Journal of Global Change Data & Discovery. 2024, 8(2): 187-192
www.geodoi.ac.cn

Daily Soil Moisture Dataset Development Using CYGNSS in Southeastern China (Jan. 2019–Oct. 2020)

Yang, T.[1, 2]

1. 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;
2. The Shandong Dongying Institute of Geographic Sciences, Dongying 257000, China

Abstract: Remote sensing-based spatial distribution monitoring of large-scale soil moisture is of great significance to the management of agricultural production, and high-precision soil moisture counts can provide data support for regional-scale climate research. In this dataset, 18°N–38°N, 97°E–122°36'E in southeastern China was selected as the study area. The spaceborne GNSS-R belongs to the cross-disciplinary category of satellite navigation applications and remote sensing, and its working band L is sensitive to soil moisture changes, which provides a new technical means for large-scale soil moisture detection. This dataset is based on publicly released spaceborne GNSS-R data, i.e., CYGNSS data, to realize an effective calculation method for complex surface soil moisture, and to generate a dataset of soil moisture changes in southern China from January 2019 to October 2020. The dataset has a temporal resolution of daily and a spatial resolution of 0.36°x0.36°. The dataset includes the following data in the study area: (1) daily soil moisture in 2019; and (2) daily soil moisture from January to October in 2020. The dataset is archived in .tif and .mdd formats, and consists of 1,338 data files with data size of 40.1 MB (compressed to one file with 21.6 MB).

Keywords: spaceborne GNSS-R; CYGNSS; soil moisture; vegetation; roughness

DOI: https://doi.org/10.3974/geodp.2024.02.09

CSTR: https://cstr.escience.org.cn/CSTR:20146.14.2024.02.09

Dataset Availability Statement:

The dataset supporting this paper was published and is accessible through the *Digital Journal of Global Change Data Repository* at: https://doi.org/10.3974/geodb.2024.08.01.V1 or https://cstr.escience.org.cn/CSTR:20146.11.2024.08.01.V1.

1 Introduction

Soil moisture is a crucial variable in land-air interactions, which interacts with the Earth's

Received: 20-10-2023; **Accepted:** 10-03-2024; **Published:** 25-06-2024

Foundation: Natural Science Foundation of China (42101376)

***Corresponding Author:** Yang, T., Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, yangt@igsnrr.ac.cn

Data Citation: [1] Yang, T. Daily soil moisture dataset development using CYGNSS in southeastern China (Jan. 2019–Oct. 2020) [J]. *Journal of Global Change Data & Discovery*, 2024, 8(2): 187–192. https://doi.org/10.3974/geodp.2024.02.09. https://cstr.escience.org.cn/CSTR:20146.14.2024.02.09.

[2] Yang, T. Daily soil moisture dataset in southeastern China using CYGNSS (201901–202010) [J/DB/OL]. *Digital Journal of Global Change Data Repository*, 2024. https://doi.org/10.3974/geodb.2024.08.01.V1. https://cstr.escience.org.cn/CSTR:20146.11.2024.08.01.V1.

climate system by affecting land surface evapotranspiration, water transport, and carbon cycle, and has a significant impact on the climate system and its changes[1–3]. Currently, due to the complex spatial and temporal characteristics of soil moisture, the acquisition of a long-time series of soil moisture data with high precision is still very challenging. Global Navigation Satellite System Reflectometry (GNSS-R) is an emerging science and technology at the intersection of satellite navigation and remote sensing disciplines, and has been a hot spot of international research in recent years. This technology uses the long-term stable L-band signals transmitted by GNSS satellites as the signal source, which can be used for the inversion of oceanic (e.g., sea surface height, lake water level, sea surface wind speed) and land-related (e.g., soil moisture, vegetation water content, snow depth) parameters, realizing innovative value-added applications of navigation satellites. With its advantages, such as low cost, high time-frequency, and wide coverage, GNSS-R technology is now becoming one of the effective means for surface soil moisture estimation[4–8].

