A Physics-Based Algorithm to Couple CYGNSS Surface Reflectivity and SMAP Brightness Temperature Estimates for Accurate Soil Moisture Retrieval

Ting Yang, Wei Wan, Jundong Wang, Baojian Liu, and Zhigang Sun

***Abstract*— Remotely sensed soil moisture (SM) with high accuracy and high spatial–temporal resolution is crucial to meteorological, agricultural, hydrological, and environmental applications. The Cyclone Global Navigation Satellite System (CYGNSS) is the first constellation that uses the L-band signal transmitted by the GNSS satellites to develop daily SM data products. In this study, a physics-based algorithm is proposed to couple CYGNSS surface reflectivity (SR) and Soil Moisture Active Passive (SMAP) brightness temperature estimates for accurate SM retrieval. The algorithm is based on the radiative transfer model and the SMAP data to derive a combined parameter of the vegetation optical depth (τ) and the surface roughness parameter (h). The CYGNSS L1 Version 2.1 data of the years 2017–2018 and 2019–2020 are used for calibration and validation, respectively. The SM estimates agree and correlate well with the SMAP SM and *in situ* SM data on a global scale ($R = 0.679$, RMSE $= 0.051$ m^{3}m^{-3}, and MAE $= 0.045$ m^{3}m^{-3} against SMAP SM; $R = 0.729$ against *in situ* SM). The proposed algorithm makes contributions from two aspects. First, the proposed algorithm provides a physics-based algorithm using SMAP brightness temperature to calibrate the attenuation due to vegetation and surface roughness on the CYGNSS-derived SR. Unlike attenuation models that have been explored previously in the context of CYGNSS, this algorithm executes the calibration without relying on observations of h or vegetation biophysical parameters as inputs but with the SMAP brightness temperature as the only observations. Second, the proposed algorithm provides a new way for the combined usage of CYGNSS and SMAP to improve the temporal and spatial coverages of global SM with temporal coverage increased by 38.2% and spatial coverage increased by 31.6%.**

***Index Terms*— Cyclone Global Navigation Satellite System (CYGNSS), soil moisture (SM), Soil Moisture Active Passive (SMAP), surface roughness, vegetation.**

Manuscript received August 5, 2021; revised October 15, 2021, November 19, 2021, February 6, 2022, and February 21, 2022; accepted February 21, 2022. Date of publication March 10, 2022; date of current version April 12, 2022. This work was supported in part by the Strategic Priority Research Program of the Chinese Academy of Sciences under Grant XDA23050102 and Grant XDA19040303; in part by the Key Project of the Chinese Academy of Sciences under Grant KJZD-SW-113; in part by the National Natural Science Foundation of China Projects under Grant 41971377, Grant 42101376, and Grant 41501360; and in part by the Fengyun Application Pioneering Project under Grant FY-APP-2021.0301. *(Corresponding authors: Wei Wan; Zhigang Sun.)*

Ting Yang is with the CAS Engineering Laboratory for Yellow River Delta Modern Agriculture, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China, and also with the Shandong Dongying Institute of Geographic Sciences, Dongying 257000, China (e-mail: yangt@igsnrr.ac.cn).

Wei Wan and Baojian Liu are with the School of Earth and Space Sciences, Institute of Remote Sensing and GIS, Peking University, Beijing 100871, China (e-mail: w.wan@pku.edu.cn; liubaojian@pku.edu.cn).

Jundong Wang is with the CAS Engineering Laboratory for Yellow River Delta Modern Agriculture, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China (e-mail: wangjd.20b@igsnrr.ac.cn).

Zhigang Sun is with the CAS Engineering Laboratory for Yellow River Delta Modern Agriculture, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China, also with the Shandong Dongying Institute of Geographic Sciences, Dongying 257000, China, also with the Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China, and also with the College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: sun.zhigang@igsnrr.ac.cn).

Digital Object Identifier 10.1109/TGRS.2022.3156959

I. Introduction

SOIL moisture (SM) is a critical component in the hydrological processes, land surface evapotranspiration, water migration, the carbon cycle, and so on [1], [2]. Remotely sensed SM products with high accuracy and high spatial–temporal resolution are essential inputs for such applications. Global SM retrieval has been the research subject of both optical remote sensing and microwave remote sensing fields. The optical remote sensing data have a high spatial resolution (i.e., MODIS with a spatial resolution of hundreds of meters) but are restricted greatly by cloud and mist. In addition, optical remote sensing only develops models for estimating SM and does not have official SM products [3], [4]. On the contrary, active/passive microwave sensors provide the all-sky capability for SM remote sensing. In recent years, numerous satellite missions have been developed, e.g., the Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E), the Advanced Scatterometer (ASCAT), the Soil Moisture and Ocean Salinity (SMOS), and the Soil Moisture Active Passive (SMAP) [5]–[7]. However, due to the design of satellite orbits, gaps exist in daily SM products provided by microwave sensors.

Global Navigation Satellite System Reflectometry (GNSS-R) is a technique that exploits the capability of GNSS satellites to act as a bistatic radar [8]. The satellites transmit L-band microwave signals, the same as the SM missions, such as SMAP and SMOS [9]. Despite drawbacks such as being vulnerable to radio frequency interference (RFI) [10], the L-band protected spectrum (1400–1427 MHz)

1558-0644 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

is optimal for near-surface (0–5 cm) SM remote sensing due to the increased ability to penetrate vegetation relative to shorter wavelengths, such as X- and C-bands. The Cyclone Global Navigation Satellite System (CYGNSS) mission, launched into space in December 2016, includes eight microsatellites with the global positioning system (GPS)-R payload, which can offer high spatial resolutions with low revisit times in measuring global ocean/land geophysical parameters. Oceanography-related CYGNSS studies mainly concentrate on characterizing wind speed, ocean altimetry, and sea surface salinity [11]–[14]. Despite the fact that CYGNSS is not dedicated for land sensing, the science community has further explored its sensitivity to land surface properties, e.g., SM, soil freeze–thaw, flood inundation, wetland, waterbody, and biomass [15]–[28]. The University Corporation for Atmospheric Research (UCAR) first developed and published the CYGNSS SM data product using surface reflectivity (SR) and its correlation with SMAP SM [26]. The UCAR SM product, hereafter named the UCAR-SM, is treated as a benchmark for the SM data developed in this study.

The physical basis of using CYGNSS for SM monitoring is that the L-band is highly sensitive to the changes in soil complex dielectric constant, which is mainly with respect to the presence of SM [17]. Meanwhile, the confounding factors of vegetation [i.e., vegetation optical depth (VOD)] and surface roughness would reduce the sensitivity of the L-band to SM. Therefore, it is a common key issue to eliminate the effects of these factors during the SM estimation process. Some recent studies have proposed different solutions to solve this problem. Eroglu *et al.* [24] built an artificial neural network (ANN) model to retrieve CYGNSS daily SM. The input features of the model included CYGNSS data and other ancillary data, such as vegetation water content (VWC), terrain elevation, and surface roughness [24]. Clarizia *et al.* [25] proposed a reflectivity–vegetation–roughness algorithm to obtain daily SM derived from CYGNSS reflectivity along with SMAP VOD and roughness coefficient. Based on the study of Clarizia *et al.* [25], Pan *et al.* [28] proposed a three-layer model of air, vegetation cover, and soil to estimate SM with the only ancillary data of VOD. However, it should be noted that these methods rely on either the direct knowledge of VOD (i.e., τ in this study) or the roughness coefficient (i.e., h in this study) derived from SMAP data to estimate the effect of vegetation and surface roughness, which makes them have limitations such as: 1) the SMAP-τ is estimated from MODIS NDVI-derived VWC, while, in fact, τ depends on not only the VWC but also the vegetation volume density [29] and 2) the SMAP-h is based on a lookup table close to the land cover classes, which makes it a fixed value lacking in temporal variability [30].

Unlike those attenuation models that have been explored in the aforementioned studies in the context of CYGNSS, this study proposes a physics-based algorithm to calibrate the CYGNSS-derived SR using a zeroth-order approximation of the radiative transfer model [30]. The algorithm executes the calibration without relying on observations of surface roughness parameters or vegetation biophysical parameters as inputs but with the SMAP brightness temperature as the only observations. Here, we assume that the reflected signal is always coherent. Four-year on-orbit CYGNSS data are used for calibration and validation (2017–2018 for calibration and 2019–2020 for validation). The resulting SM is validated by the SMAP SM products and the *in situ* measurements from the International Soil Moisture Network (ISMN). The improvement of the temporal and spatial coverages of global SM after involving CYGNSS into SMAP is also shown. Several issues related to the algorithm are also discussed, including the contributions of the surface roughness scale, the land type on the modeling accuracy, and the advantages and limitations of the algorithm.

II. Data

A. CYGNSS Data

The CYGNSS constellation observes the land characterizations between 38° N and 38° S. The delay-Doppler map (DDM) instruments of CYGNSS are designed to map the signal scattered from the ocean and land surfaces. A DDM is defined as the power of the correlator output as a function of the applied delay and Doppler offsets [17]. The signal-to-noise ratio (SNR) is an observable derived from DDM, which is used to estimate the SR and the resulting SM at the specular point based on the bistatic radar equation [24], [27]. The spatial resolution of CYGNSS over land is an open question at present, which depends on the roughness of the surface at and near the reflection point. For the retrievals dominated by coherent reflections, the spatial footprint can be defined by the along-track first Fresnel zone, which is $\sim 6 \times 0.5$ km before July 2019 and $\sim 3 \times 0.5$ km after July 2019, depending on the integration time of the signal of either 1 or 0.5 s, respectively [31].

Nearly four years (January 2017–November 2020) of the CYGNSS L1 Version 2.1 data are used in this study, downloaded via https://podaac.jpl.nasa.gov/. The data from 2017 to 2018 and 2019 to 2020 are used for calibration and validation, respectively. Two criteria are used to filter the CYGNSS data: 1) the antenna gain greater than 0 dB (corresponding to uncertainties reported in the measured antenna gain patterns) and 2) the quality flags [i.e., S-band transmitter powered up, spacecraft attitude error, blackbody DDM, which is a test pattern, direct signal in DDM, and low confidence in the GPS equivalent isotopically radiated power (EIRP)] in the CYGNSS L1b data are used to select the good data acquisition. All data from eight different acquisition times in one day are used for each pixel. The median value of the revisit times is 2.8 h, and the mean revisit time is 7.2 h. For each grid, the CYGNSS observations are completely sampled to 36-km Equal-Area Scalable Earth (EASE) Grid 2.0 cell used by SMAP; then, a simple averaging is used to resample CYGNSS observations to match with the SMAP retrieval.

