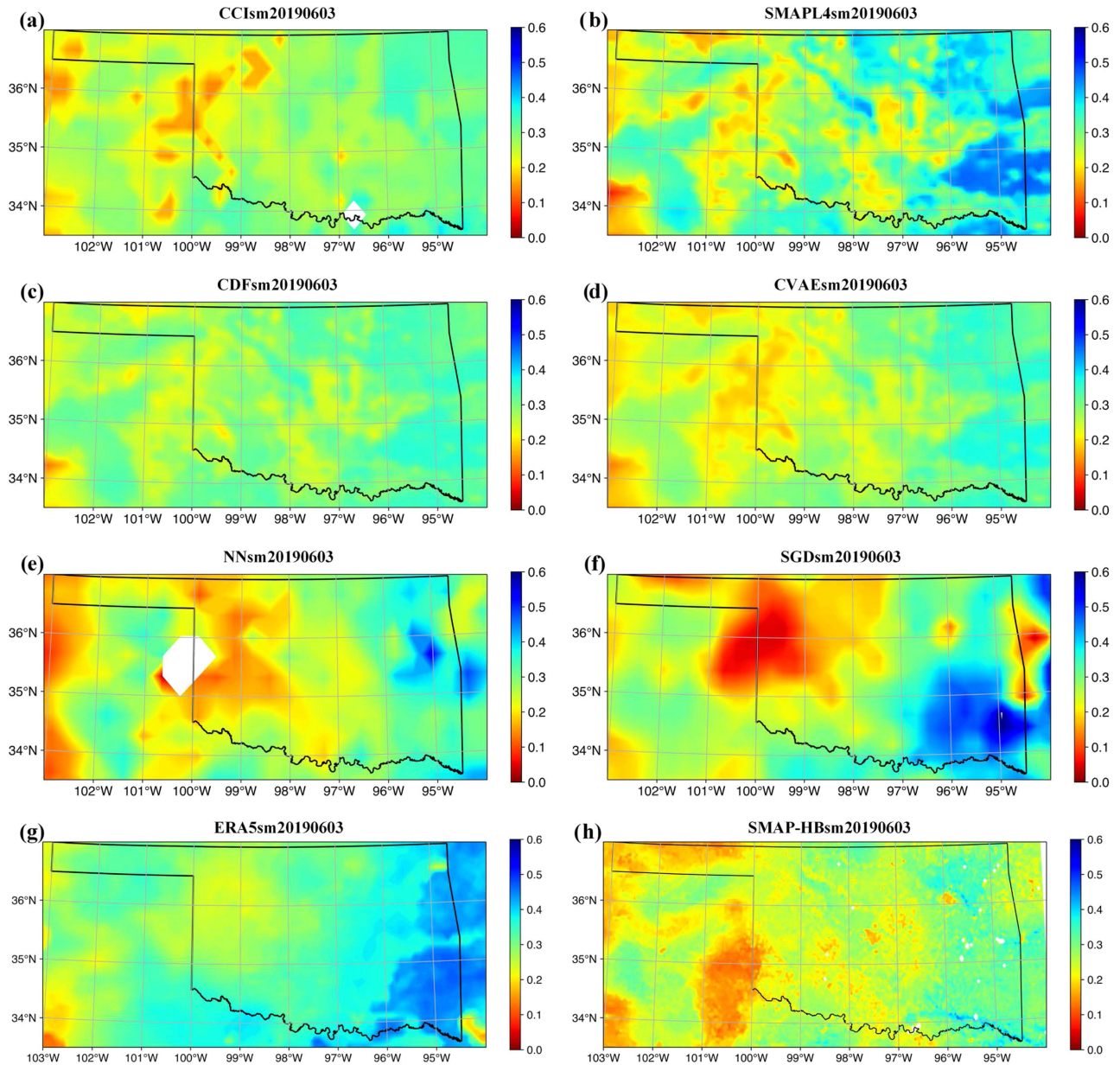
Fig. 15. Map of CCIsm, SMAPsm, CDFsm, CVAEsm, NNsm, SGDsm, ERA5sm, and SMAP-HBsm across Oklahoma, USA on June 3, 2019.