CYGNSS is one of the operational GNSS-R payloads currently operating in orbit and making data available simultaneously. With its high spatial and temporal resolution, wide coverage, and other advantages, CYGNSS has become an effective data source for soil water and salt parameter inversion. The mechanism of soil moisture inversion by CYGNSS is that its L-band is sensitive to the change of soil dielectric constant, and the signal-to-noise ratio (SNR) parameter based on the CYGNSS Delay Doppler Map (DDM) is obtained from the correlation power of the land surface reflected signal peaks. The Signal-to-Noise Ratio (SNR) parameter obtained based on the CYGNSS Delay Doppler Map (DDM), which is calculated from the correlation power of the peak of the reflected signal from the land surface, is greatly affected by the soil dielectric constant. From the data application point of view, CYGNSS global day-by-day soil moisture inversion products based on SNR inversion have been publicly released, but the effects of vegetation and roughness are neglected[9]. Based on the original CYGNSS data from October 2019 to 2020, this dataset improves the spaceborne GNSS-R soil moisture inversion method, develops the surface correction method of physical mechanism, and realizes the daily output of soil moisture in southeastern China, which is aimed to provide data support for climate change research and agricultural production management of southeast China.

2 Metadata of the Dataset

The metadata of the Daily soil moisture dataset in southeastern China using CYGNSS (201901–202010) [10] is summarized in Table 1. It includes the dataset's full name, short name, authors, year of the dataset, temporal resolution, spatial resolution, data format, data size, data files, data publisher, and data sharing policy, etc.

3 Methods

3.1 Data Processing

3.1.1 Derive of Surface Reflectivity

Here, the surface reflectivity (SR) is calculated using the bistatic radar equation, and the SR after vegetation and surface roughness correlation can be expressed as[8, 9]:

$$\begin{aligned}\sigma_0 &= \sigma_{cali} \cdot \exp\left(-2\tau \sec\theta\right) \cdot \exp(-h\cos^2\theta) \\ &= \sigma_{cali} \cdot \exp(-2\tau \sec\theta - h\cos^2\theta)\end{aligned} \tag{1}$$

where σ_0 represents the difference between DDM peak power and noise; σ_{cali} represents the SR; τ is the vegetation optical depth; h is the surface roughness; and θ represents the incidence angle. Here, exp(–2τsecθ–hcos2θ) is called the correction factor.

Table 1 Metadata summary of the daily soil moisture dataset in southeastern China using CYGNSS (201901–202010)

Items	Description
Dataset full name	Daily soil moisture dataset in southeastern China using CYGNSS (201901–202010)
Dataset short name	SM_SEChina201901-202010
Authors	Yang, T., Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, yangt@igsnrr.ac.cn
Geographical region	Southeastern China: 18°N–38°N, 97°E–122°36′E
Year	From January 2019 to October 2020
Temporal resolution	Daily
Spatial resolution	36 km
Data format	.tif, .mdd
Data size	40.1 MB (20.6 MB after compression)
Data files	(1) daily soil moisture in 2019; and (2) daily soil moisture from January to October, 2020
Foundation	National Natural Science Foundation of China (42101376)
Computing environment	Matlab, ArcGIS
Data publisher	Global Change Research Data Publishing & Repository, http://www.geodoi.ac.cn
Address	No. 11A, Datun Road, Chaoyang District, Beijing 100101, China
Data sharing policy	(1) *Data* are openly available and can be free downloaded via the Internet; (2) End users are encouraged to use *Data* subject to citation; (3) Users, who are by definition also value-added service providers, are welcome to redistribute *Data* subject to written permission from the GCdataPR Editorial Office and the issuance of a *Data* redistribution license; and (4) If *Data* are used to compile new datasets, the 'ten per cent principal' should be followed such that *Data* records utilized should not surpass 10% of the new dataset contents, while sources should be clearly noted in suitable places in the new dataset[11]
Communication and searchable system	DOI, CSTR, Crossref, DCI, CSCD, CNKI, SciEngine, WDS, GEOSS, PubScholar, CKRSC