B. SMAP Data

SMAP incorporated a radar and a radiometer, both operating at the L-band at the incidence (observation) angle $\theta = 40°$. The revisit time of SMAP is two to three days. After the failure of the SMAP radar in July 2015, a procedure was developed to replace the radar-based data from the SMAP radiometer

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

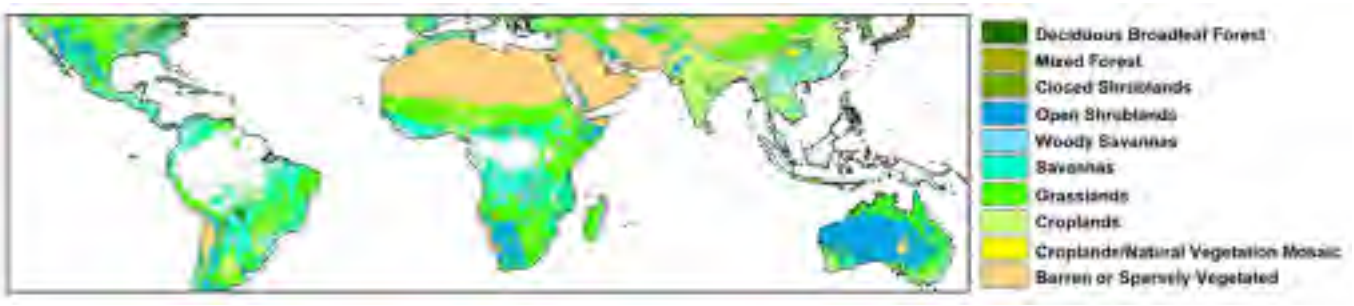

Fig. 1. Ten IGBP classes provided by MODIS in the year 2019 at a 5-km spatial resolution.

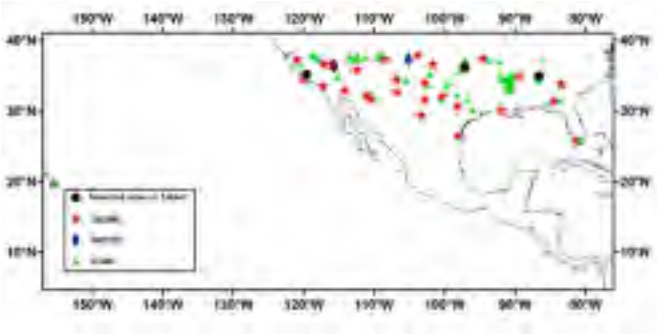

Fig. 2. Distribution of the 97 sites derived from the ISMN in 2019.

data [28], [29]. In this study, the 2017–2020 Level 3 36-km SMAP gridded data produced from the radiometer are used. Several parameters, i.e., the SM, and H and V polarization brightness temperatures, are used to build the model and to validate the CYGNSS SM results.

In this study, the "retrieval_qual_flag" variable in the SMAP product is used to identify retrievals to be of recommended quality. A "retrieval_qual_flag" value of either 0 or 8 indicates high-quality retrievals, with "0" indicating that "SM retrieval has recommended quality" and "8" indicating that "SM retrieval has recommended quality but with an unsuccessful freeze/thaw state retrieval" [30]. Subsequently, an averaging over three consecutive days is implemented to produce seamless globally covered brightness temperature data as inputs for the proposed model.

C. Land Cover Data

The International Geosphere-Biosphere Programme (IGBP) land classifications supplied by the MODIS MCD12Q3 product in 2019 with a 5-km spatial resolution are used in this study. As shown in Fig. 1, ten IGBP classes representing different vegetation distributions are selected to show the capability of the algorithm for SM retrieval on various land classifications.

D. ISMN Data In situ

SM measurements at the depths of 5 and 10 cm from 97 sites of the ISMN are used to validate the final CYGNSS-derived SM results (see Fig. 2). Sites with data point less than 50 for each year are not used. The hourly *in situ* SM data are averaged to derive daily values. Detailed information on the sites is summarized in the Appendix. All the data are downloaded from the ISMN website (http://ismn.geo.tuwien.ac.at/).

TABLE I
CHARACTERISTICS OF THE SELECTED SIX IN SITU SITES

Site Name	Lat, Lon(°)	Land Cover (IGBP)	VWC (kg m^{-2})	*Roughness* (*h*) parameter
CochoraRanch	35.12, -119.6	Grasslands	0.72	0.11
Stillwater-2-W	36.12, -97.09	Savannas	0.94	0.15
Stillwater-5-WNW	36.13, -97.11	Grasslands	0.94	0.15
AAMU-jtg	34.90, -86.60	Cropland/ Natural Vegetation Mosaic	4.49	0.13
BraggFarm	34.78, -86.55	Savannas	3.98	0.13
WTARS	34.90, -86.53	Cropland/ Natural Vegetation Mosaic	3.98	0.13

Six sites representing three examples of S-km grids collocated with different numbers of sites (i.e., one grid contains one, two, and three sites, respectively) are chosen to further analyze the temporal variations of the estimated SM (see Fig. 2). These sites also represent different climate conditions, roughness, and vegetation densities. The characteristics of the six *in situ* sites are listed in Table I. These sites are distributed across the contiguous United States (CONUS) from three different ground networks (i.e., the SCAN, SNOTEL, and USCRN), and they have different levels of vegetation cover as represented by VWC values of 0.72, 0.94, 0.94, 4.49, 3.98, and 3.98 kg $\cdot$ m^{-2}, respectively. The surface roughness (*h*) parameter varies from 0.11 to 0.15. The VWC and surface roughness of each site are determined as the mean SMAP value during the observed period of 2019. For validation of the CYGNSS-derived gridded SM, using *in situ* measurements from more than one site in the same grid can balance the difference of spatial scales between the two. The Stillwater-2-W and Stillwater-5-WNW are located in the same 36-km EASE-2 SMAP grid. The AAMU-jtg, BraggFarm, and WTARS are located in the same 36-km EASE-2 SMAP grid.

III. METHODS

The flowchart of the method in this study is shown in Fig. 3. The method consists of three steps.

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

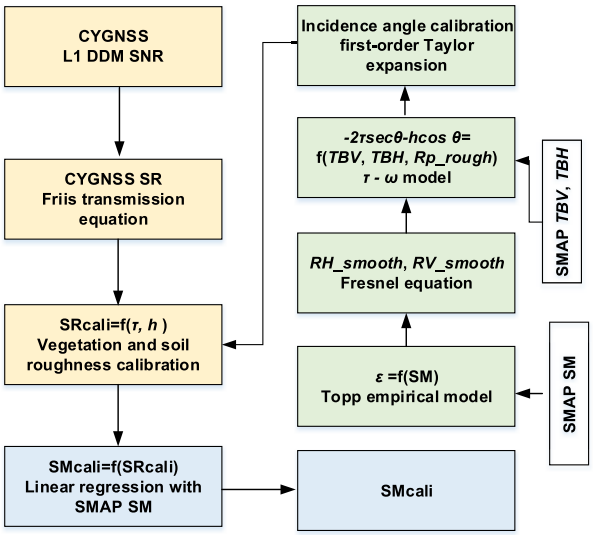

Fig. 3. Flowchart of the proposed algorithm.

1) Using the Friis transmission equation [see (1)] to estimate the CYGNSS SR and presenting the combined parameter (i.e., $\exp(-2\tau\sec\theta - h\cos^2\theta)$ [30] in (2), where τ is the VOD, h is the surface roughness parameter, and θ is the incident angle), which is used to derive the calibrated CYGNSS SR ($\mathrm{SR_{cali}}$).
2) Employing the radiative transfer model (i.e., the τ–ω model) and SMAP auxiliary data to derive the combined parameter presented in (2) [see (3)–(11)]. The SMAP data are collocated with CYGNSS data to derive a point-by-point calibration.
3) Calculating the resulting SM from CYGNSS $\mathrm{SR_{cali}}$ [see (12)].

A. Combined Parameter Used to Calibrate the CYGNSS SR

Suppose that the land reflections are dominated by the coherent component from the surface, as proposed in previous studies [17], and the CYGNSS observable used to estimate the SM is the SR. Based on the Friis transmission equation, the SR can be solved by converting all terms to dB as follows [27]:

$$\mathrm{SR(dB)} = 10\log P_r - 10\log P_t - 10\log G_t - 10\log G_r + 20\log(R_{\mathrm{ts}} + R_{\mathrm{sr}}) - 20\log\lambda + 20\log 4\pi \quad (1)$$

where P_r (dB, ddm_snr) is the peak value of the DDM minus the noise N, P_t (dB, gps_tx_power) is the transmitted RHCP power, R_{ts} (m, rx_to_sp_range) is the range from the transmitter to the specular reflection point, R_{sr} (m, tx_to_sp_range) is the range from the specular reflection point to the receiver, λ (m) is the wavelength of the GPS L1 band (19 cm), and G_t (dB, gps_ant_gain) and G_r (dB, sp_rx_gain) are the gain of the transmitter antenna and the gain of the receiver antenna, respectively. To produce SR values in the range that intuitively makes sense, we subtract 140 from the raw SR value derived in (1).

In most real situations, the original SR is a combined value of reflections from the rough surface of soil and vegetation [25], [26]. The original SR can be calibrated using the following expression:

$$\begin{aligned}\mathrm{SR} &= \mathrm{SR_{cali}} * \exp\left(-2\tau_p \sec\theta\right)\exp\left(-h\cos^2\theta\right)\\ &= \mathrm{SR_{cali}} * \exp\left(-2\tau_p \sec\theta - h\cos^2\theta\right)\end{aligned} \quad (2)$$

where $\exp(-2\tau_p\sec\theta)$ is the correction of vegetation and $\exp(-h\cos^2\theta)$ is the correction of surface roughness. τ_p is the VOD, h is the surface roughness parameter, and θ is the incidence angle. Thus, the combined parameter $\exp(-2\tau_p\sec\theta - h\cos^2\theta)$ can be used to calibrate the SR considering the effect of both surface roughness and vegetation at the same time. It should be noted that, in the practical application of (2), SR and $\mathrm{SR_{cali}}$ are in dB unit, and the combined parameter $\exp(-2\tau_p\sec\theta - h\cos^2\theta)$ should also be converted to dB.

B. Derivation of the Combined Parameter Using the Radiative Transfer Model

The zeroth-order radiative transfer model, known as the τ–ω model, is shown in (3). In this model, the brightness temperature T_{Bp} is described as the sum of three terms: 1) soil upward emission; 2) vegetation upward emission; and 3) the vegetation downward emission

$$T_{\mathrm{Bp}} = T_s e_p \gamma_p + T_s\left(1-\omega_p\right)\left(1-\gamma_p\right) + T_v\left(1-\omega_p\right)\left(1-\gamma_p\right)R_{p\_\mathrm{rough}}\gamma_p \quad (3)$$

where the subscript p represents the vertical or horizontal polarization, T_s and T_v are the soil effective temperature and the vegetation temperature, respectively, γ_p is the one-way vegetation transmissivity that can be expressed as $\gamma_p = \exp(-\tau_p\sec\theta)$, τ_p is the vegetation opacity depth, θ is the incidence angle, ω_p is the vegetation single scattering albedo, and R_{p_rough} is the rough soil reflectivity. R_{p_rough} is related to the emissivity (e_p) by $e_p = (1 - R_{p\_\mathrm{rough}})$ [30], [32]–[34].