TABLE X
METRICS FOR THE COMPARISONS BETWEEN CVAESM AND ERA5SM OVER
THE CORE VALIDATION SITES

Core	CVAE		ERA5	
validation sites	CC	ubRMSE	CC	ubRMSE
Fort Cobb	0.882	0.025	0.824	0.040
Walnut Gulch	0.856	0.025	0.744	0.049
Little Washita	0.900	0.023	0.829	0.041
REMEDHUS	0.860	0.026	0.933	0.053
Shandian River	0.661	0.029	0.696	0.033
Maqu	0.755	0.034	0.613	0.053
Yanco	0.964	0.031	0.915	0.035
Benin	0.958	0.045	0.877	0.107
Reynolds Creek	0.897	0.031	0.850	0.035
Little River	0.898	0.019	0.850	0.035
Average	**0.863**	**0.029**	0.813	0.048

SMAPsm was used as the reconstructed data source of CVAEsm, and its original spatial resolution is 9 km. To quantitatively express the reconstructed spatial texture quality, we introduced PSNR and SSIM to compare the spatial texture quality of SM products with SMAPsm as the benchmark in the CONUS from 2016 to 2020. Among the four types of the reconstructed products and ERA5sm, according to the PSNR and SSIM calculations, the spatial texture quality of CVAEsm and CDFsm is close to that of SMAP L4. The average PNSR and SSIM of the CVAEsm and the CDFsm products over the CONUS are 79.096 and 0.945, 77.340, and 0.954, correspondingly. The spatial texture quality of SGMsm, NNsm, and ERA5sm is worse than that of SMAPsm, with PNSR and SSIM of 61.978, 67.102, 68.217 and 0.557,0.646, 0.656, respectively.
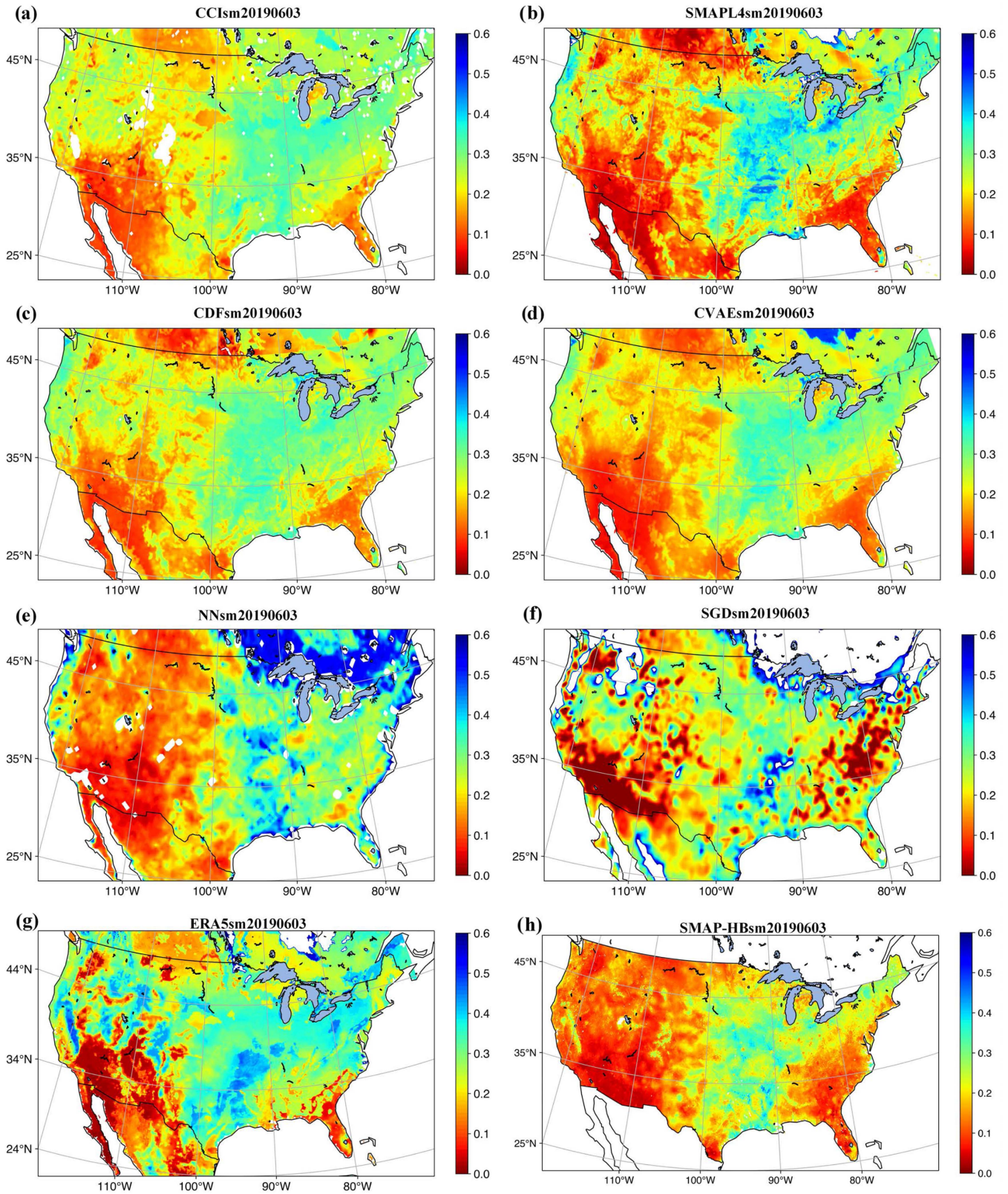
Fig. 16.   Map of CCIsm, SMAPsm, CDFsm, CVAEsm, NNsm, SGDsm, ERA5sm, and SMAP-HBsm across the CONUS on June 3, 2019.