3.1.2 Surface Reflectivity Correction

Here, a zero-order radiative transfer model (τ-ω model) is introduced to obtain the calculation of the correction factor. The algorithm relies on the bright temperature data from SMAP (Soil Moisture Active and Passive) as the only parameter input value. τ-ω model expresses the bright temperature data as an integrated value of the signals from the soil, vegetation, and vegetation reflected from the soil, expressed as follows[12]:

$$T_{Bp} = T[1 - R_{p\_rough} \exp(-2\tau \sec\theta)] \tag{2}$$

where T_{Bp} represents the SMAP bright temperature; the soil and vegetation temperatures are assumed to be equal, denoted as T; τ is the vegetation optical thickness; θ is the angle of incidence; and R_{p_rough} is the soil reflectivity of a rough surface, which can be expressed as the soil reflectivity of a smooth surface with a roughness influence factor:

$$R_{p\_rough} = R_{p\_smooth} \cdot \exp(-h\cos^2\theta) \tag{3}$$

where R_{p_smooth} is the soil reflectivity of the smooth surface.

The correlation factor can be expressed as:

$$\exp\left(-2\tau\sec\theta - h\cos^2\theta\right) = \frac{T_{BV} - T_{BH}}{T_{BV}R_{H\_smooth} - T_{BH}R_{V\_smooth}} \tag{4}$$

where T_{BV} and T_{BH} represent the V-polarized and H-polarized SMAP bright temperatures, respectively, and R_{H_smooth} and R_{V_smooth} refer to the Fresnel reflectance coefficients of the V-polarized and H-polarized smooth surfaces, respectively. R_{H_smooth} and R_{V_smooth} can be computed by using the Fresnel equation. Subsequently, to correct for the difference in incident angles between CYGNSS and SMAP, the incidence angle of CYGNSS is normalized, and finally, σ_{cali} is obtained using the correction factor brought into the Equation (1).

3.1.3 Soil Moisture Calculation

Finally, a linear empirical model is developed using SMAP soil moisture data with σ_{cali} to obtain CYGNSS soil moisture. Here, the daily CYGNSS data for 2021 matches the SMAP data to obtain the parameters of the empirical formula for each grid point, which is subsequently applied to the retrieval of the CYGNSS soil moisture data for May 2020[9]. The equation is as follows:

$$SM_{CYGNSS} = a \cdot \sigma_{cali} + b \quad (5)$$

where a and b represent the slope and intercept of each grid, respectively, and σ_{cali} is the CYGNSS SR.

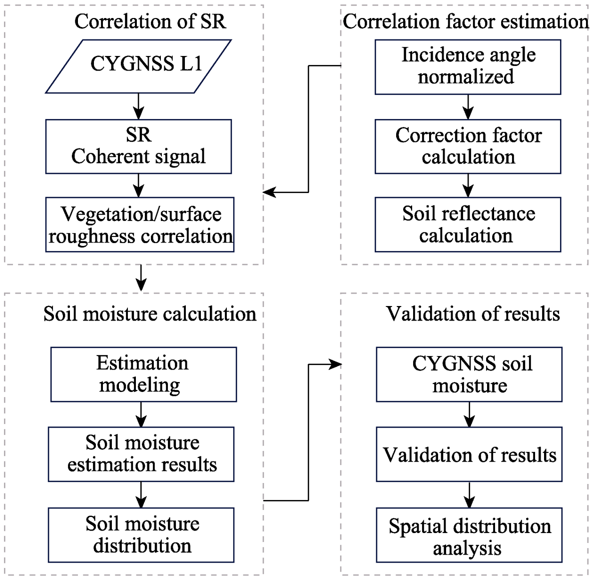

Figure 1 Flow chart of the dataset development

3.2 Technical Route

In the model training stage, the SR is first calculated using the CYGNSS coherent signals from August to December 2018; then, the SMAP bright temperature data is coupled to correct the vegetation and surface roughness errors based on the radiative transfer model; and finally, the soil moisture inversion model is built by linearly fitting with the SMAP soil moisture data from August to December 2018.