Since the value of ω_p is relatively small at the L-band, in this study, ω_p is assumed to be zero for simplicity [30]. In addition, the soil and vegetation temperatures are assumed to be equal and represented as T. T_{Bp} in (3) can be simplified as

$$T_{\mathrm{Bp}} = T\left[1 - R_{p\_\mathrm{rough}}\exp\left(-2\tau_p\sec\theta\right)\right]. \quad (4)$$

The soil reflectivity R_{p_rough} is modeled as

$$R_{p\_\mathrm{rough}} = R_{p\_\mathrm{smooth}} * \exp\left(-h\cos^2\theta\right) \quad (5)$$

where p refers to polarization (for SMAP, V or H) and h parameterizes the intensity of the roughness effects.

By combining (4) and (5), the parameters τ_p and h can be combined as a single parameter as

$$T_{\mathrm{Bp}} = T\left[1 - R_{p\_\mathrm{smooth}} * \exp\left(-2\tau_p\sec\theta - h\cos^2\theta\right)\right]. \quad (6)$$

For SMAP data, at a given frequency with H and V polarizations, (6) can be expressed as [28]

$$\begin{aligned}T_{\mathrm{BH}} &= T\left[1 - R_{H\_\mathrm{smooth}} * \exp\left(-2\tau_p\sec\theta - h\cos^2\theta\right)\right]\\ T_{\mathrm{BV}} &= T\left[1 - R_{V\_\mathrm{smooth}} * \exp\left(-2\tau_p\sec\theta - h\cos^2\theta\right)\right].\end{aligned} \quad (7)$$

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

The combined parameter can be derived using (7) as

$$\exp\left(-2\tau_p \sec\theta - h\cos^2\theta\right) = \frac{T_{\mathrm{BV}} - T_{\mathrm{BH}}}{T_{\mathrm{BV}} R_{H\_\mathrm{smooth}} - T_{\mathrm{BH}} R_{V\_\mathrm{smooth}}}. \tag{8}$$

According to the Fresnel equations, R_{p_smooth} describes the performance of an electromagnetic wave at a smooth dielectric boundary, and it is only related to the soil complex dielectric constants at a specific incidence angle. R_{H_smooth} and R_{V_smooth} can be expressed as

$$\begin{aligned} R_{H\_\mathrm{smooth}} &= \left| \frac{\cos\theta - \sqrt{\varepsilon - \sin^2\theta}}{\cos\theta + \sqrt{\varepsilon - \sin^2\theta}} \right|^2 \\ R_{V\_\mathrm{smooth}} &= \left| \frac{\varepsilon\cos\theta - \sqrt{\varepsilon - \sin^2\theta}}{\varepsilon\cos\theta + \sqrt{\varepsilon - \sin^2\theta}} \right|^2 \end{aligned} \tag{9}$$

where ε is the soil complex dielectric constant. The Level 3 36-km gridded SMAP SM data are used to calculate ε using the Topp model [35] as

$$\mathrm{SM} = -0.053 + 0.0292\varepsilon - 0.00055\varepsilon^2 + 0.0000043\varepsilon^3. \tag{10}$$

Thus, by substituting (9) into (8), $\exp(-2\tau_p \sec\theta - h\cos^2\theta)$ [30] can be obtained. Finally, $\mathrm{SR_{cali}}$ can be calculated using (2). It should be noted that θ in (8) is the incidence angle of SMAP (i.e., 40°), which is not the same one as in (2) (i.e., the incidence angle of CYGNSS).

The SMAP provides a constant incidence angle of 40°, while the relevant incidence angle of CYGNSS varies from 5° to 35°. This, thus, leads to errors when substituting (8) into (2). To obtain the expected VOD and surface roughness calibration, the uncertainty of the incidence angle must be propagated. Here, a first-order Taylor expansion [36] for reducing the uncertainty from the incidence angle is used. The modified equation is

$$f(\theta_C) = f(\theta_S) + \left[\tan\theta_S \cdot \ln\left(f(\theta_S)\right) + \tan\theta_S \cdot 3h \cdot \cos^2\theta_S\right] \cdot (\theta_C - \theta_S) \cdot f(\theta_S) \tag{11}$$

where θ_S is the incidence angle of SMAP (40°), θ_C is the incidence angle of CYGNSS, and $f(\theta_S) = \exp(-2\tau_p \sec 40° - h\cos^2 40°)$. Here, the h value is set to 1.28 according to mode statistical. This method can eliminate the influence of the angle and minimize the impact of h changes.

C. SM Estimation From the Calibrated CYGNSS SR

$\mathrm{SR_{cali}}$ is converted to SM using the concept proposed in [27]. The algorithm is based on the linear regression between $\mathrm{SR_{cali}}$ and SMAP SM. The CYGNSS data points are spatially and temporally matched with the S-km grid. The descending SMAP data (6:00 A.M.)from one specific day are matched up with the CYGNSS data ranging from 6:00 P.M. of the previous day to 6:00 A.M. of the current day. The mean values of both $\mathrm{SR_{cali}}$ and SMAP for each grid over the two years are first calculated. For each grid, the maximum number of match-ups of CYGNSS $\mathrm{SR_{cali}}$ and SMAP SM is 254, and the average value is 128. Grids with less than ten observations are not used for calibration. In each SMAP grid over the two-year (2017–2018) period, the CYGNSS $\mathrm{SR_{cali}}$ and SMAP SM are correlated after having subtracted the mean values of both $\mathrm{SR_{cali}}$ and SMAP SM in that grid, and a linear regression is established between $\Delta\mathrm{SR_{cali}}$ and ΔSM for each grid calculated from the number of points in the calibrate period. The final calibrated SM ($\mathrm{SM_{cali}}$) can be expressed as

$$\Delta\mathrm{SR_{cali}} = \mathrm{SR_{cali}} - \mathrm{SR_{average}} \tag{12}$$

$$\mathrm{SM_{cali}} = \beta * \Delta\mathrm{SR_{cali}} + \alpha \tag{13}$$

where β is the slope of the linear regression for a particular grid and α is the mean SMAP SM during the entire time period. Note that α and β vary in each grid.

IV. Results

A. Improvement of the CYGNSS SR

Similar to the UCAR-SM ($\mathrm{SM_{UCAR}}$) product, the CYGNSS observations are upscaled to the 36-km SMAP EASE-2 grid resolution through averaging. Fig. 4(a) and (b) shows the SR and $\mathrm{SR_{cali}}$ against SMAP SM for each grid over the years 2019 and 2020, respectively. Fig. 4(c) illustrates the numerical distributions of R for SR and $\mathrm{SR_{cali}}$ corresponding to Fig. 4(a) and (b). Overall, it is obvious that, compared with R of SR against SMAP, R of $\mathrm{SR_{cali}}$ against SMAP SM is significantly improved to a higher level.

It is worth noting that the consistency varies over different regions. The R-value is low in the arid region [box in Fig. 4(b)]. This is due to the limited variability in SM throughout the observation period.

B. Evaluation of Daily SM Estimates

Fig. 5 shows the slope of the best fit linear regression between $\Delta\mathrm{SR_{cali}}$ and ΔSM of SMAP in 2019–2020. Overall, the spatial distribution of the slope varies significantly. The low slope values mainly occur in the regions with relatively low vegetation (blue box in Fig. 5, i.e., the Sahara Desert). The high slope values occur in the regions where the VWC values are relatively high, e.g., savannas and grasslands. Compared with Fig. 4(b), the slope can represent the sensitivity of $\mathrm{SR_{cali}}$ to SM with a lower slope occurring in with a lower R in Fig. 4(b) (blue box in Fig. 5), indicating lower sensitivity of $\mathrm{SR_{cali}}$ to SM and vice versa (black box in Fig. 5) [13].

The spatial distributions of the $\mathrm{SM_{UCAR}}$, $\mathrm{SM_{cali}}$, and SMAP SM on January 1, 2019, are shown in Fig. 6(a)–(c), respectively. The results of $\mathrm{SM_{UCAR}}$ minus SMAP SM and $\mathrm{SM_{cali}}$ minus SMAP SM are shown in Fig. 6(d) and (e), respectively. Both the $\mathrm{SM_{UCAR}}$ and $\mathrm{SM_{cali}}$ show good consistency of fluctuations with the SMAP SM. The absolute delta values tend to be lower in many arid regions (e.g., North Africa and Western Australia), with values generally varied less than 0.05 $\mathrm{m^3 m^{-3}}$, corresponding to the areas with the lowest SM. The delta value varies obviously in tropical rain forests (i.e., South Africa); insufficient open water masking or incoherent scattering due to vegetation may lead to this phenomenon [15]. In other places, such as the eastern USA and southern South America, the delta values vary relatively obvious with $-0.39 \sim 0.05$ $\mathrm{m^3 m^{-3}}$.

Fig. 7 shows the density plots of the $\mathrm{SM_{UCAR}}$ and $\mathrm{SM_{cali}}$ versus SMAP SM on the first day of each month in 2019. The color represents point density. Both $\mathrm{SM_{UCAR}}$ and $\mathrm{SM_{cali}}$

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

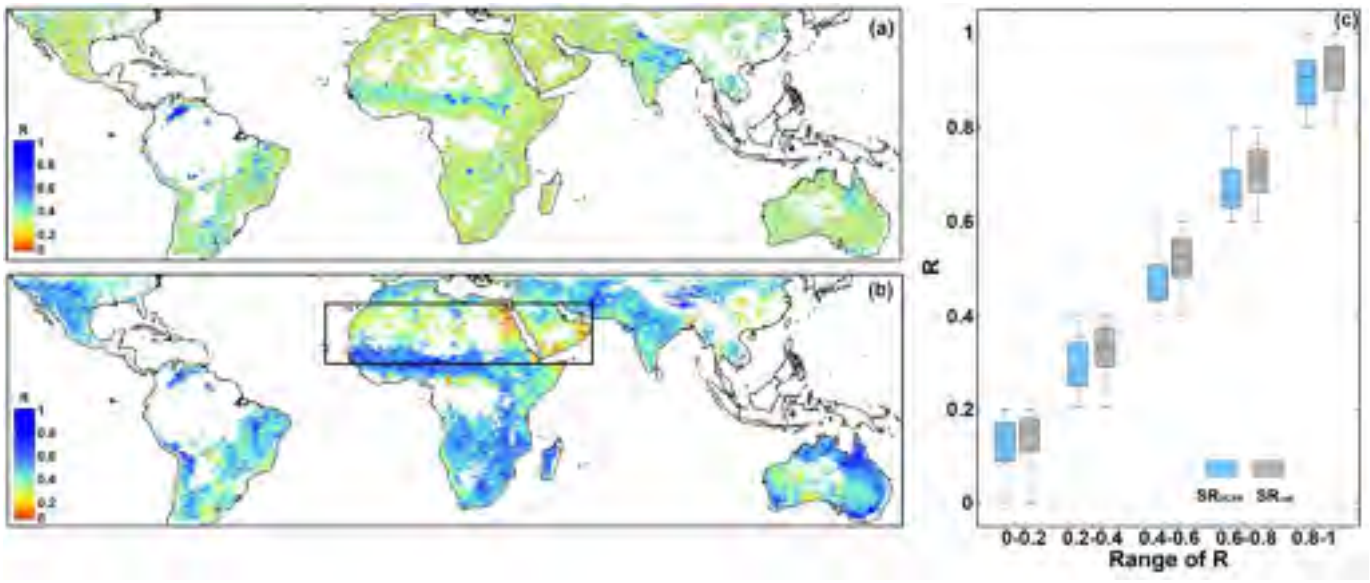

Fig. 4. Comparison of R values between SR (SR_{cali}) and SMAP. (a) R between SR and SMAP. (b) R between SR_{cali} and SMAP. (c) Boxplot of the R values in (a) and (b).