TABLE XI
METRICS FOR THE COMPARISONS BETWEEN CVAEsm AND ERA5sm OVER THE SPARSE NETWORKS

Sparse network	CVAE		ERA5	
	CC	ubRMSE	CC	ubRMSE
SCAN	0.781	0.087	0.663	0.112
USCRN	0.678	0.099	0.702	0.106
SNOTEL	0.516	0.103	0.444	0.117
OK Mesonet	0.637	0.071	0.578	0.084
SMOSMANIA	0.937	0.024	0.950	0.023
TAHMO	0.726	0.053	0.712	0.078
LAB-net	0.925	0.062	0.941	0.042
Average	**0.743**	**0.071**	0.713	0.080

TABLE XII
METRICS FOR THE COMPARISONS BETWEEN CVAEsm, CCIsm, AND SMAPsm AT DIFFERENT SPATIAL SCALE

Scale	CVAE		CCI		SMAP	
	CC	ubRMSE	CC	ubRMSE	CC	ubRMSE
Globe	0.855	0.029	0.811	0.035	0.840	0.033
CONUS	0.678	0.092	0.663	0.094	0.650	0.094
Oklahoma	0.637	0.071	0.600	0.074	0.617	0.077
Average	**0.726**	**0.064**	0.694	0.067	0.705	0.068

TABLE XIII
METRICS FOR THE COMPARISONS BETWEEN CVAEsm AND CVAEsm WITHOUT PROBAILISTIC DISTRIBUTION LEARNING

Probailistic distribution learning	✓	✗
CC	**0.639**	0.598
Bias	0.037	**0.033**
RMSE	**0.081**	0.088
ubRMSE	**0.072**	0.081

TABLE XIV
METRICS FOR THE COMPARISONS BETWEEN CVAEsm WITH DIFFERENT NUMBER OF BASE CHANNELS OF DISTRIBUTION ENCODER AND DISTRIBUTION DECODER

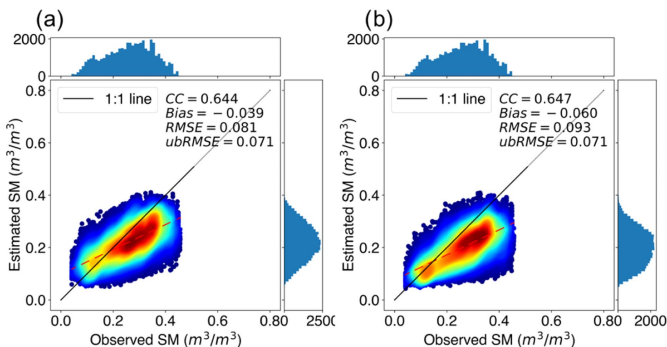| Number of Base channels	32	64	96	128
CC	0.639	**0.642**	0.637	**0.642**
Bias	0.037	**0.034**	0.040	0.038
RMSE	0.081	**0.079**	0.082	0.081
ubRMSE	0.072	**0.071**	**0.071**	**0.071**

Fig. 17.   Scatterplot of CVAEsm and SMAP-HB products versus sparse in situ SM observations in Oklahoma, USA. (a) Enhanced by CVAE. (b) SMAP-HydroBlocks.

D. Parameter Sensitivity Analysis

In order to demonstrate the effectiveness of CVAE distribution learning for our proposed model, we conducted ablation experiments to compare the experimental results of the model after removing the distribution encoder and distribution decoder with the original CVAE model. The result is shown in the table XIII.