The data development process of this study (Figure 1) consists of four steps: 1) the model output stage, which uses the CYGNSS coherent signals from January 2019 to November 2020 to calculate the surface albedo; 2) coupling the SMAP bright temperature data to correct the vegetation and surface roughness errors based on the radiative transfer model using the SMAP soil moisture data for validation; and 3) using the model training stage of the linear fit to output CYGNSS soil moisture; 4) validating the result using SMAP soil moisture data from January 2019 to October 2020

4 Data Results and Validation

4.1 Data Composition

The data results consist of four data files, including:

(1) Daily soil moisture in 2019;

(2) Daily soil moisture from January to October 2020.

4.2 Data Products

4.2.1 Soil Moisture in Southeastern China

Based on the above algorithm and process, the spatial distribution of soil moisture from Jan. 2019 to Nov. 2020 in southeastern China is obtained with a spatial resolution of 36 km. In order to show the soil moisture distribution in southeastern China more clearly, Figure 2 shows the monthly average soil moisture in April, July, and October 2019 and January 2020. Soil moisture values are larger in the central region, with a range of variation between 0.3 and 0.5 $m^3 \cdot m^{-3}$, and smaller values in the north, with a range of variation between 0 and 0.2 $m^3 \cdot m^{-3}$. The overall trend is that soil moisture gradually increased from northward.

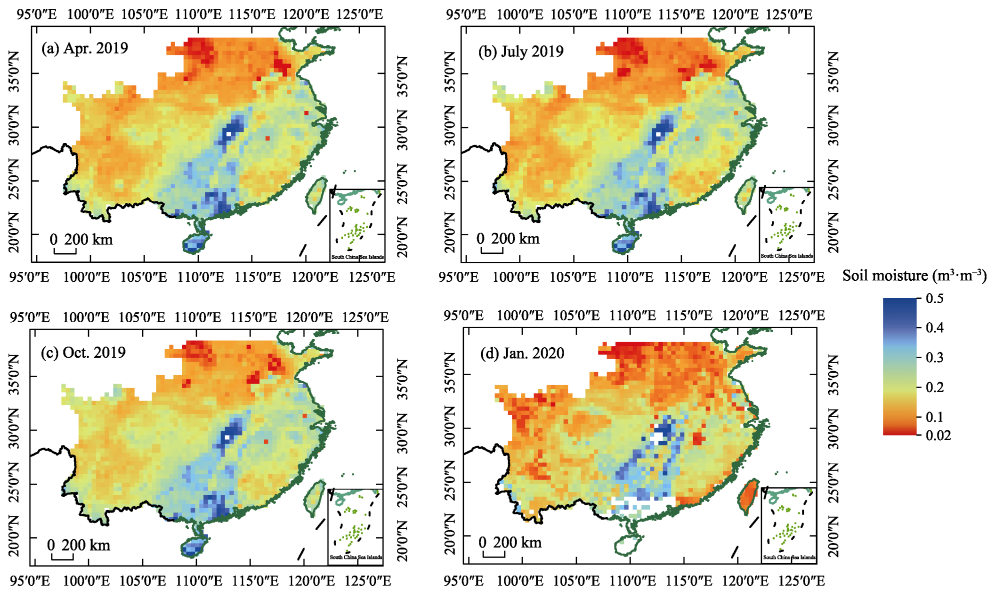

Figure 2 Soil moisture maps in southeastern China: (a) April 2019, (b) July 2019, (c) October 2019, (d) January 2020

4.3 Data Validation

Figure 3 shows the correlation coefficients (R) and ubRMSE of CYGNSS and SMAP soil moisture calculated on a daily basis from Jan. 2019 to Oct. 2020. Overall, CYGNSS soil moisture and SMAP soil moisture fit well, and the CYGNSS soil moisture is more closely aligned and numerically close to the SMAP soil moisture (i.e., R is 0.768; ubRMSE is 0.052 $m^3 \cdot m^{-3}$), suggesting that the methodology proposed in this study can be used to generate soil moisture products with high accuracy.