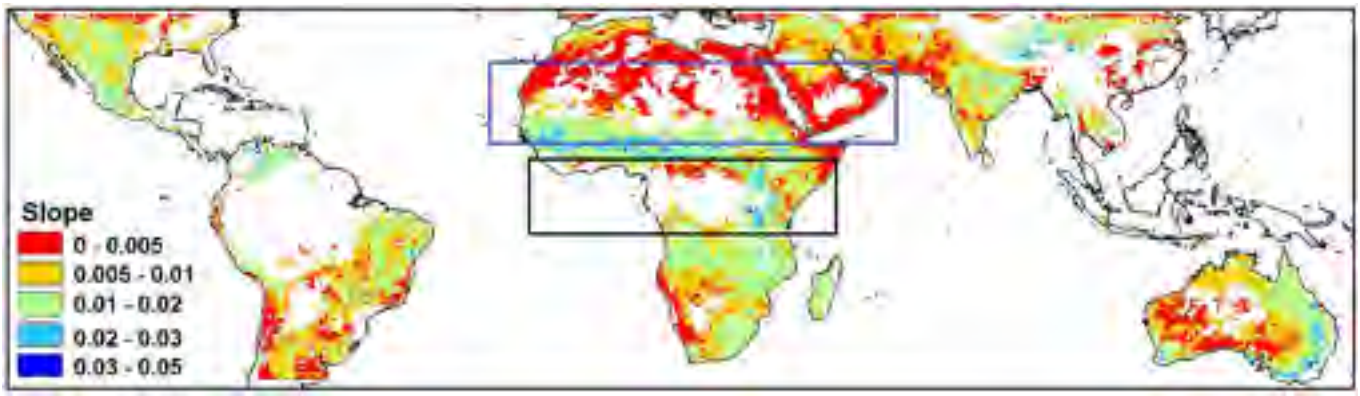

Fig. 5. Slope of the best fit linear regression between ΔSR_{cali} and ΔSM.

are close to the 1:1 line against SMAP SM. Specifically, the R-value varies from 0.732 to 0.909, respectively. The SM_{cali} also shows better root mean square error (RMSE) values than that of the SM_{UCAR} (RMSE varies from 0.051 to 0.078 m^3 m^{-3} versus 0.045 to 0.057 m^3 m^{-3}).

The average values of the three statistical indices [i.e., R, RMSE, and the mean absolute error (MAE)] per month in 2019 are shown in Table II. The SM_{cali} performs better than SM_{UCAR} with mean $R = 0.875$ versus 0.834, mean RMSE $= 0.050$ m^3m^{-3} versus 0.060 m^3m^{-3}, and mean MAE $= 0.041$ m^3m^{-3} versus 0.053 m^3m^{-3}. For SM_{cali}, the R-value is higher than 0.082, and RMSE and MAE values are less than 0.056 and 0.047 m^3m^{-3} over the entire time period, respectively.

TABLE II
R, RMSE, AND MAE VALUES FOR EACH MONTH IN 2019

	SM_{UCAR}			SM_{cali}		
	R	**RMSE (m^3 m^{-3})**	**MAE (m^3 m^{-3})**	**R**	**RMSE (m^3 m^{-3})**	**MAE (m^3 m^{-3})**
Jan.	0.789	0.070	0.062	0.820	0.053	0.044
Feb.	0.792	0.056	0.056	0.844	0.053	0.040
Mar.	0.847	0.051	0.047	0.863	0.050	0.047
Apr.	0.830	0.054	0.051	0.852	0.053	0.043
May.	0.841	0.056	0.052	0.921	0.049	0.039
Jun.	0.862	0.064	0.057	0.892	0.048	0.039
Jul.	0.823	0.060	0.052	0.876	0.053	0.041
Aug.	0.884	0.056	0.045	0.901	0.056	0.040
Sept.	0.862	0.063	0.051	0.891	0.052	0.039
Oct.	0.798	0.071	0.060	0.882	0.046	0.037
Nov.	0.834	0.058	0.049	0.879	0.045	0.037
Dec.	0.851	0.055	0.052	0.873	0.047	0.041
Mean	0.834	0.060	0.053	0.875	0.050	0.041

C. Evaluation of SM Estimates Over the Entire Year

For the SM estimates over January 2019 and November 2020, the spatial patterns of the three metrics (R, RMSE, and MAE) are shown in Fig. 8. The spatial coverages of Fig. 8(b1) and (c1) are different from that of Fig. 8(a1). It is because Fig. 8(a1) is derived using (1), the same algorithm to produce the SM_{UCAR}. In contrast, Fig. 8(b1) and (c1) are directly from the official SM_{UCAR}.

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

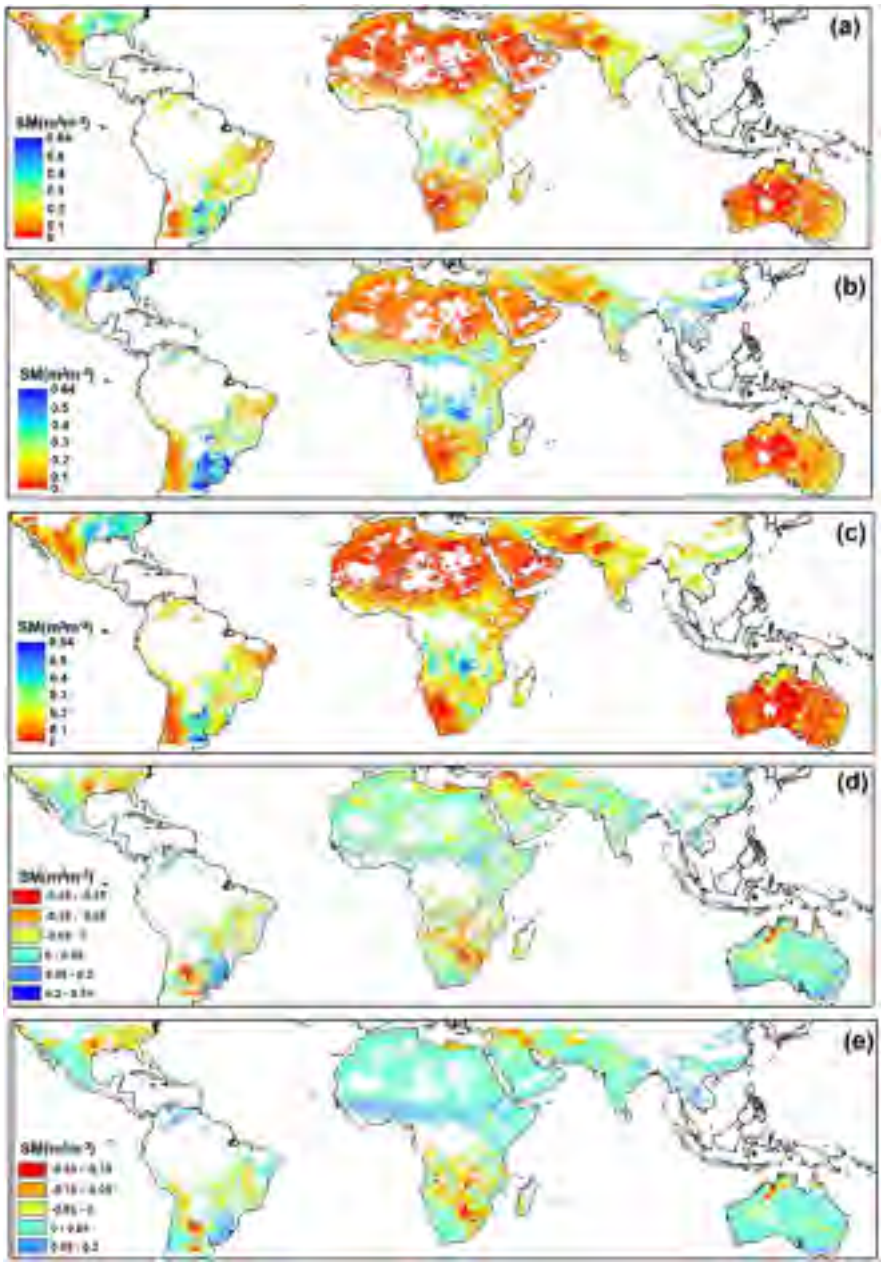

Fig. 6. Comparison of SM values on January 1, 2019. (a) SM_{UCAR}. (b) SM_{cali}. (c) SMAP SM (an averaging over January 1–3, 2019). (d) SM_{UCAR} minus SMAP SM. (e) SM_{cali} minus SMAP SM.

The difference is that the official SM_{UCAR} product filtered the raw CYGNSS data using some additional empirical quality controls besides the quality flags used in this study.

Generally, SM_{cali} performs better spatial patterns for all three indices than those of SM_{UCAR} ($R = 0.679$, RMSE $= 0.051$ m^3m^{-3}, and MAE $= 0.045$ m^3m^{-3} for SM_{cali}; $R = 0.565$, RMSE $= 0.067$ m^3m^{-3}, and MAE $= 0.051$ m^3m^{-3} for SM_{UCAR}). As shown in Fig. 8(a1) and (a2), lower R values occur in forests and desert regions. For SM_{cali}, over 70% of the R values are higher than 0.6, and the RMSE values are smaller than 0.07 m^3m^{-3} over most areas. The MAE values of SM_{cali} are smaller than 0.05 m^3m^{-3} over most areas of the globe.

Fig. 8(d) illustrates the numerical distributions of the three statistics for SM_{UCAR} and SM_{cali} at a daily scale in 2019 and 2020 [i.e., in Fig. 8(a1)–(c1) and (a2)–(c2)]. Generally, the R values of SM_{UCAR} are smaller than those of SM_{cali} [see Fig. 8(a)], and RMSE and MAE values of SM_{cali} are overall larger than those of SM_{UCAR} [see Fig. 8(b) and (c)], which results in better performances of the SM_{cali}. In addition, the RMSE and MAE of both SM_{UCAR} and SM_{cali} are smaller than 0.2 m^3m^{-3}.