Since the original CVAE model has better performance on three out of the four metrics (Bias, RMSE, and CC), the ablation experiments demonstrate that the distribution encoder and decoder play an important role in the effectiveness of the model for probabilistic distribution learning.

We also analyzed the effect of the number of base convolution channels of the input distribution encoder and distribution decoder on the fusion accuracy of the SM map. As shown in the table XIV, while the increase in CC between 32 and 128 channels is not very large (0.003), it suggests that a higher number of channels might allow the model to learn more complex features from the input data, leading to a slightly better agreement with the actual SM map. It is important to note that increasing the number of channels can also lead to overfitting, where the model memorizes the training data rather than learning generalizable patterns. The optimal number of channels would likely need to be determined through experimentation on a specific dataset.

VII. CONCLUSION

In this study, a new SM reconstruction scheme based on the variational autoencoder was proposed to reconstruct a global high quality three-day seamless SM product from 2015 to 2021 by using the CCI and SMAP L4 SM products. The scheme does not rely on ground measurements of SM and other geographic auxiliary variables, takes the statistical distribution characteristics of 16 regions around the world into account, thereby reduces the impact of missing areas, and combines the high spatial resolution of SMAP L4 with the high accuracy of CCI fusion products across a wide range of ground cover types. Validations of the reconstructed SM product in both global sparse networks and core validation sites show that the reconstructed SM product has a higher correlation with ground-based observations and smaller ubRMSE than the baseline product, and better spatial detail than the recently proposed NNsm, SGDsm, and ERA5sm SM products.

APPENDIX

Researchers in hydrology, climatology, ecology, and related fields can access this three-day soil moisture product online,[3] facilitating valuable insights and applications in these domains. The code utilized in this research is accessible online.[4]

---

[3][Online]. Available: https://doi.org/10.5281/zenodo.8000601
[4][Online]. Available: https://github.com/YangTze7/CVAE-SM

CONUS. The authors would also like to thank Dr. P. Yao at the Aerospace Information Research Institute, Chinese Academy of Sciences, who kindly helped to access the NNsm dataset.
Author Contributions: C. Shi and Z. Zhang: Data curation, methodology, software, writing original draft. W. Zhang and Z. Zhang: Conceptualization, methodology, supervision, writing review, and editing. S. Xiong: Data curation, proofreading, and editing.

REFERENCES

[1] C. A. Williams and J. D. Albertson, "Soil moisture controls on canopy-scale water and carbon fluxes in an African savanna," *Water Resour. Res.*, vol. 40, no. 9, Sep. 2004, doi: 10.1029/2004WR003208.

[2] D. B. Clark and N. Gedney, "Representing the effects of subgrid variability of soil moisture on runoff generation in a land surface model," *J. Geophys. Res. Atmos.*, vol. 113, no. D10, May 2008, doi: 10.1029/2007JD008940.

[3] E. E. Small and S. A. Kurc, "Tight coupling between soil moisture and the surface radiation budget in semiarid environments: Implications for land-atmosphere interactions," *Water Resour. Res.*, vol. 39, no. 10, Oct. 2003, doi: 10.1029/2002WR001297.

[4] X. Jia et al., "Carbon and water exchange over a temperate semi-arid shrubland during three years of contrasting precipitation and soil moisture patterns," *Agricultural Forest Meteorol.*, vol. 228-229, pp. 120–129, Nov. 2016.

[5] T. W. Ford, A. D. Rapp, S. M. Quiring, and J. Blake, "Soil moisture–precipitation coupling: Observations from the oklahoma mesonet and underlying physical mechanisms," *Hydrol. Earth Syst. Sci.*, vol. 19, no. 8, pp. 3617–3631, Aug. 2015.

[6] 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. 185–215, Dec. 2017.

[7] V. Masson-Delmotte et al., "IPCC, 2021: Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change," vol. 2, no. 1, pp. 2391, Jun. 2021.

[8] C. B. Baker et al., "Working toward a national coordinated soil moisture monitoring network: Vision, progress, and future directions," *Bull. Amer. Meteorol. Soc.*, vol. 103, no. 12, pp. E2719–E2732, Sep. 2022.

[9] J. P. Walker, G. R. Willgoose, and J. D. Kalma, "In situ measurement of soil moisture: A comparison of techniques," *J. Hydrol.*, vol. 293, no. 1, pp. 85–99, Jul. 2004.