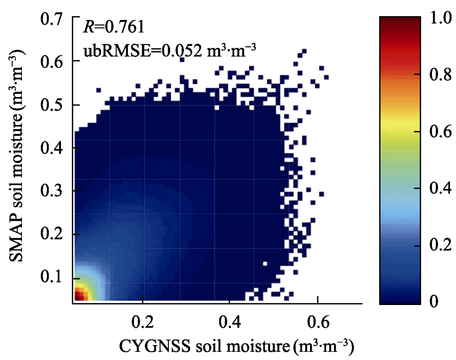

Figure 3 Validation of CYGNSS soil moisture based on SMAP soil moisture for 2020 in southeastern China

5 Discussion and Conclusion

The remote sensing technology of spaceborne GNSS-R is developing rapidly, and its theoretical system and methodological research have been gradually improved, but there are still some deficiencies. Regarding soil moisture retrieval, the authors have improved the inversion method of soil moisture. Meanwhile, distinguishing from the empirical or semi-empirical models, the authors combined the physical radiative transfer model to realize the correction of vegetation cover and surface roughness. The results show that soil moisture is highly spatially heterogeneous and varies significantly from month to month in southeastern China.

In terms of accuracy, the dataset provided by the authors has high accuracy, in which *R* can reach 0.768, and ubRMSE can reach 0.052 $m^3 \cdot m^{-3}$ compared with SMAP soil moisture. This dataset is expected to provide technical support for the improvement of the accuracy of remotely sensed soil moisture, and support the study of climate change in southeastern China and the management of agricultural production, which is of practical application value.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Shi, J. C., Du, Y., Du, J. Y., *et al*. Progresses on microwave remote sensing of land surface parameters [J]. *Science China*: *Earth Sciences*, 2012, 42(6): 814–842.

[2] Zhang, S. C., Guo, Q. Y., Ma, Z. M., *et al*. Research advances and some thoughts on soil moisture retrieval by space-borne GNSS-R [J]. Geomatics and information science of Wuhan University, 2013, 49(1): 15–26.

[3] Zeng, J. Y, Shi, P. F, Chen, K. S., *et al*. Assessment and error analysis of satellite soil moisture products over the third pole [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 60: 1–18.

[4] Yan, Q. Y., Jin, S. G., Huang, W. M., *et al*. Retrievals of soil moisture and vegetation optical depth using CyGNSS data [J]. *Journal of Nanjing University of Information Science & Technology* (*Natural Science Edition*), 2021, 13(2): 194–203.

[5] Yin, C. Soil Moisture Remote Sensing with Global Navigation Satellite System Reflectometry [D]. Nanjing: Nanjing University of Information Science and Technology, 2019.

[6] Zhang, Y., Zhang, D. D., Meng W. T., *et al*. Soil moisture retrieval using CYGNSS/SMAP data Fusion semi-empirical model [J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2023, 49(11): 2873–2882.

[7] Clarizia, M. P., Pierdicca, N., Costantini, F., *et al*. Analysis of CYGNSS data for soil moisture retrieval [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12(7): 2227–2235.

[8] Yang, T., Wan, W., Wang, J. D., *et al*. A physics-based algorithm to couple CYGNSS surface reflectivity and SMAP brightness temperature estimates for accurate soil moisture retrieval [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1–15.

[9] Chew, C., Small, E. Description of the UCAR/CU soil moisture product [J]. *Remote Sensing*, 2020, 12(10), 1558.

[10] Yang, T. Daily soil moisture dataset in Southeastern China using CYGNSS (201901–202010) [J/DB/OL]. *Digital Journal of Global Change Data Repository*, 2024. https://doi.org/10.3974/geodb.2024.08.01.V1. https://cstr.escience.org.cn/CSTR:20146.11.2024.08.01.V1.

[11] GCdataPR Editorial Office. GCdataPR data sharing policy [OL]. https://doi.org/10.3974/dp.policy.2014.05 (Updated 2017).

[12] O'Neill, P., Bindlish, R., Chan, S., *et al*. Algorithm Theoretical Basis Document. Level 2 & 3 Soil Moisture (Passive) Data Products [M]. Pasadena, CA: Jet Propulsion Laboratory, 2018.