D. Validation Using the in Situ Measurements

For each *in situ* site shown in Fig. 2 during the entire observation period, the CYGNSS observations being located

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

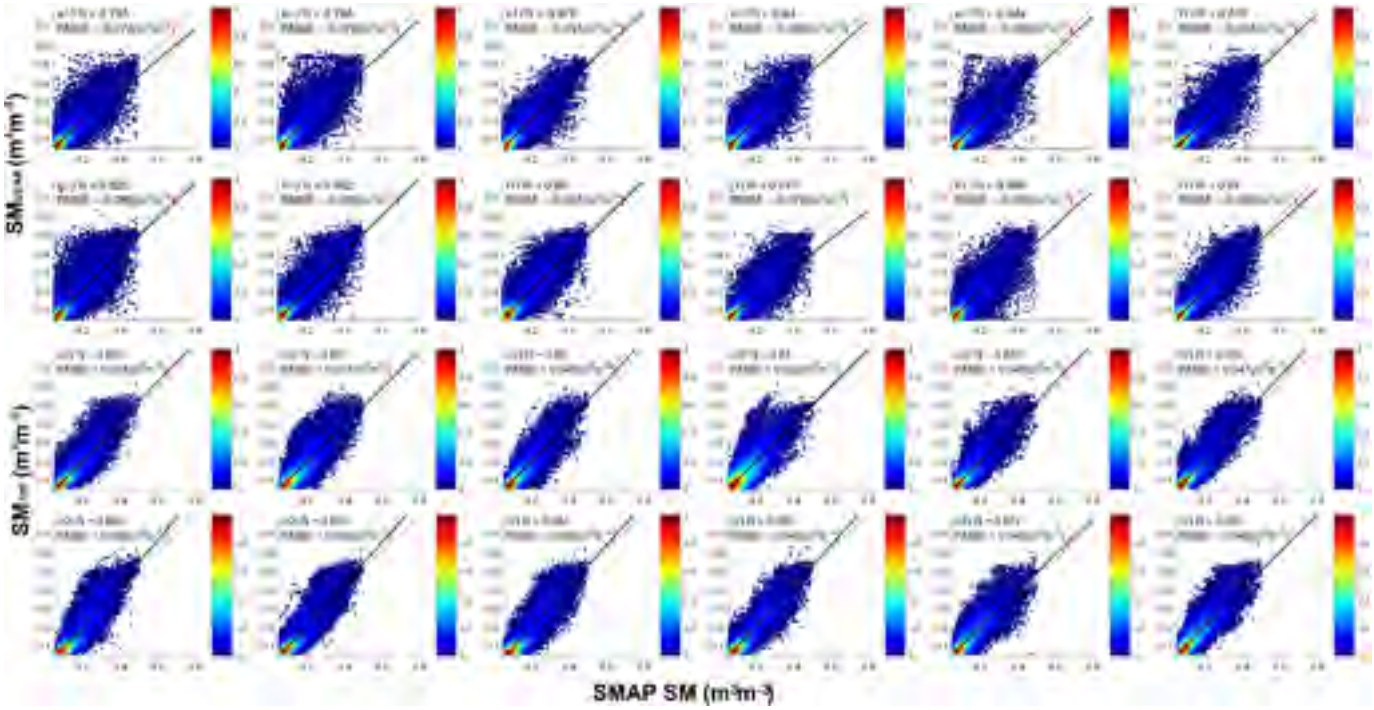

Fig. 7. Comparison of CYGNSS SM with SMAP SM on the first day of each month in 2019. (a1)–(i1) SM_{UCAR}. (a2)–(i2) SM_{cali}.

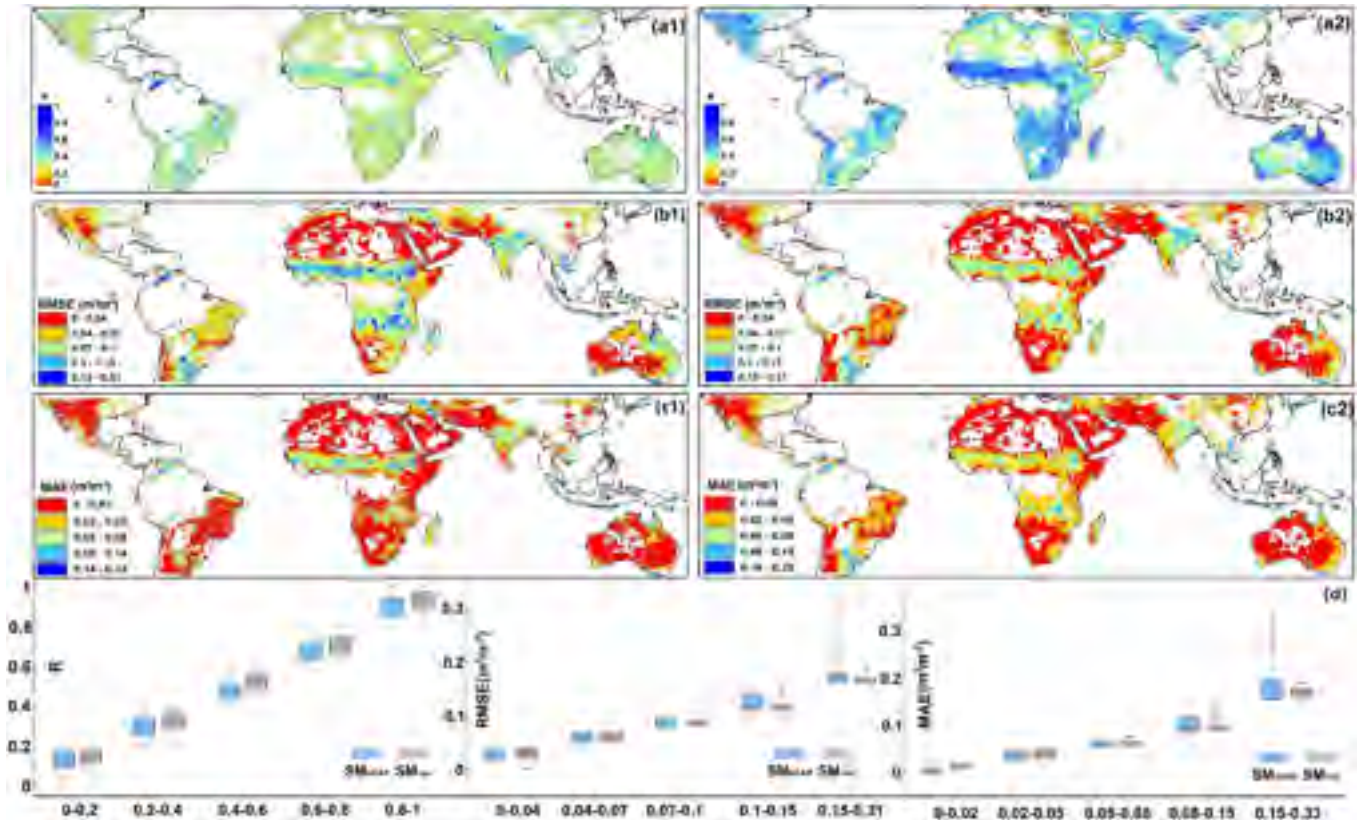

Fig. 8. Statistical indices for the SM derived from SM_{UCAR} (SM_{cali}) against the SMAP SM. (a1) and (a2) R. (b1) and (b2) RMSE. (c1) and (c2) MAE. (d) Boxplot of R, RMSE, and MAE.

less than 0.5 km from this site with VWC less than 5 kg · m^{-2} are selected. The SR_{cali} of these data points is correlated with the *in situ* SMs at the corresponding date. When more than one *in situ* point matches up with one SR_{cali}, a final value is calculated using the inverse distance weighting method considering the distances between the *in situ*

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

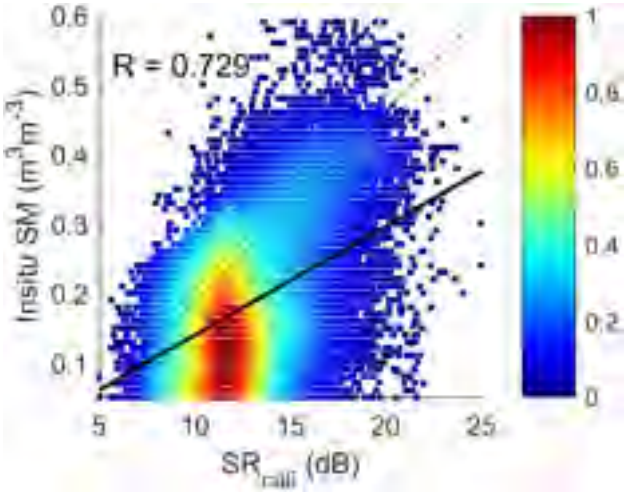

Fig. 9. Density plots of collocated SR_{cali} versus *in situ* SM from 97 ISMN sites.

measurements and the SR_{cali} point, rather than a simple average.

Using *in situ* data and km-scale satellite data for absolute error evaluation can be significantly biased [38], [39]. Therefore, for the ISMN network location, only R values are calculated here, rather than RMSE and MAE. The correlation results of SR_{cali} and the *in situ* SM are shown in Fig. 9. The correlation of SR_{cali} against the *in situ* SM is 0.729.

Because the CYGNSS SM is derived based on the 36-km EASE-2 gridded SMAP data as a reference, a significant mismatch of the spatial scale exists between the CYGNSS SM and the *in situ* points. Therefore, the comparison between the CYGNSS SM and the *in situ* SM of the 97 sites is not shown in this study. To evaluate the accuracy of CYGNSS SM compared to *in situ* measurements, Fig. 10 shows the comparisons of the gridded 36-km SM_{cali} and the *in situ* measurements for the six sites in Table I during 2019–2020. When one grid includes more than one site, different colored lines are used to indicate SM values from different sites.

Overall, good agreement is achieved between the two data even for the case of the relatively heavy vegetated site with R varying from 0.605 to 0.759. It is clear that SM_{cali} performs well in capturing the temporal dynamics of *in situ* SM, even for the case of the densely vegetated site. It is particularly noteworthy that, for the sites of AAMU-jtg, Bragg Farm, and WTARS, compared to the *in situ* SM, the SM_{cali} overestimates over half of the observed time. This is possibly due to the humid climate condition in this region. Previous researches show that SMAP usually overestimates SM in humid and cold regions, and this leads to the overestimation of SM_{cali} because it relies on the SMAP SM as a reference [30].

E. Improvement of the Temporal and Spatial Coverages Compared With SMAP

Although the SM estimation method of CYGNSS itself relies on SMAP SM data as a reference, the significant benefit of this topic is that it allows the combined usage of CYGNSS and SMAP to improve the temporal and spatial coverages of global SM. The improvement of the temporal and spatial coverages is shown in Fig. 11. From the perspective of the spatial distribution, the CYGNSS can increase the spatial coverage of the SMAP at an average value of 38.2% [see Fig. 11(a)].

The filled numbers of the daily CYGNSS observations against the daily SMAP SM product for each grid cell over the year 2019 are shown in Fig. 11(b). Except for the high-altitude regions (e.g., the Tibet Plateau) where there is a lack of valid data for CYGNSS, the daily CYGNSS can fill the gaps of SMAP with an average value of 113.7 days per 36-km grid with an increase by 31.6%.

V. Discussion

A. Contributions of the Surface Roughness

The SMAP brightness temperature can convey useful information about surface roughness. However, their monostatic mode of operation makes those estimates primarily sensitive to microwave-level roughness. In contrast, the CYGNSS's specular land measurements are excepted for macroroughness. Here, the contributions of the surface roughness are discussed based on two levels of comparisons:

First, the CYGNSS SR calibration results with and without the roughness correction are compared. Fig. 12(a) shows the results of SM with corrections of VOD and surface roughness, and Fig. 12(b) shows the results of SM with correction of VOD only. This experiment uses VOD and h parameters from SMAP rather than the physics-based model proposed in this study. The correlation coefficient has improved slightly when considering the surface roughness (from $R = 0.599$ to $R = 0.621$). This convinces two points: 1) the effect of vegetation on SM is much larger than that of surface roughness and 2) despite the limitation, the microroughness could calibrate the SR to some extent.