[10] R. D. Koster, Z. Guo, R. Yang, P. A. Dirmeyer, K. Mitchell, and M. J. Puma, "On the nature of soil moisture in land surface models," *J. Climate*, vol. 22, no. 16, pp. 4322–4335, Aug. 2009.

[11] G. Dumedah and J. P. Walker, "Assessment of model behavior and acceptable forcing data uncertainty in the context of land surface soil moisture estimation," *Adv. Water Resour.*, vol. 101, pp. 23–36, Mar. 2017.

[12] A. J. Hartley, N. MacBean, G. Georgievski, and S. Bontemps, "Uncertainty in plant functional type distributions and its impact on land surface models," *Remote Sens. Environ.*, vol. 203, pp. 71–89, Dec. 2017.

[13] L. Prihodko, A. S. Denning, N. P. Hanan, I. Baker, and K. Davis, "Sensitivity, uncertainty and time dependence of parameters in a complex land surface model," *Agricultural Forest Meteorol.*, vol. 148, no. 2, pp. 268–287, Feb. 2008.

[14] E. G. Njoku and J.-A. Kong, "Theory for passive microwave remote sensing of near-surface soil moisture," *J. Geophys. Res.*, vol. 82, no. 20, pp. 3108–3118, Jul. 1977.

[15] J. P. Wigneron, J. C. Calvet, T. Pellarin, A. A. Van de Griend, M. Berger, and P. Ferrazzoli, "Retrieving near-surface soil moisture from microwave radiometric observations: Current status and future plans," *Remote Sens. Environ.*, vol. 85, no. 4, pp. 489–506, Jun. 2003.

[16] L. Brocca et al., "A review of the applications of ASCAT soil moisture products," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 2285–2306, May 2017.

[17] D. Entekhabi et al., "The soil moisture active passive (SMAP) mission," *Proc. IEEE*, vol. 98, no. 5, pp. 704–716, May 2010.

[18] K. Imaoka, M. Kachi, M. Kasahara, N. Ito, K. Nakagawa, and T. Oki, "Instrument performance and calibration of AMSR-E and AMSR2," *Int. Arch. Photogramm. Remote Sens. spatial Inf. Sci.—ISPRS Arch.*, vol. 38, no. 8, pp. 13–18, 2010.

[19] Y. H. Kerr et al., "The SMOS soil moisture retrieval algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1384–1403, May 2012.

[20] K. C. Kornelsen and P. Coulibaly, "Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications," *J. Hydrol.*, vol. 476, pp. 460–489, Jan. 2013.

[21] E. G. Njoku, J. M. Stacey, and F. T. Barath, "The seasat scanning multichannel microwave radiometer (SMMR): Instrument description and performance," *IEEE J. Ocean. Eng.*, vol. 5, no. 2, pp. 100–115, Apr. 1980.

[22] S. Paloscia, G. Macelloni, E. Santi, and T. Koike, "A multifrequency algorithm for the retrieval of soil moisture on a large scale using microwave data from SMMR and SSM/I satellites," *IEEE Trans. Geosci. Remote Sens*, vol. 39, no. 8, pp. 1655–1661, Aug. 2001.

[23] S. Paloscia, S. Pettinato, E. Santi, C. Notarnicola, L. Pasolli, and A. Reppucci, "Soil moisture mapping using sentinel-1 images: Algorithm and preliminary validation," *Remote Sens. Environ.*, vol. 134, pp. 234–248, Jul. 2013.

[24] 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, Jan. 2015.

[25] J. Shi et al., "Physically based estimation of bare-surface soil moisture with the passive radiometers," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3145–3153, Nov. 2006.

[26] W. Wagner, G. Lemoine, and H. Rott, "A method for estimating soil moisture from ERS scatterometer and soil data," *Remote Sens. Environ.*, vol. 70, no. 2, pp. 191–207, Nov. 1999.

[27] L. Karthikeyan, M. Pan, N. Wanders, D. N. Kumar, and E. F. Wood, "Four decades of microwave satellite soil moisture observations: Part 1. a review of retrieval algorithms," *Adv. Water Resour.*, vol. 109, pp. 106–120, 2017.

[28] J. Peng, A. Loew, O. Merlin, and N. E. Verhoest, "A review of spatial downscaling of satellite remotely sensed soil moisture," *Rev. Geophys.*, vol. 55, no. 2, pp. 341–366, 2017.

[29] W. A. Dorigo et al., "Error characterisation of global active and passive microwave soil moisture datasets," *Hydrol. Earth Syst. Sci.*, vol. 14, no. 12, pp. 2605–2616, Dec. 2010.