Second, the CYGNSS SR calibration results of using microwave level roughness and macroscale roughness are compared. Taken the Mainland of China as an example, the impacts of different scales of roughness are shown in Fig. 13. The macroroughness here is calculated as $H_{\text{macro}}(i) = (\text{std.curv}(i))^a$ [41]. Here, "i" is per pixel; *std.curv* is the second derivative of a surface; "a" determines the scale of macroroughness contributing toward total surface roughness; and the value of "a" is set to 2.

As shown in Fig. 13, the correlation coefficient between CYGNSS SR and SMAP SM has similar values with slight improvement when using macroroughness ($R = 0.54$ versus $R = 0.578$). It should be noted that the ability in total CYGNSS SR calibration is mainly determined by VOD, and the calibration of surface roughness on SM accuracy may be limited. In addition, compared to the surface roughness derived from SMAP, the macrosurface roughness is more suitable for CYGNSS SM calibration. Future research should consider the impact of the surface roughness scale.

B. Influences of Land Types

The R-values estimated from SR_{cali} against *in situ* according to Fig. 9 are classified into ten different land types derived

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

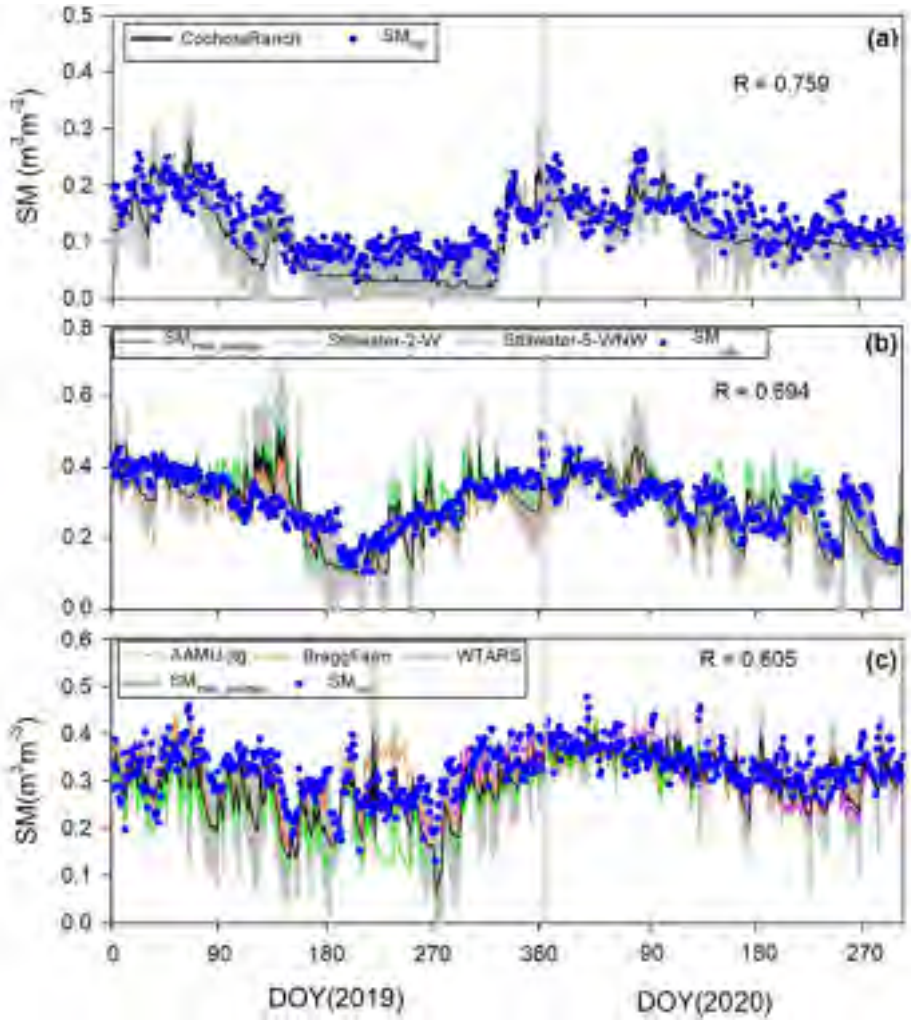

Fig. 10. Comparisons of the SM_{cali} and the *in situ* SM for six sites in the years 2019 and 2020. (a) CochoraRanch. (b) Stillwater-2-W and Stillwater-5-WNW. (c) AAMU-jtg, BraggFarm, and WTARS. The error bar is the absolute difference between the 5- and 10-cm SM values. DOY: day of year.

TABLE III
R-VALUE FOR DIFFERENT LAND TYPES

Land Type	R
Deciduous Broadleaf Forest	0.681
Mixed Forest	0.525
Closed Shrublands	0.703
Open Shrublands	0.652
Woody Savannas	0.768
Savannas	0.753
Grasslands	0.814
Croplands	0.789
Croplands/Natural Vegetation Mosaic	0.715
Barren or Sparsely Vegetated	0.558

from IBGP, as listed in Table III. The R-values are higher than 0.5 over the entire period time. L-bands have a common issue of "poor quality" when used in monitoring SM in dense vegetation and barren regions, and it appears to be very sensitive to regions with low vegetation height. In this study, SR_{cali} shows a weaker correlation with *in situ* measurements in densely forested regions than that in regions with less vegetation (e.g., shrublands and grassland). When comparing the similarity of Fig. 8 visually, the dependence of errors on different land types is obvious.

C. Advantages and Limitations

This study uses SMAP EASE-2 grids (36 km) to aggregate CYGNSS observations and exploit auxiliary data, such as SM, H, and V polarization brightness temperature. This is a limiting factor against the much higher potential of CYGNSS for spatiotemporal resolution. Also, because the proposed method uses a dielectric constant value, which was calculated from the SMAP SM product, the quality of the SMAP SM data will directly affect the quality of the current CYGNSS-based SM retrieval. However, the proposed algorithm can improve the SM estimation accuracy compared to the UCAR product, and it can also provide a new way to allow the combined usage of CYGNSS and SMAP to improve the temporal and spatial coverages of global SM.

Moreover, since most signals are probably a combination of incoherent and coherent scattering, here, we assume the

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

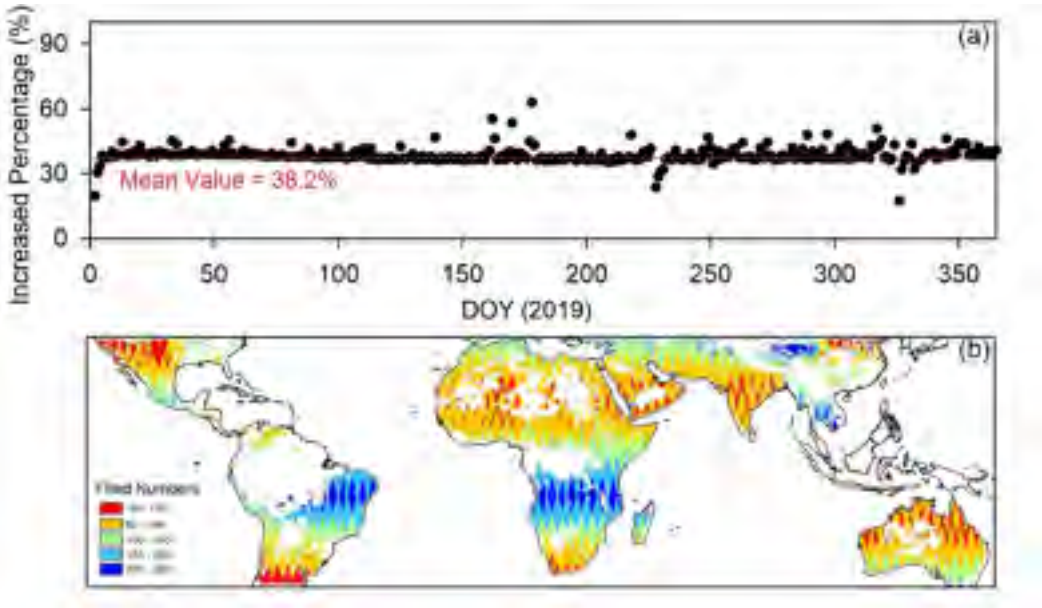

Fig. 11. Improvement of spatial and temporal coverages of CYGNSS compared with SMAP in the year 2019. (a) Increased percentage of the spatial coverage. (b) Filled numbers of the daily CYGNSS observations against the daily SMAP SM product for each grid cell.

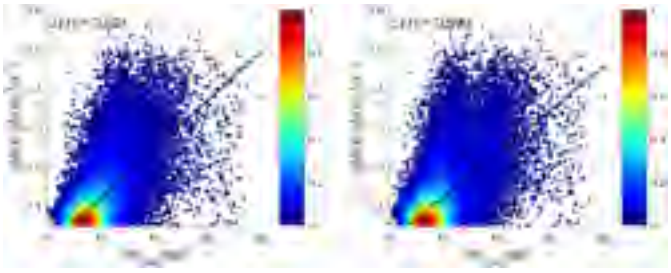

Fig. 12. Correlations of calibrated CYGNSS SR and SMAP on January 1, 2019. (a) SR calculated with corrections of VOD and surface roughness collected from SMAP. (b) SR calculated only with VOD collected from SMAP.

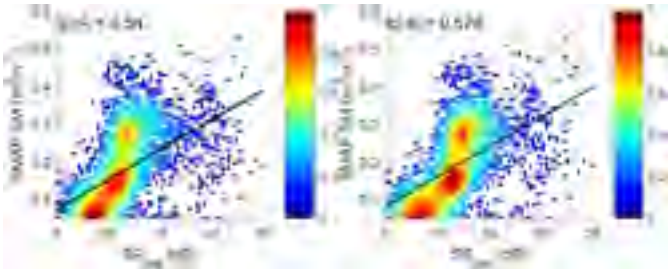

Fig. 13. Correlations of calibrated CYGNSS SR and SMAP on January 1, 2019, over Mainland China. (a) SR calculated with surface roughness collected from SMAP. (b) SR calculated with the macrosurface roughness.

reflected signals are coherent, which may lead the errors in final SM retrievals. Future work should consider the impact of both incoherent and coherent signals on SM estimation.

In addition, the Voronoi diagram is widely used in weighting the *in situ* measurements and upscale them for validation of the km-scale data [42]. However, sparse network measurements present challenges that compromise the evaluation of the performance in the absolute sense using the Voronoi diagram. Future work will consider the Voronoi diagram with a denser distribution of sites.