[30] L. Karthikeyan, M. Pan, N. Wanders, D. N. Kumar, and E. F. Wood, "Four decades of microwave satellite soil moisture observations: Part 2. product validation and inter-satellite comparisons," *Adv. Water Resour.*, vol. 109, pp. 236–252, Nov. 2017.

[31] L. Zhuo and D. Han, "Multi-source hydrological soil moisture state estimation using data fusion optimisation," *Hydrol. Earth Syst. Sci.*, vol. 21, no. 7, pp. 3267–3285, Jul. 2017.

[32] X. Zhan et al., "Fusing microwave and optical satellite observations for high resolution soil moisture data products," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2017, pp. 2519–2522.

[33] D. Hong et al., "Cross-city matters: A multimodal remote sensing benchmark dataset for cross-city semantic segmentation using high-resolution domain adaptation networks," *Remote Sens. Environ.*, vol. 299, 2023, Art. no. 113856.

[34] A. Gruber, W. A. Dorigo, W. Crow, and W. Wagner, "Triple collocation-based merging of satellite soil moisture retrievals," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 6780–6792, Dec. 2017.

[35] A. Gruber, T. Scanlon, R. van der Schalie, W. Wagner, and W. Dorigo, "Evolution of the ESA CCI soil moisture climate data records and their underlying merging methodology," *Earth Syst. Sci. Data*, vol. 11, no. 2, pp. 717–739, May 2019.

[36] R. H. Reichle, R. D. Koster, J. Dong, and A. A. Berg, "Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation," *J. Hydrometeorol.*, vol. 5, no. 3, pp. 430–442, Jun. 2004.

[37] K. E. Mitchell, "The multi-institution north american land data assimilation system (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system," *J. Geophys. Res.*, vol. 109, Apr. 2004, Art. no. D07S90.

[38] Y. Xia et al., "Evaluation of multi-model simulated soil moisture in NLDAS-2," *J. Hydrol.*, vol. 512, pp. 107–125, May 2014.

[39] C. Shi, Z. Xie, H. Qian, M. Liang, and X. Yang, "China land soil moisture EnKF data assimilation based on satellite remote sensing data," *Sci. China Earth Sci.*, vol. 54, no. 9, pp. 1430–1440, Sep. 2011.

[40] J. Blyverket, P. Hamer, L. Bertino, C. Albergel, D. Fairbairn, and W. Lahoz, "An evaluation of the EnKF vs. EnOI and the assimilation of SMAP, SMOS and ESA CCI soil moisture data over the contiguous US," *Remote Sens.*, vol. 11, no. 5, Feb. 2019, Art. no. 478.

[41] W. T. Crow, E. F. Wood, and M. Pan, "Multiobjective calibration of land surface model evapotranspiration predictions using streamflow observations and spaceborne surface radiometric temperature retrievals," *J. Geophys. Res. Atmos.*, vol. 108, Dec. 2003, doi: 10.1029/2002JD003292.

[42] L. Almendra-Martín, J. Martínez-Fernández, M. Piles, and A. González-Zamora, "Comparison of gap-filling techniques applied to the CCI soil moisture database in southern Europe," *Remote Sens. Environ.*, vol. 258, Jul. 2021, Art. no. 112377.

[43] D. Long et al., "Generation of spatially complete and daily continuous surface soil moisture of high spatial resolution," *Remote Sens. Environ.*, vol. 233, Nov. 2019, Art. no. 111364.

[44] L. Gao et al., "A deep neural network based SMAP soil moisture product," *Remote Sens. Environ.*, vol. 277, Aug. 2022, Art. no. 113059.

[45] Q. Yuan, H. Xu, T. Li, H. Shen, and L. Zhang, "Estimating surface soil moisture from satellite observations using a generalized regression neural network trained on sparse ground-based measurements in the continental U.S.," *J. Hydrol.*, vol. 580, Jan. 2020, Art. no. 124351.

[46] M. ElSaadani, E. Habib, A. M. Abdelhameed, and M. Bayoumi, "Assessment of a spatiotemporal deep learning approach for soil moisture prediction and filling the gaps in between soil moisture observations," *Front. Artif. Intell.*, vol. 4, Mar. 2021, Art. no. 636234.

[47] D. Hong et al., "SpectralGPT: Spectral remote sensing foundation model," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Apr. 2024, doi: 10.1109/TPAMI.2024.3362475.