VI. Conclusion

To accurately estimate SM from CYGNSS data, this study explores a physics-based algorithm coupling CYGNSS SR and SMAP brightness temperature. This algorithm executes the calibration without relying on observations of h or vegetation biophysical parameters as inputs but with the SMAP brightness temperature as the only observations, which can provide an effective and efficient way to calibrate CYGNSS reflectivity from a new perspective. In addition, it provides a new way for the combined usage of CYGNSS and SMAP to improve the temporal and spatial coverages of global SM. The accuracy of the resulting SM shows good agreement with the SMAP SM ($R = 0.679$, RMSE $= 0.051$ m^{3}m^{-3}, and MAE $=$ 0.045 m^{3}m^{-3}) and with the *in situ* measurements ($R = 0.729$).

The high-quality CYGNSS SM derived in this study is expected to be used for a better understanding of the role of SM in the water and energy cycles. This study demonstrates that, as supplementary to SM-dedicated missions (e.g., SMAP and SMOS), GNSS-R missions, such as CYGNSS and future ones, could provide powerful data sources for producing global SM products with high accuracy and high temporal resolution. Furthermore, the current method of retrieving SM data from CYGNSS shows great potential for improving the quality of land surface models. It has long been thought that satellite-based SM data, when appropriately assimilated into land surface models, play a pivotal role in improving the quality of SM data [43]. Recently, Kim *et al.* [44] showed that CYGNSS-based SM data can significantly improve the quality of land surface models' SM outputs and can potentially improve the quality of other hydrometeorological variables,

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

TABLE IV
DETAILED INFORMATION FOR THE 97 ISMN SITES USED FOR VALIDATION

Site Name	Lat/Lon(°)	Land cover classes	VWC (kg m^{-2})	Roughness (*h*) parameter	Site Name	Lat/Lon(°)	Land cover classes	VWC (kg m^{-2})	Roughness (*h*) parameter
Charkiln	36.37/-115.83	Savannas	0.06	0.11	MonoclineR	36.54/-120.55	Grasslands	1.42	0.12
Stovepip	36.60/-117.14	Barren or Sparsely Vegetated	0.11	0.13	Fallbroo	33.44/-117.19	Closed Shrublands	1.51	0.10
LosLunasPm	34.77/-106.77	Croplands	0.14	0.13	Merced-2	37.24/-120.88	Croplands	1.54	0.12
Yuma-27-	32.84/-114.19	Barren or Sparsely Vegetated	0.15	0.11	SandyRidge	33.67/-90.57	Croplands	1.60	0.11
Essex	34.67/-115.17	Open Shrublands	0.18	0.12	SilverCity	33.08/-90.52	Croplands	1.67	0.11
Socorro-	34.36/-106.89	Grasslands	0.19	0.13	Vance	34.07/-90.35	Croplands	1.69	0.12
Las-Cruc	32.61/-106.74	Open Shrublands	0.19	0.11	SandHollow	37.10/-113.35	Open Shrublands	1.89	0.12
Williams	35.76/-112.34	Grasslands	0.22	0.15	UAPBEarle_	35.28/-90.45	Croplands	1.97	0.11
Mercury-	36.62/-116.02	Open Shrublands	0.23	0.11	UAPBMarian	34.78/-90.82	Croplands	2.01	0.11
Eastland	37.78/-109.17	Grasslands	0.26	0.16	PrairieVie	30.08/-95.98	Grasslands	2.01	0.14
Spooky	37.52/-111.27	Open Shrublands	0.27	0.12	Tunica	34.68/-90.42	Permanent Wetlands	2.08	0.12
Panther-	29.35/-103.21	Open Shrublands	0.28	0.11	NorthIssaq	33.00/-91.07	Croplands	2.09	0.11
DeepSpring	37.37/-117.97	Open Shrublands	0.31	0.12	MtVernon	37.07/-93.88	Croplands	2.14	0.14
CaveValley	37.35/-113.12	Grasslands	0.32	0.15	Joplin-2	37.43/-94.58	Croplands	2.17	0.12
La-Junta	37.86/-103.82	Grasslands	0.32	/	Cortez-8	37.26/-108.50	Grasslands	2.19	0.15
Monahans	31.62/-102.81	Grasslands	0.34	0.11	KemoleGulc	19.92/-155.58	Grasslands	2.33	0.15
AdamsRanch	34.25/-105.42	Grasslands	0.36	0.14	SilverSwor	19.77/-155.42	Open Shrublands	2.33	0.14
KyleCanyon	36.27/-115.62	Grasslands	0.37	0.12	WaimeaPlai	20.02/-155.60	Grasslands	2.33	0.15
LovellSumm	36.17/-115.62	Open Shrublands	0.37	0.11	DoeRidge	37.63/-118.83	Savannas	2.41	0.15
TroughSpri	36.37/-115.80	Woody Savannas	0.37	0.11	Everglades	25.50/-80.55	Grasslands	2.45	0.08
BRISTL	36.32/-115.70	Savannas	0.37	0.12	Lafayett	30.09/-91.87	Savannas	2.58	/
LEECAN	36.31/-115.68	Grasslands	0.37	0.12	BeasleyLak	33.38/-90.65	Croplands	2.68	0.11
RAINBO	36.25/-115.63	Grasslands	0.37	0.12	Mayday _	32.87/-90.52	Deciduous Broadleaf Forest	2.75	0.11
Enterprise	37.63/-113.65	Grasslands	0.38	0.15	Onward	32.75/-90.93	Grasslands	2.78	0.13
MarbleCree	37.78/-118.42	Open Shrublands	0.38	0.14	Panguitch	37.87/-112.43	Savannas	2.81	0.15
Muleshoe	33.96/-102.78	Grasslands	0.43	0.15	RiverRoadF	31.02/-85.03	Croplands	2.86	0.12
Bushland#1	35.17/-102.10	Grasslands	0.43	0.16	Scott	33.62/-91.10	Croplands	3.30	0.11
Tucson-1	32.24/-111.17	Open Shrublands	0.44	0.08	Holly-Sp	34.82/-89.43	Woody Savannas	3.36	0.13
WalnutGulc	31.73/-110.05	Grasslands	0.47	0.13	UAPBLonoke	34.85/-91.88	Woody Savannas	3.40	0.12
Goodwell	36.57/-101.61	Grasslands	0.50	0.16	Watkinsvil	33.88/-83.43	Savannas	3.49	0.12
Levelland	33.55/-102.37	Grasslands	0.61	0.13	Newton-8	31.31/-84.47	Croplands	3.57	0.12

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

TABLE IV
(Continued.) DETAILED INFORMATION FOR THE 97 ISMN SITES USED FOR VALIDATION

KnoxCity	33.45/-99.87	Grasslands	0.61	0.15	UAPBDewitt	34.28/-91.35	Croplands	3.71	0.12
Vernon	34.02/-99.25	Grasslands	0.61	0.16	LittleRive	31.50/-83.55	Cropland - Natural Vegetation Mosaic	3.75	0.13
Bronte-1	32.04/-100.25	Grasslands	0.62	0.16	JournaganR	37.00/-92.26	Savannas	3.90	0.13
SanAngelo	31.55/-100.51	Grasslands	0.66	0.15	Perthshire	33.97/-90.90	Deciduous Broadleaf Forest	3.95	0.12
CochoraRan	35.12/-119.60	Grasslands	0.72	0.11	BraggFarm	34.90/-86.60	Croplands	3.98	0.13
Elgin-5-	31.59/-110.51	Open Shrublands	0.85	0.13	McalisterF	35.07/-86.58	Croplands	3.98	0.13
Edinburg	26.53/-98.06	Grasslands	0.89	0.14	WTARS	34.90/-86.53	Croplands	3.98	0.13
WHISKE	37.21/-105.12	Grasslands	0.94	0.16	UAPBCampus	34.25/-92.03	Mixed Forests	4.05	0.13
Stillwat	36.12/-97.09	Grasslands	0.94	/	Watkinsv	33.78/-83.39	Savannas	4.09	0.14
Weslaco	26.16/-97.96	Urban Areas	0.95	0.10	AllenFarms	35.07/-86.90	Savannas	4.12	0.13
FortReno#1	35.55/-98.02	Grasslands	0.97	0.15	Cullman-NA	34.20/-86.80	Savannas	4.24	0.13
AlkaliMesa	37.67/-109.37	Grasslands	1.12	0.16	StanleyFar	34.43/-86.68	Deciduous Broadleaf Forest	4.24	0.13
MccrackenM	37.45/-109.33	Open Shrublands	1.12	0.14	AAMU-jtg	34.78/-86.55	Croplands	4.49	0.13
Vermillion	37.18/-112.20	Grasslands	1.13	0.13	MammothCav	37.18/-86.03	Deciduous Broadleaf Forest	4.50	0.13
Riesel _	31.48/-96.88	Grasslands	1.16	0.15	Everglad	25.90/-81.32	Permanent Wetlands	4.84	0.04
Stephenvil	32.25/-98.20	Grasslands	1.17	0.16	Stillwater-2-W	36.12/ -97.09	Savannas	0.94	0.15
Austin-3	30.62/-98.08	Savannas	1.23	0.15	Stillwater-5-WNW	36.13/ -97.11	Grasslands	0.94	0.15
Abrams	37.13/-97.08	Croplands	1.36	0.15					

such as energy flux, water flux, and carbon flux. Considering these findings, the superior quality of SM data from CYGNSS has great potential for improving land surface models.

APPENDIX

See Table IV.

ACKNOWLEDGMENT

The authors would like to thank the Cyclone Global Navigation Satellite System (CYGNSS), the Soil Moisture Active Passive (SMAP), the International Geosphere-Biosphere Programme (IGBP), and the International Soil Moisture Network (ISMN) teams for archiving and providing the data used in this study.

REFERENCES

[1] M. M. Al-Khaldi, J. T. Johnson, A. J. O'Brien, A. Balenzano, and F. Mattia, "Time-series retrieval of soil moisture using CYGNSS," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4322–4331, Jan. 2019.

[2] M. Owe, R. D. Jeu, and J. Walker, "A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 8, pp. 1643–1654, Aug. 2001.

[3] P. Rahimzadeh-Bajgiran, A. A. Berg, C. Champagne, and K. Omasa, "Estimation of soil moisture using optical/thermal infrared remote sensing in the Canadian prairies," *ISPRS J. Photogramm. Remote Sens.*, vol. 83, pp. 94–103, Sep. 2013.

[4] L. Brocca, L. Ciabatta, C. Massari, S. Camici, and A. Tarpanelli, "Soil moisture for hydrological applications: Open questions and new opportunities," *Water*, vol. 9, no. 2, p. 140, Feb. 2017.

[5] C. S. Draper, R. H. Reichle, G. J. M. De Lannoy, and Q. Liu, "Assimilation of passive and active microwave soil moisture retrievals," *Geophys. Res. Lett.*, vol. 39, no. 4, Feb. 2012, Art. no. L04401.

[6] G. Portal *et al.*, "A spatially consistent downscaling approach for SMOS using an adaptive moving window," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 6, pp. 1883–1894, Jun. 2018.

[7] L. Brocca *et al.*, "SM2RAIN-ASCAT (2007–2018): Global daily satellite rainfall data from ASCAT soil moisture observations," *Earth. Syst. Sci. Data.*, vol. 11, no. 4, pp. 1583–1601, Sep. 2019.

[8] M. Clarizia and C. S. Ruf, "Wind speed retrieval algorithm for the Cyclone Global Navigation Satellite System (CYGNSS) mission," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4419–4432, Aug. 2016.