[48] W. Jing, P. Zhang, and X. Zhao, "Reconstructing monthly ECV global soil moisture with an improved spatial resolution," *Water Resour. Manage.*, vol. 32, no. 7, pp. 2523–2537, May 2018.

[49] Q. Li et al., "A 1 km daily soil moisture dataset over China using in situ measurement and machine learning," *Earth Syst. Sci. Data*, vol. 14, no. 12, pp. 5267–5286, Nov. 2022.

[50] Y. Qu et al., "Rebuilding a microwave soil moisture product using random forest adopting AMSR-E/AMSR2 brightness temperature and SMAP over the Qinghai–Tibet Plateau, China," *Remote Sens.*, vol. 11, no. 6, Mar. 2019, Art. no. 683.

[51] H. Yang, Q. Wang, W. Zhao, X. Tong, and P. M. Atkinson, "Reconstruction of a global 9 km, 8-day SMAP surface soil moisture dataset during 2015–2020 by spatiotemporal fusion," *J. Remote Sens.*, vol. 2022, pp. 1–23, Jul. 2022.

[52] M. Bordoni, M. Bittelli, R. Valentino, S. Chersich, M. G. Persichillo, and C. Meisina, "Soil water content estimated by support vector machine for the assessment of shallow landslides triggering: The role of antecedent meteorological conditions," *Environ. Model. Assess.*, vol. 23, no. 4, pp. 333–352, Aug. 2018.

[53] A. Hachani, M. Ouessar, S. Paloscia, E. Santi, and S. Pettinato, "Soil moisture retrieval from Sentinel-1 acquisitions in an arid environment in Tunisia: Application of artificial neural networks techniques," *Int. J. Remote Sens.*, vol. 40, no. 24, pp. 9159–9180, Dec. 2019.

[54] C. s. Lee, E. Sohn, J. D. Park, and J.-D. Jang, "Estimation of soil moisture using deep learning based on satellite data: A case study of South Korea," *GISci. Remote Sens.*, vol. 56, no. 1, pp. 43–67, Jan. 2019.

[55] L. Zhang et al., "Reconstruction of ESA CCI satellite-derived soil moisture using an artificial neural network technology," *Sci. Total Environ.*, vol. 782, Aug. 2021, Art. no. 146602.

[56] Q. Zhang, Q. Yuan, J. Li, Y. Wang, F. Sun, and L. Zhang, "Generating seamless global daily AMSR2 soil moisture (SGD-SM) long-term products for the years 2013–2019," *Earth Syst. Sci. Data*, vol. 13, no. 3, pp. 1385–1401, Mar. 2021.

[57] K. Fang, M. Pan, and C. Shen, "The value of SMAP for long-term soil moisture estimation with the help of deep learning," *IEEE Trans. Geosci. Remote Sens*, vol. 57, no. 4, pp. 2221–2233, Apr. 2019.

[58] O. Sungmin and R. Orth, "Global soil moisture data derived through machine learning trained with in-situ measurements," *Sci. Data*, vol. 8, no. 1, Jul. 2021, Art. no. 170.

[59] 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, Jan. 2019.

[60] P. Yao et al., "A long term global daily soil moisture dataset derived from AMSR-E and AMSR2 (2002–2019)," *Sci. Data*, vol. 8, no. 1, Dec. 2021, Art. no. 143.

[61] D. Hong, J. Yao, C. Li, D. Meng, N. Yokoya, and J. Chanussot, "Decoupled-and-coupled networks: Self-supervised hyperspectral image super-resolution with subpixel fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, Oct. 2023, Art. no. 5527812.

[62] M. H. Cosh, T. J. Jackson, R. Bindlish, and J. H. Prueger, "Watershed scale temporal and spatial stability of soil moisture and its role in validating satellite estimates," *Remote Sens. Environ.*, vol. 92, no. 4, pp. 427–435, Sep. 2004.

[63] F. Lei et al., "Assessment of the impact of spatial heterogeneity on microwave satellite soil moisture periodic error," *Remote Sens. Environ.*, vol. 205, pp. 85–99, Feb. 2018.

[64] R. H. Reichle et al., "Assessment of the SMAP Level-4 surface and root-zone soil moisture product using in situ measurements," *J. Hydrometeorol.*, vol. 18, no. 10, pp. 2621–2645, Oct. 2017.

[65] K. Sohn, H. Lee, and X. Yan, "Learning structured output representation using deep conditional generative models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015, pp. 3483–3491.

[66] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," Apr. 2015, *arXiv:1409.1556 [cs]*.

[67] M. Hu, C. Wu, L. Zhang, and B. Du, "Hyperspectral anomaly change detection based on autoencoder," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3750–3762, 2021.

[68] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[69] A. Gruber et al., "Validation practices for satellite soil moisture retrievals: What are (the) errors?," *Remote Sens. Environ.*, vol. 244, Jul. 2020, Art. no. 111806.

[70] S. Xu and J. Cheng, "A new land surface temperature fusion strategy based on cumulative distribution function matching and multiresolution Kalman filtering," *Remote Sens. Environ.*, vol. 254, Mar. 2021, Art. no. 112256.

[71] J. Muñoz-Sabater et al., "ERA5-land: A state-of-the-art global reanalysis dataset for land applications," *Earth Syst. Sci. Data*, vol. 13, no. 9, pp. 4349–4383, Sep. 2021.

Changjiang Shi (Graduate Student Member, IEEE) received the B.Eng. degree in remote sensing science and technology from the Shandong University of Science and Technology, Qingdao, China, in 2020. He is currently working toward the Ph.D. degree in cartography and geographical information system with the Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China.

His research interests include intelligent understanding of remote sensing images and data assimilation.

Zhijie Zhang received the B.S. degree in electrical engineering from the Nanjing University of Post and Telecommunications, Nanjing, China, in 2014, the M.S. degree in computer engineering from Rutgers University, New Brunswick, NJ, USA, in 2018, and the Ph.D. degree in geography from the University of Connecticut, Storrs, CT, USA, in 2022.

From 2016 to 2017, he was a Research Student with WINLAB, Rutgers, NJ, USA. He is involved in two Internet of Things sensor network projects. From 2018 to 2022, he was a Research Assistant with the Geography Department, University of Connecticut. In 2022, he was a Postdoctoral Research Associate with the School of Geography, Development and Environment, University of Arizona. He is currently an Assistant Professor with the Department of Environment and Society, Quinney College of Natural Resources, Utah State University, UT, USA. His research interest includes applying deep learning methods in remote sensing, hydrology studies, and landslide studies.

Shengqing Xiong received the B.S. degree in hydrogeology and engineering geology from the Chengdu College of Geology (now Chengdu University of Technology), Chengdu, China, in 1983, and the Ph.D. degree in applied geophysics from the China University of Geosciences, Beijing, China, in 1997.

He holds several esteemed positions within academia and professional societies. He is a Doctoral Supervisor with the School of Geological Engineering and Surveying and Mapping, Chang'an University, Xi'an, China, and a special Professor and Doctoral Supervisor with the China University of Geosciences (Beijing) and Chengdu University of Science and Technology. He has authored 18 monographs, authored more than 100 papers, and held five patents. His research interests include theoretical and applied studies of airborne geophysical survey methods and techniques.

Dr. Xiong serves as a Member of the Expert Advisory Committee of the Ministry of Land and Resources, as well as a Member of the Higher Education Teaching Steering Committee of the Ministry of Education. He is the Vice-President of the Geophysics Society of China, a Member of the Council of the China Society of Geosciences, and the Editor-in-Chief of "*Physical Prospecting and Chemistry*." He has been recognized for his contributions, having received one special prize and two second prizes of the National Science and Technology Progress Award.

Wanchang Zhang (Member, IEEE) received the first M.S. degree in geography from The Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences (CAS), Lanzhou, China, in 1992, the second M.S. degree in hydrospheric-atmospheric science through the IHAS Special Program sponsored by UNESCO, Nagoya University, Nagoya, Japan, in 1996, and the Ph.D. degree in Earth system science from the Department of Earth and Planetary Sciences, Graduate School of Science, Nagoya University, in 2000.

He holds a diverse educational background and extensive experience in the field of Earth system science. After completing his doctoral degree, he become a Full Professor and Ph.D. Supervisor with the International Institute of Earth System Science, Nanjing University. In 2005, he was promoted under the prestigious "100 Talent Researcher Program" sponsored by the CAS in 2005, focusing on his research on global changes with the Institute for Atmospheric Physics, CAS. Since 2012, he has been with the Aerospace Information Research Institute, CAS. He specializes specifically in the integration of remote sensing (RS) and geographic information systems (GIS) technologies to develop distributed ecohydro-climatic models for hydrology, water resources, and environmental studies. In addition, his expertise extends to global disaster and environmental studies, utilizing RS and GIS techniques in conjunction with modeling schemes. His research interests include RS and GIS applications in Earth system sciences.