[9] N. Rodriguez-Alvarez *et al.*, "Land geophysical parameters retrieval using the interference pattern GNSS-R technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 71–84, Jan. 2011.

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

[10] J. R. Piepmeier *et al.*, "Radio-frequency interference mitigation for the soil moisture active passive microwave radiometer," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 761–775, Jan. 2014.
[11] C. S. Ruf and R. Balasubramaniam, "Development of the CYGNSS geophysical model function for wind speed," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 1, pp. 66–77, Jan. 2019.
[12] J. Mashburn *et al.*, "Improved GNSS-R ocean surface altimetry with CYGNSS in the seas of Indonesia," *IEEE Trans. Geosci. Remote Sens.*, vol. 1, no. 17, pp. 1–17, Mar. 2020.
[13] W. Li, E. Cardellach, F. Fabra, S. Ribo, and A. Rius, "Assessment of spaceborne GNSS-R ocean altimetry performance using CYGNSS mission raw data," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 238–250, Jan. 2020.
[14] B. Liu, W. Wan, and Y. Hong, "Can the accuracy of sea surface salinity measurement be improved by incorporating spaceborne GNSS-reflectometry?" *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 1, pp. 3–7, Jan. 2021, doi: 10.1109/LGRS.2020.2967472.
[15] C. C. Chew and E. E. Small, "Soil moisture sensing using spaceborne GNSS reflections: Comparison of CYGNSS reflectivity to SMAP soil moisture," *Geophys. Res. Lett.*, vol. 45, no. 9, pp. 4049–4057, May 2018.
[16] X. Wu *et al.*, "First measurement of soil freeze/thaw cycles in the Tibetan plateau using CYGNSS GNSS-R data," *Remote Sens.*, vol. 12, no. 15, p. 2361, 2020.
[17] C. Chew, J. T. Reager, and E. Small, "CYGNSS data map flood inundation during the 2017 Atlantic hurricane season," *Sci. Rep.*, vol. 8, pp. 1–8, Jun. 2018.
[18] C. Chew and E. Small, "Estimating inundation extent using CYGNSS data: A conceptual modeling study," *Remote Sens. Environ.*, vol. 246, Sep. 2020, Art. no. 111869.
[19] W. Wan *et al.*, "Using CYGNSS data to monitor China's flood inundation during typhoon and extreme precipitation events in 2017," *Remote Sens.*, vol. 11, no. 7, p. 854, Apr. 2019.
[20] E. Loria, A. O'Brien, V. Zavorotny, B. Downs, and C. Zuffada, "Analysis of scattering characteristics from inland bodies of water observed by CYGNSS," *Remote Sens. Environ.*, vol. 245, Aug. 2020, Art. no. 111825.
[21] H. Carrenoluengo, G. Luzi, and M. Crosetto, "Above-ground biomass retrieval over tropical forests: A novel GNSS-R approach with CyGNSS," *Remote Sens.*, vol. 12, no. 9, p. 1369, Apr. 2020.
[22] T. Yang, W. Wan, Z. Sun, B. Liu, S. Li, and X. Chen, "Comprehensive evaluation of using TechDemoSat-1 and CYGNSS data to estimate soil moisture over mainland China," *Remote Sens.*, vol. 12, no. 11, p. 1699, May 2020.
[23] H. Kim and V. Lakshmi, "Use of cyclone global navigation satellite system (CyGNSS) observations for estimation of soil moisture," *Geophys. Res. Lett.*, vol. 45, no. 16, pp. 8272–8282, 2018.
[24] O. Eroglu, M. Kurum, D. Boyd, and A. C. Gurbuz, "High spatio-temporal resolution CYGNSS soil moisture estimates using artificial neural networks," *Remote Sens.*, vol. 11, no. 19, p. 2272, Sep. 2019.
[25] M. P. Clarizia, N. Pierdicca, F. Costantini, and N. Floury, "Analysis of CYGNSS data for soil moisture retrieval," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 7, pp. 2227–2235, Jul. 2019.
[26] S. H. Yueh, R. Shah, M. J. Chaubell, A. Hayashi, X. Xu, and A. Colliander, "A semiempirical modeling of soil moisture, vegetation, and surface roughness impact on CYGNSS reflectometry data," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–17, 2022, doi: 10.1109/TGRS.2020.3035989.
[27] C. Chew and E. Small, "Description of the UCAR/CU soil moisture product," *Remote Sens.*, vol. 12, no. 10, pp. 1558, May 2020.
[28] Q. Yan, W. Huang, S. Jin, and Y. Jia, "Pan-tropical soil moisture mapping based on a three-layer model from CYGNSS GNSS-R data," *Remote Sens. Environ.*, vol. 247, Sep. 2020, Art. no. 111944.
[29] E. Santi *et al.*, "Vegetation water content retrieval by means of multifrequency microwave acquisitions from AMSR2," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 9, pp. 3861–3873, Sep. 2017.
[30] O. Neill *et al.*, "Soil moisture active passive (SMAP) algorithm theoretical basis document: Level 2 & 3 soil moisture (Passive) data products," NASA Goddard Space Flight Center, Greenbelt, MD, USA, Tech. Rep. JPL D-66480, 2018.
[31] E. Loria, A. O'Brien, V. Zavorotny, B. Downs, and C. Zuffada, "Analysis of scattering characteristics from inland bodies of water observed by CYGNSS," *Remote Sens. Environ.*, vol. 245, Aug. 2020, Art. no. 111825.
[32] R. Panciera *et al.*, "The soil moisture active passive experiments (SMAPEx): Toward soil moisture retrieval from the SMAP mission," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 490–507, Jan. 2014.
[33] J. Li, "Accounting for unresolved clouds in a 1D infrared radiative transfer model—Part I: Solution for radiative transfer, including cloud scattering and overlap," *J. Atmos. Sci.*, vol. 59, no. 23, pp. 3302–3320, Dec. 2002.
[34] J. Y. Zeng, Z. Li, Q. Chen, and H. Y. Bi, "A simplified physically-based algorithm for surface soil moisture retrieval using AMSR-E data," *Frontiers Earth Sci.*, vol. 8, no. 3, pp. 427–438, Sep. 2014.
[35] G. C. Topp, J. L. Davis, and A. P. Annan, "Electromagnetic determination of soil water content: Measurements in coaxial transmission lines," *Water Resour. Res.*, vol. 16, no. 3, pp. 574–582, Jun. 1980.
[36] L. Xu, "A proportional differential control method for a time-delay system using the Taylor expansion approximation," *Appl. Math. Comput.*, vol. 236, pp. 391–399, Jun. 2014.
[37] R. Bindlish *et al.*, "GCOM-W AMSR2 soil moisture product validation using core validation sites," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 1, pp. 209–219, Jan. 2018.
[38] J. Dong *et al.*, "Comparison of microwave remote sensing and land surface modeling for surface soil moisture climatology estimation," *Remote Sens. Environ.*, vol. 242, Jun. 2020, Art. no. 111756, doi: 10.1016/j.rse.2020.111756.
[39] H. Kim *et al.*, "Global scale error assessments of soil moisture estimates from microwave-based active and passive satellites and land surface models over forest and mixed irrigated/dryland agriculture regions," *Remote Sens. Environ.*, vol. 251, Dec. 2020, Art. no. 112052.
[40] C. Ma, X. Li, L. Wei, and W. Wang, "Multi-scale validation of SMAP soil moisture products over cold and arid regions in northwestern China using distributed ground observation data," *Remote Sens.*, vol. 9, no. 4, p. 327, Apr. 2017.
[41] M. Neelam, A. Colliander, B. P. Mohanty, M. H. Cosh, S. Misra, and T. J. Jackson, "Multiscale surface roughness for improved soil moisture estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5264–5276, Aug. 2020.
[42] A. Colliander *et al.*, "Validation of SMAP surface soil moisture products with core validation sites," *Remote Sens. Environ.*, vol. 191, pp. 215–231, Mar. 2017.
[43] R. H. Reichle, W. T. Crow, and C. L. Keppenne, "An adaptive ensemble Kalman filter for soil moisture data assimilation," *Water Resour. Res.*, vol. 44, no. 3, Mar. 2008, Art. no. W03423.
[44] H. Kim, V. Lakshmi, Y. Kwon, and S. V. Kumar, "First attempt of global-scale assimilation of subdaily scale soil moisture estimates from CYGNSS and SMAP into a land surface model," *Environ. Res. Lett.*, vol. 16, no. 7, Jul. 2021, Art. no. 074041.

Ting Yang received the M.E. degree in cartography and geography information systems from the Aerospace Information Research Institute, Chinese Academy of Sciences (old name was the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences), Beijing, China, in 2014, and the Ph.D. degree in photogrammetry and remote sensing from Peking University, Beijing, in 2018.

Since 2019, she has been a Researcher with the CAS Engineering Laboratory for Yellow River Delta Modern Agriculture, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences. Her recent research interests include global navigation satellite system (GNSS)-reflectometry remote sensing and hydrological remote sensing.

Wei Wan received the M.E. degree in photogrammetry and remote sensing from Nanjing University, Nanjing, China, in 2010, and the Ph.D. degree in photogrammetry and remote sensing from Peking University, Beijing, China, in 2014.

She was a Visiting Scholar with the University of Colorado Boulder, Boulder, CO, USA, from 2012 to 2013. From 2014 to 2017, she was a Post-Doctoral Researcher with Tsinghua University, Beijing. Since 2017, she has been a Researcher with the Institute of Remote Sensing and GIS, Peking University. Her recent research interests include global navigation satellite system (GNSS)-reflectometry remote sensing and hydrological remote sensing.

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.

Jundong Wang received the M.S. degree in geographical information science from the Wuhan University of Technology, Wuhan, China, in 2020. He is currently pursuing the Ph.D. degree in ecology with the Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographical Sciences and Natural Resources Research, University of Chinese Academy of Sciences, Beijing, China.

His research interests include global navigation satellite system reflectometry (GNSS-R) application on soil salinity and GNSS signal processing.

Baojian Liu received the B.S. degree in geographic information science from the School of Earth and Space Science, Peking University, Beijing, China, in 2017, where he is currently pursuing the Ph.D. degree in photogrammetry and remote sensing.

His research interests include global navigation satellite system (GNSS)-reflectometry remote sensing and hydrological remote sensing.

Zhigang Sun received the M.E. degree from the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, in 2004, and the Ph.D. degree in environmental modeling and remote sensing from the University of Tsukuba, Tsukuba, Tsukuba, in 2008.

From 2009 to 2011, he was a Post-Doctoral Researcher with the University of Connecticut, Storrs, CT, USA. From 2011 to 2014, he was an Associate Researcher with the Center for Regional Environmental Research, Tsukuba. Since 2014, he has been a Professor with the CAS Engineering Laboratory for Yellow River Delta Modern Agriculture, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences. His recent research interests include multiscale observation, modeling, and evaluation of agroecosystem and big data applications in agriculture.

Authorized licensed use limited to: Peking University. Downloaded on May 28,2022 at 08:35:47 UTC from IEEE Xplore. Restrictions apply.